Investigating the roles of gene dosage and stem cell maintenance in the regulation of plant shoot and inflorescence architecture

A Dissertation Presented

by

Katie Lynn Liberatore

National Science Foundation Graduate Research Fellow

and

Starr Foundation Fellow

to

The Watson School of Biological Sciences

at

Cold Spring Harbor Laboratory

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

in

Biological Sciences

Cold Spring Harbor Laboratory

December 2013

Table of contents

List of figures and tables

Figures

Tables

Abbreviations

CAPS	cleaved amplified polymorphic sequences
cDNA	complementary deoxyribonucleic acid
CTAB	cetrimonium bromide
CZ	central zone
DAG	days after germination
dCAPS	derived cleaved amplified polymorphic sequences
DNA	deoxyribonucleic acid
EVM	early vegetative meristem
FM	floral meristem
IF	inflorescence
IM	inflorescence meristem
Indel	insertion-deletion
LOD	logarithm of odds
LVM	late vegetative meristem
mRNA	messenger ribonucleic acid
MVM	middle vegetative meristem
NIL	near isogenic line
OC	organizing center
PCR	polymerase chain reaction
PSM	primary shoot meristem
PZ	peripheral zone
qPCR	quantitative PCR
QTL	quantitative trait locus
RAM	root apical meristem
RIL	recombinant inbred line
RNA	ribonucleic acid
RNA-seq	RNA sequencing
RT-PCR	reverse transcription PCR
SAM	shoot apical meristem
SIM	sympodial inflorescence meristem
SSR	simple sequence repeats
SYM	sympodial shoot meristem
T-DNA	transfer DNA
TM	transition meristem

Acknowledgements

I have had an amazing time at Cold Spring Harbor Laboratory during the past several years and am humbled to have met, learned from, and worked with so many wonderful scientists. First, I am thankful to the Watson School of Biological Sciences, including the current and former deans Alex Gann and Leemor Joshua-Tor, executive committee members, and course instructors for directing the school, training me to think like a scientist, and giving me the opportunity to earn my doctoral degree at Cold Spring Harbor Laboratory. To the rest of the current and former administration, Alyson Kass-Eisler, Kim Geer, Kim Creteur, Carrie Cowan, Dawn Meehan Pologruto, Keisha John and Uwe Hilgert; from the little conversations and treats to organizing classes and scheduling committee meetings, thank you for everything you have done that has made it easy and enjoyable to focus on my studies. To my fantastic group of classmates (Beth, Carrie, Mel, Zina, Jiahao, Mitch, Dario, Philippe, Felix, Nil, Ozlem, Sang, Elvin and Susan) - I am happy to have you as my colleagues and friends.

To my thesis committee (Dave Jackson, Marja Timmermans, Doreen Ware, Adrian Krainer) - I have learned a lot from our conversations and am grateful for your time and suggestions that helped shape this work. Thanks also to my external examiner Rüdiger Simon for the time and effort spent reviewing this work and for traveling a long way for the defense.

To Zach, I could not have imagined a better Ph.D. advisor. I am continually inspired by your intense passion for science and I am honored to have had the opportunity to work in your lab and to learn from you. Thank you for your unwavering support and encouragement, and for constantly challenging me to improve myself.

Much thanks to all past and present members of the Lippman lab. It has been a pleasure and honor to work with and learn from you all. Special thanks to Cora MacAlister for collaboration with *FIN*, to Ke Jiang for the collaboration on the heterosis project, Soon-Ju Park for lots of general help and conversations and Sophie Thomain and Cao Xu for great conversation

and recent collaboration on the *FIN* story. My projects were phenotyping intensive and would not have been possible without help from everyone in the field. I am also very thankful for the plant care provided by Tim Mulligan, Paul Hanlon, Sharron Zhou and field and greenhouse assistance from Andrew Krainer and Will Ware. Many thanks also to Junko Hige, Maya Strahl, Jason Wong and Christopher Brooks for technical assistance along the way and for keeping the lab running smoothly, and to J. Van Eck and colleagues for generating transgenic plants. Furthermore, to the Delbrück community, thanks for providing a great environment to work in and for conversation, encouragement, and help along the way. Special thanks to Marja and the rest of the Timmermans lab for an introduction to the beauty of plant development during my first rotation. Much thanks to Damianos, Marie, Fang, and Sasha who taught me techniques crucial to this work. Thanks also to Mike, Aman, Eugene, and Joe for fruitful discussions, moral support, and fun throughout the journey. To many wonderful friends not already named who have made this journey fun: Jessica, Jeremy, Austin, Shane, Leah, Anna, Paloma, Felix, Ian, Kaja, Fred, Amy, Roberto, and Colleen. Thanks for good food, conversation, and laughs along the way.

I am also grateful for a suite of stellar science teachers and advisors who fueled my interest in science early on. Special thanks to my undergraduate advisors Stephanie W. Ruby and Ursula Shepherd who provided early training and encouraged me to pursue my Ph.D.

To my parents, thank you for always believing in me and encouraging me to follow my passions. I could not ask for more loving and supportive parents and I appreciate everything you have done for me along the way. To my Sissy, thanks for being my friend and partner-in-crime for all these years and for always making me laugh. Thanks also to my Kramer family and the rest of my extended family for continued encouragement and support.

Finally, to my amazing husband, best friend, and thesis formatting guru, Brian. From coffee in the Jemez to bagels and *cawfee* on the harbor, it has been an amazing ride. Thanks for everything, including keeping me well fed and warmed by the fire while writing, anniversaries spent in the field, and to you and the bugs for reminding me to stop and smell the roses.

Abstract

Plants develop a striking diversity of architecture types including variations in branching of the vegetative shoots and branching within the flower bearing branches called inflorescences. The work presented in this dissertation focuses on mechanisms underlying these diverse architectures. In the first results chapter, I focus on the controversial theory of single-gene overdominance as a mechanism for driving hybrid vigor/heterosis involving a flowering hormone called florigen. I show that while a dosage effect on plant growth is conserved across divergent species, the yield effect is dependent on the sympodial growth habit of tomato and does not translate to the monopodial plant *Arabidopsis*. In the latter chapters, I focus on the characterization of genetic factors underlying inflorescence branching variants in tomato, specifically those that have increased branching. This work includes QTL mapping in *bifurcating inflorescence* (*bif*), which maps to a region containing *COMPOUND INFLORESCENCE* (*S*), mutants in which are known to delay meristem maturation during the transition to flowering and thus alter inflorescence architecture. I hypothesize that regulatory changes decrease the dosage of *S* to modulate branching. Finally, I present detailed characterization of two mutants that develop branched inflorescences with extra floral organs and enlarged flowers called, *fasciated and branched* (*fab*) and *fasciated inflorescence* (*fin*). Both mutants have enlarged stem cell populations marked by expanded expression of the stem cell promoting factor, *WUSCHEL (WUS)*, providing a link between the control of meristem size and inflorescence architecture. However, while *FAB* encodes the tomato ortholog of CLAVATA1 and is the first member of the CLV-WUS meristem maintenance pathway to be identified in tomato, *FIN* encodes a transmembrane protein of unknown function, adding a novel factor to the suite of plant stem cell regulators. Combined, my work demonstrates that multiple mechanisms, from modulation of plant growth habit, to meristem maturation and maintenance, converge on the regulation of shoot branching.

1 Introduction

Introduction

As sessile organisms, plants are at the mercy of the environment in which they are born and therefore, they have evolved a number of clever mechanisms to maximize their reproductive success. For instance, plants have adapted to different ecological niches and have developed coping mechanisms for tolerating biotic and abiotic stress in uncertain and ever-changing environments. In addition, plants have an enormous range in life cycles ranging from those that have simple body plans and transition to flowering once in their lifetime, like sunflower, to plants like trees that live many years, generate complex branching systems and reiterate vegetative and reproductive growth each season. Beyond the wide range of overall plant size and the number and arrangement of vegetative shoots they can produce, flowering plants also display a tremendous diversity in the number and arrangement of flowers developed on reproductive shoots called inflorescences. It is this amazing developmental plasticity to renew growth and the diversity of architecture types that particularly caught my attention and is a main focus of this dissertation.

The ability to continually generate organs and the magnificent array of morphological diversity this affords the plant is dependent on meristems, stem cell reservoirs at the growing tips of the plant that direct all post-embryonic growth (reviewed in Barton 2010; Ha et al. 2010; Stahl and Simon 2010; Aichinger et al. 2012). Two main meristems direct belowground and aboveground apical growth, the root apical meristem (RAM) and the shoot apical meristem (SAM), respectfully, the latter of which will be the focus herein. The SAM is a dynamic microenvironment responsible for integrating endogenous signals with environmental cues to determine the timing and frequency of lateral organ formation. For continuous growth, the SAM is maintained under tight regulation via overlapping signaling pathways to strike a fine balance between stem cell self-renewal and differentiation of cells to the flanks of the meristem. These daughter cells go on to develop the shoot and lateral organs in two phases of growth, first shoots and leaves during the vegetative stage, and then flowers upon the reproductive transition. While

much progress has been made to uncover the genetic and molecular mechanistic basis for specification and development of particular organ types such as leaves and flowers, much less is known about the dynamics of the transition from vegetative to reproductive growth and the mechanisms governing plant architecture including shoot and inflorescence branching. The combination of vegetative shoot branching and branching within the inflorescence influences how the plant interacts with biotic and abiotic factors in the environment, which directly impacts reproductive success in many facets including pollination rate, fruit and seed production and the range of seed dispersal (Wyatt 1982).

Investigating the processes by which plants regulate their growth, including shoot and inflorescence architecture, is interesting from both a basic science perspective of understanding fundamental mechanisms in plant development and from an applied perspective of how these processes may be harnessed for increasing crop yield. Fundamental questions include: How do plants decide when, where and how often to make new organs? What are the mechanisms that regulate how shoots and organs are arranged? What are the possible variations in architecture types that can be formed? What are the global molecular dynamics underlying these processes? What are the major contributing genetic factors? How do changes in gene expression patterns impact these developmental processes? How conserved are the mechanisms governing these processes across diverse species? How can these processes be harnessed and fine-tuned for improving agriculture?

Focusing on the model systems *Arabidopsis* and tomato, I explored these questions from two main angles. First, I considered the controversial theory of single-gene overdominance in driving hybrid vigor/heterosis in *Arabidopsis*, which relates to how changes in gene dosage for a single gene might change plant growth habit, shoot and inflorescence architecture and overall yield. Second, I explored tomato variants with mild to moderately branched inflorescences in the search for novel genes and pathways involved in regulating inflorescence architecture. Characterization of these branching variants uncovered links to mechanisms underlying the

control of meristem size and the rate of meristem maturation. Therefore, an understanding of meristem form and function, meristem maintenance and maturation, the floral transition, plant growth habit and inflorescence architecture is indispensible background to this work. This chapter provides a brief introduction to these topics and ends with a summary of the specific areas I pursued and the major outcomes. Each results chapter provides a more specific introduction that expands on these topics as pertinent to the questions addressed in each section. A detailed review of hybrid vigor and the theory of single-gene overdominance in driving heterosis can be found in Appendix (Liberatore et al. 2013, see Appendix A) and a summary of these concepts are also provided in the introduction to Chapter 2, and therefore, is not included in this introduction.

The shoot apical meristem: organization, maintenance and maturation

All post-embryonic growth arises from meristems, pluripotent stem cell niches at the growing tips of the plant (reviewed in Barton 2010; Ha et al. 2010; Stahl and Simon 2010; Aichinger et al. 2012). Achieving a fine balance between stem cell self-renewal and cell differentiation, not to mention deciding when and what type of organs to form, is a complex task. Cellular organization combined with intercellular signaling directs both maintenance of the stem cell niche and maturation to allow for formation of lateral organs and proliferation of shoots. Here I describe the *Arabidopsis* meristem as the stereotypical SAM – subtle differences are apparent between dicotyledonous plants like *Arabidopsis* and monocotyledonous plants such as rice and maize however, the basic concepts are conserved across this divide.

Aboveground growth is initiated from the vegetative shoot apical meristem (SAM), which appears as a small dome of cells at the shoot apex (reviewed in Barton 2010; Ha et al. 2010; Stahl and Simon 2010; Aichinger et al. 2012). Plant cells are not motile and therefore, the organization and signaling between individual cells and amongst populations of cells is crucial for proper development. The SAM is organized in distinct layers and sub-domains each with a

particular function. First, the cells are formed in three clonally distinct layers L1-L3 (Figure 1.1A). The L1 and L2 each form a single layer of cells that propagate by anticlinal divisions towards the periphery of the meristem and will form the epidermis and sub-epidermis of developing organs, respectfully, while the underlying L3 cells can divide in all directions to form the internal tissues. Second, cells within the meristem can be categorized into zones defined by characteristics of the cells themselves and by functional properties of the domain (Figure 1.1B). The stem cells are situated at the apex of the meristem in a region called the central zone (CZ). These cells divide slowly to both replenish themselves and provide cells for developing organs. Daughter cells that are pushed out of the central zone and towards the flanks of the meristem begin to divide more rapidly in an area called the peripheral zone (PZ) and are directed into lateral organ formation. Cells also divide downwards into the rib zone (RZ) and are canalized in shoot formation. Directly underlying the stem cell niche is an area called the organizing center (OC), which consists of a small number of stem cell promoting cells marked by expression of *WUSCHEL* (*WUS*), a homeobox transcription factor that is expressed within the meristem from early in embryogenesis (Mayer et al. 1998).

Decades of work in plant models, primarily through forward genetic studies of mutants, have uncovered many factors involved in various stages of plant development. An ever-growing body of literature continues to provide deeper understanding of core factors and regulatory mechanisms and continues to add new layers of complexity to the overlapping and interconnected pathways by which the meristem forms, is maintained, matures and allows for the development of specific organ types. Here, I focus on the fundamentals of meristem maintenance and maturation.

Studies of *Arabidopsis* mutants with altered meristem size were particularly important for uncovering a major regulatory feedback loop called the CLAVATA-WUSCHEL (CLV-WUS) pathway at the core of maintaining proper meristem size (reviewed in Barton 2010; Ha et al. 2010; Stahl and Simon 2010). Work in other models, particularly the grasses has demonstrated that core components and basic mechanism of this pathway are conserved across the monocot-

dicot divide with some subtle differences (reviewed in Pautler et al. 2013). *clavata* mutants (*clv1*, *clv2* and *clv3*) and *wus* mutants develop larger and smaller meristems, respectfully. *CLV3* encodes a small peptide that is modified, cleaved, and exported to the apoplast where it is bound by various receptor complexes including homodimers of CLV1 (a leucine-rich repeat receptor kinase), a complex comprised of CLV2 (an LRR-receptor like protein lacking a kinase domain) and a pseudokinase CORYNE (CRN)/SOL2, or RECEPTOR-LIKE PROTEIN KINASE 2 (RPK2)/TOAD2 (Clark et al. 1997; Fletcher et al. 1999; Jeong et al. 1999; Kondo et al. 2006; Miwa et al. 2008; Muller et al. 2008; Ogawa et al. 2008; Ohyama et al. 2009; Bleckmann et al. 2010; Kinoshita et al. 2010) (Figure 1.1C). The perception of CLV3 triggers a signaling cascade to restrict *WUS* expression to the OC. The WUS protein feeds back on this system non-cell autonomously to positively regulate *CLV3* in the stem cell niche (Brand et al. 2000; Schoof et al. 2000). While this signaling pathway is at the core of SAM maintenance, there are many other factors and signaling pathways involved in establishing and maintaining the meristem (reviewed in Barton 2010; Ha et al. 2010; Stahl and Simon 2010; Aichinger et al. 2012) including important functions of auxin and cytokinin, KNOTTED1-like homeodomain (KNOX) signaling, and small-RNA signaling and that are not described here. Importantly, increased meristem size due to defects in the *CLV* genes can also lead to changes in inflorescence architecture, however, no such mutants have been previously described in tomato.

In parallel to regulating their size and maintaining stemness, meristems must also mature. Young vegetative meristems start out as shallow domes and grow wider and taller as they mature. Morphologically, aside from size, the meristem does not look dramatically different until it transitions to a reproductive program to generate flowers. However, the molecular profile changes significantly in these vegetative meristems as the plant matures (Park et al. 2012), and tellingly, these meristem dynamics are paralleled by morphological and molecular changes of the leaves they produce (Lifschitz et al. 2006; Efroni et al. 2008). Indeed, tomato leaves reflect these dynamics beautifully; the first leaves formed on the main shoot are rather simple, while leaves

formed just prior to the transition to flowering are much more complex (Figure 1.2). We have just begun to elucidate the molecular dynamics that underlie the phase transition from vegetative to reproductive growth and how these dynamics impact the resulting inflorescence architectures. For instance, recent profiling of tomato meristems at defined stages from early vegetative growth through the transition to flowering has revealed dynamic changes in gene expression (Park et al. 2012). Based on these dynamics, Park et al. (2012) defined a "meristem maturation clock" and explored the hypothesis that the rate of maturation regulates inflorescence branching. This is described in more detail below in the context of using tomato as a model to investigate the factors that regulate inflorescence architecture.

Tomato as a model for sympodial growth

For some plants, when the decision is made to flower and the SAM transitions to a reproductive state it is committed to this reproductive phase for the rest of its life, while other plants are able to alternate between periods of vegetative and reproductive growth. These two different growth habits are known as monopodial ("single-footed") or sympodial ("united feet") growth (Bell 1993); (Figure 1.3). Monopodial plants, for example *Arabidopsis*, have an indeterminate shoot apical meristem (SAM), which generates a certain number of leaves in a spiral pattern around the flanks of the meristems. The SAM then undergoes a phase change from a vegetative SAM that produces leaves, to a reproductive inflorescence meristem (IM) that continually produces floral meristems (FM) around its flanks to make flowers until the plant ceases growth (Figure 1.3A). Therefore, only a single monopodial or "single-footed" shoot is formed. While the meristem itself is indeterminate, monopodial growth is considered determinate in habit, because once the shoot transitions from vegetative to reproductive growth, additional shoot growth cannot be renewed from that axis (Weberling 1989).

On the other hand, sympodial plants, such as the model plant tomato[§] discussed herein, begin post-embryonic growth from a determinate primary shoot apical meristem (PSM) that terminates in the formation of the first flower (Figure 1.3B and C). The PSM starts out as a vegetative meristem (VM) and generates an average of eight leaves in a spiral pattern along the flanks of the meristem. Meristems form in the axil of each leaf and have the potential to create side shoots. Over time, the VM gradually matures to a molecularly distinct transition meristem (TM) and prepares to switch to reproductive growth. The meristem switches to an inflorescence meristem that is consumed in the formation of the first floral meristem (FM) and hence the PSM terminates in the first flower. However, before this termination event occurs, additional meristems form to renew both apical and lateral growth. Apical growth renews in the axil of the last leaf formed by the PSM from a specialized axillary meristem called the sympodial shoot meristem (SYM), which produces a few leaves and another inflorescence (Pnueli et al. 1998); (Figure 1.3B and C). While the PSM generates approximately 8 leaves in tomato before transitioning to flowering, the SYM has a molecular profile more similar to late vegetative stages of meristem development (Park et al. 2012), and therefore typically produces only three leaves before transitioning and itself terminating in the formation of the first flower of the subsequent inflorescence. This SYM reiteration continues to produce several "sympodial units" that each contains typically three leaves and an inflorescence (Pnueli et al. 1998). Again, while the meristems themselves are determinate, the renewal of apical growth from these specialized meristems give sympodial plants an overall indeterminate growth habit, because, as long as environmental conditions are right, the plant can continue to reiterate growth along the main shoot indefinitely.

In addition to the continuation of apical growth, tomatoes produce multi-flowered inflorescences, and, therefore, meristem reiteration must also occur laterally within the inflorescence (Lippman et al. 2008; Park et al. 2012). Prior to the PSM becoming consumed in the formation of the first flower, a sympodial inflorescence meristem (SIM) forms lateral to the

FM to renew growth within the inflorescence (Figure 1.3C). The SIM will itself terminate in the next FM, but not before reserving stem cells for an additional SIM (Figure 1.3D). This SIM reiteration occurs several times to generate the typical multi-flowered tomato inflorescence containing 6-10 flowers arranged in a zig-zag pattern along a single truss.

Interestingly, mutants have been found that alter the tomato growth habit. By nature, tomato has an "indeterminate" growth habit, meaning that it can continue growing like a vine reiterating between producing leaves and producing flowers. However, one mutant, *self pruning* (*sp*) transitions to flowering faster than wild type, produces only a limited number of SYMs, and develops a reduced number of leaves within each SYM unit and terminates flowering (Pnueli et al. 1998). The result is a "determinate" tomato bush that bears only a few tightly packed clusters of fruits on each branch. On the other hand, *single flower truss* (*sft*) mutants are late flowering, highly enlarged and vegetative plants that do not undergo SYM reiteration and revert to a vegetative inflorescence (Lifschitz et al. 2006). Intriguingly, *SP* encodes the tomato ortholog of *TERMINAL FLOWER 1* (*TFL*) and *SFT* encodes the tomato ortholog of the *TFL1* antagonist, *FLOWERING LOCUS T* (*FT*) (Pnueli et al. 1998; Lifschitz et al. 2006). While *SFT* is completely epistatic over *SP*, a mysterious and exciting result was found in *sft*/+ heterozygous mutants in an *sp* background; *sp* growth termination was partially relieved generating a determinate plant, but with greater SYM reiteration than in *sp* single mutants, and strikingly, this also led to increased yield over both parents (Krieger et al. 2010). However, this result left unanswered questions as to how heterozygosity for a single gene in an otherwise isogenic background caused such a dramatic effect. For instance, is dosage of *SFT* sensed by the plant? If so, how does this change the transcriptional profile? And how does this translate to altered plant habit? Moreover, as *FT* and *TFL* are highly conserved regulators of flowering, does this effect translate to other species? These questions are explored further in Chapter 2 and in the Appendix (Jiang et al. 2013; see Appendix B).

Tomato as a model for inflorescence development and architecture

Flowers are arranged on reproductive branches called inflorescences (Weberling 1989; Benlloch et al. 2007). The variety of ways in which plants can arrange their flowers on these specialized shoots is one of the most striking features of flowering plants and has a major impact on reproductive success (Wyatt 1982). These architecture types are determined by when, where and how often flowers are formed. Generally, inflorescences are classified based on two main factors, first, whether or not the shoot terminates in a flower and second, the degree of branching within the structure (Weberling 1989; Benlloch et al. 2007). The first distinction is determined by the determinacy of the inflorescence. When the inflorescence does not terminate in the formation of a flower, the inflorescence is termed "indeterminate" and in this case, it grows apically and continually generates flowers laterally on the flanks of the meristem along a single axis. This type of growth is typified by the raceme, which is the inflorescence structure produced by *Arabidopsis*. A different structure arises when the inflorescence meristem is "determinate" because its growth is terminated apically in the formation of a flower, however, growth can be renewed by specialized axillary meristems to continue lateral growth and produce additional flowers within the inflorescence. Determinate inflorescences are split into two categories, cymes and panicles (Benlloch et al. 2007; Prusinkiewicz et al. 2007). Panicles have growth along a single axis, while cymes do not, however, both terminate in the formation of a flower (Prusinkiewicz et al. 2007). Typical examples of plants that have determinate inflorescence types are tomato (cyme) and rice (panicle). Inflorescence structure is also described in terms of the complexity of the shoot architecture; when constrained to a single truss, the inflorescence is termed "simple," while a shoot structure comprised of multiple trusses is termed "compound" (Weberling 1989).

These major classifications of inflorescence types, racemes, cymes and panicles, describe the most commonly displayed architectures, however, they only represent a limited amount of the theoretically possible diversity in architecture types (Prusinkiewicz et al. 2007). This modeling

suggested that the rate at which apical and lateral meristems mature and terminate, measured by a meristem state they called *vegetativeness* (*veg*), determines the architecture type (Prusinkiewicz et al. 2007). The tomato inflorescence is a nice example of a simple cyme (single trussed, yet multi-flowered inflorescence) and provides a baseline for comparison of inflorescence architecture complexities and a great opportunity to test this theory. Within the cyme, there is an enormous amount of variation in the number of flowers and branching architectures that can develop and this is illustrated beautifully by the diversity of inflorescences produced by different Solanaceae species. For instance, species such as petunia and pepper typically develop just a single flower per inflorescence. Again, we use tomato as the "average" inflorescence, which has multiple flowers arranged along a single truss (Figure 1.4A). However, some species deviate from a simple linear truss, to generate branched inflorescences that bear a dozen to several dozen flowers. For example, several wild tomato species, like *S. peruvianum* have mildly branched inflorescences (Peralta and Spooner 2005); (Figure 1.4A). In addition to the variation in branching in diverse species a number of mutants have been identified in tomato that convert the typical single-trussed inflorescence with approximately 6-10 flowers to either a single-flowered inflorescence, as in the *terminating flower* (*tmf*) mutant (MacAlister et al. 2012) or a highly branched inflorescence as in *compound inflorescence* (*s*) (Lippman et al. 2008); (Figure 1.4B).

Focusing on branching variation in tomato, transcriptome profiling of wild type (non-branched), *S. peruvianum* (a weakly branched wild tomato species) and *s* mutant (highly branched) meristems from the vegetative stage through the termination in the first flower was able to experimentally demonstrate *veg* (Park et al. 2012). Many dynamically expressed genes were found in wild type and clustered based on their dynamics throughout the different stages to generate a "meristem maturation clock" to which the transcriptome dynamics of *S. peruvianum* and *s* were compared. It was hypothesized that *s*, a highly branched mutant would have increased *veg* (delayed maturation), which would allow for the proliferation of additional inflorescence meristems and hence branching within the inflorescence. Indeed, *s* is delayed at two time points,

within the primary meristem and within the SIM, allowing for proliferation at both stages. *S. peruvianum*, on the other hand, which is a weakly branched, is only delayed in the primary meristem, allowing for one or two extra SIMs to form, but the SIMs mature at a normal rate, hence a branching phenotype intermediate to WT and *s* (Park et al. 2012).

Importantly, this profiling revealed that many genes are dynamically expressed during this transition and that manipulation of the timing of expression of particular genes can have a profound effect on architecture. A handful of branching mutants have been cloned recently including the highly branched mutants *compound inflorescence* (*s*), *anantha* (*an*) and *falsiflora* (*fa*) (Molinero-Rosales et al. 1999; Lippman et al. 2008); (Figure 1.4B). On the other end of the spectrum is an unusual mutant that terminates in the formation of a single flower on the primary shoot, but produces normal inflorescences from sides shoots, which is called *terminating flower* (*tmf*) (MacAlister et al. 2012); (Figure 1.4B). All of the underlying genes participate in regulating meristem maturation (reviewed in Chapter 3). However, a number of tomato branching variants for which the genetic basis is unknown remain suggesting the exciting possibility of uncovering novel components of the "meristem maturation clock" or other pathways underlying branching within the inflorescence, which was a main focus of my work (see Chapters 3 and 4).

Summary

Many mechanisms converge to regulate plant architecture. The work presented in the following three results chapters focuses on mechanisms that modulate plant growth habit and inflorescence branching. First, I investigate the controversial theory of single-gene overdominance as a mechanism for driving hybrid vigor/heterosis. Results from previous and accompanying work (Krieger et al. 2010; Jiang et al. 2013; see Appendix B) demonstrate that changes in dosage for a single gene, the flowering hormone florigen, can have profound effects on plant architecture and drive yield heterosis in tomato. However, in Chapter 2, I show that while a dosage effect on plant

growth is conserved across divergent species, the yield effect is dependent on the sympodial growth habit of tomato and does not translate to the monopodial plant *Arabidopsis*. In Chapters 3 and 4, I focus on the mapping and characterization of genetic factors underlying inflorescence branching variants in tomato. Chapter 3 describes rough mapping and characterization in a handful of variants that have increased branching within the inflorescence. This work includes QTL mapping in *bifurcating inflorescence* (*bif*), which maps to a region containing *COMPOUND INFLORESCENCE* (*S*), mutants in which are known to delay the "meristem maturation clock" and alter inflorescence architecture. I hypothesize that regulatory changes in this gene may underlie the *bif* branching phenotype. Detailed characterization of an additional mutant, *fasciated and branched* (*fab*), is presented in Chapter 4 alongside a related mutant *fasciated inflorescence* (*fin*); both mutants develop branched inflorescences with extra floral organs and enlarged fruits. Interestingly, this work draws a connection between the regulation of meristem size and inflorescence architecture. *FAB* encodes the tomato ortholog of CLAVATA 1 and is the first member of the CLV-WUS meristem maintenance pathway to be identified in tomato. *FIN* encodes a transmembrane protein of unknown function and adds a new layer of complexity to the regulation of meristem size. Combined, my work demonstrates that multiple mechanisms, from modulation of plant growth habit, to meristem maturation and maintenance, converge on the regulation of shoot branching.

[§]Note: Tomato as a model organism

Cultivated tomato (*Solanum lycopersicum*) originates from South America and was domesticated from wild cherry tomato *S. lycopersicum* var. *cerasiforme* (Peralta and Spooner 2007), which is widely considered to be an admixture of *S. lycopersicum* and the closest wild relative *S. pimpinellifolium* (Ranc et al. 2008; Blanca et al. 2012). By the time of the first botanical descriptions of tomato in 1544 by an Italian doctor and naturalist, Pietro Andrea Mattioli (Mattioli 1544), fruit size increased from small, currant-sized fruits of the wild ancestor to large

beefsteak-sized fruits. Tomato is an attractive model system for many reasons (reviewed in Peralta and Spooner 2007; Kimura and Sinha 2008). Briefly, it is currently one of the most cultivated food crops in the world and a member of the Solanaceae family, which includes a number of other important food crops such as potato, pepper and eggplant. Domesticated tomato develops complex leaves, fleshy fruits in a variety of colors and shapes and a variety of arrangements within the flower-bearing branches called inflorescences. Moreover, it is a genetically tractable system; it is diploid, self-compatible and cross compatible with several of its closest wild relatives, which is useful for creating genetic mapping populations. It is also amenable for transformation using a variety of backgrounds (Sun et al. 2006; Van Eck et al. 2006). A number of genetics resources are available to aid in both forward and reverse genetics studies including saturated mutagenesis populations (Menda et al. 2004) and TILLING collections (Minoia et al. 2010; Okabe et al. 2011). Moreover, recent sequencing of the tomato and potato genomes (Xu et al. 2011; Consortium 2012), with sequencing underway for a number of other Solanaceae species (SGN 2013), has advanced genetics, genomics and evolutionary genomics studies. Tomato is therefore an attractive model fruit crop that is gaining popularity in the plant developmental genetics realm.

References

Aichinger E, Kornet N, Friedrich T, Laux T. 2012. Plant stem cell niches. *Annu Rev Plant Biol* **63**: 615-636.

Barton MK. 2010. Twenty years on: the inner workings of the shoot apical meristem, a developmental dynamo. *Developmental biology* **341**: 95-113.

Bell AD. 1993. *Plant Form: An Illustrated Guide to Flowering and Plant Morphology*. Oxford University Press, Oxford.

Benlloch R, Berbel A, Serrano-Mislata A, Madueno F. 2007. Floral initiation and inflorescence architecture: a comparative view. *Ann Bot* **100**: 659-676.

Blanca J, Canizares J, Cordero L, Pascual L, Diez MJ, Nuez F. 2012. Variation revealed by SNP genotyping and morphology provides insight into the origin of the tomato. *PLoS One* **7**: e48198.

Bleckmann A, Weidtkamp-Peters S, Seidel CA, Simon R. 2010. Stem cell signaling in Arabidopsis requires CRN to localize CLV2 to the plasma membrane. *Plant Physiol* **152**: 166-176.

Brand U, Fletcher JC, Hobe M, Meyerowitz EM, Simon R. 2000. Dependence of Stem Cell Fate in Arabidopsis on a Feedback Loop Regulated by CLV3 Activity. *Science* **289**: 617-619.

Clark SE, Williams RW, Meyerowitz EM. 1997. The CLAVATA1 gene encodes a putative receptor kinase that controls shoot and floral meristem size in Arabidopsis. *Cell* **89**: 575-585.

Consortium TTG. 2012. The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* **485**: 635-641.

Efroni I, Blum E, Goldshmidt A, Eshed Y. 2008. A protracted and dynamic maturation schedule underlies Arabidopsis leaf development. *The Plant cell* **20**: 2293-2306.

Fletcher JC, Brand U, Running M, Simon R, Meyerowitz EM. 1999. Signaling of Cell Fate Decisions by CLAVATA3 in Arabidopsis Shoot Meristems. *Science* **283**: 1911-1914.

Ha CM, Jun JH, Fletcher JC. 2010. Shoot Apical Meristem Form and Function. **91**: 103-140.

Jeong S, Trotochaud AE, Clark SE. 1999. The Arabidopsis CLAVATA2 gene encodes a receptor-like protein required for the stability of the CLAVATA1 receptor-like kinase. *The Plant cell* **11**: 1925-1934.

Jiang K, Liberatore KL, Park SJ, Alvarez JP, Lippman ZB. 2013; see Appendix B. Tomato Yield Heterosis is Triggered by a Dosage Sensitivity of the Florigen Pathway that Fine-Tunes Shoot Architecture. *PLoS Genet* **9**: e1004043.

Kimura S, Sinha N. 2008. Tomato (Solanum lycopersicum): A Model Fruit-Bearing Crop. *CSH Protoc* **2008**: pdb emo105.

Kinoshita A, Betsuyaku S, Osakabe Y, Mizuno S, Nagawa S, Stahl Y, Simon R, Yamaguchi-Shinozaki K, Fukuda H, Sawa S. 2010. RPK2 is an essential receptor-like kinase that transmits the CLV3 signal in Arabidopsis. *Development* **137**: 3911-3920.

Kondo T, Sawa S, Kinoshita A, Mizuno S, Kakimoto T, Fukuda H, Sakagami Y. 2006. A plant peptide encoded by CLV3 identified by in situ MALDI-TOF MS analysis. *Science* **313**: 845-848.

Krieger U, Lippman ZB, Zamir D. 2010. The flowering gene SINGLE FLOWER TRUSS drives heterosis for yield in tomato. *Nat Genet* **42**: 459-463.

Liberatore KL, Jiang K, Zamir D, Lippman ZB. 2013. Heterosis: the case for single gene overdominance. in *Hybrid and Polyploid Genomics* (eds. ZJ Chen, JA Birchler). Wiley-Blackwell.

Lifschitz E, Eviatar T, Rozman A, Shalit A, Goldshmidt A, Amsellem Z, Alvarez JP, Eshed Y. 2006. The tomato FT ortholog triggers systemic signals that regulate growth and flowering and substitute for diverse environmental stimuli. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 6398-6403.

Lippman ZB, Cohen O, Alvarez JP, Abu-Abied M, Pekker I, Paran I, Eshed Y, Zamir D. 2008. The making of a compound inflorescence in tomato and related nightshades. *PLoS biology* **6**: e288.

MacAlister CA, Park SJ, Jiang K, Marcel F, Bendahmane A, Izkovich Y, Eshed Y, Lippman ZB. 2012. Synchronization of the flowering transition by the tomato TERMINATING FLOWER gene. *Nature genetics* **44**: 1393-1398.

Mattioli PA. 1544. Discorsi.

Mayer KF, Schoof H, Haecker A, Lenhard M, Jurgens G, Laux T. 1998. Role of WUSCHEL in regulating stem cell fate in the Arabidopsis shoot meristem. *Cell* **95**: 805-815.

Menda N, Semel Y, Peled D, Eshed Y, Zamir D. 2004. In silico screening of a saturated mutation library of tomato. *Plant J* **38**: 861-872.

Minoia S, Petrozza A, D'Onofrio O, Piron F, Mosca G, Sozio G, Cellini F, Bendahmane A, Carriero F. 2010. A new mutant genetic resource for tomato crop improvement by TILLING technology. *BMC Res Notes* **3**: 69.

Miwa H, Betsuyaku S, Iwamoto K, Kinoshita A, Fukuda H, Sawa S. 2008. The receptor-like kinase SOL2 mediates CLE signaling in Arabidopsis. *Plant Cell Physiol* **49**: 1752-1757.

Molinero-Rosales N, Jamilena M, Zurita S, Gomez P, Capel J, Lozano R. 1999. FALSIFLORA, the tomato orthologue of FLORICAULA and LEAFY, controls flowering time and floral meristem identity. *Plant J* **20**: 685-693.

Muller R, Bleckmann A, Simon R. 2008. The receptor kinase CORYNE of Arabidopsis transmits the stem cell-limiting signal CLAVATA3 independently of CLAVATA1. *The Plant cell* **20**: 934-946.

Ogawa M, Shinohara H, Sakagami Y, Matsubayashi Y. 2008. Arabidopsis CLV3 peptide directly binds CLV1 ectodomain. *Science* **319**: 294.

Ohyama K, Shinohara H, Ogawa-Ohnishi M, Matsubayashi Y. 2009. A glycopeptide regulating stem cell fate in Arabidopsis thaliana. *Nat Chem Biol* **5**: 578-580.

Okabe Y, Asamizu E, Saito T, Matsukura C, Ariizumi T, Bres C, Rothan C, Mizoguchi T, Ezura H. 2011. Tomato TILLING technology: development of a reverse genetics tool for the efficient isolation of mutants from Micro-Tom mutant libraries. *Plant Cell Physiol* **52**: 1994-2005.

Park SJ, Jiang K, Schatz MC, Lippman ZB. 2012. Rate of meristem maturation determines inflorescence architecture in tomato. *Proceedings of the National Academy of Sciences of the United States of America* **109**: 639-644.

Pautler M, Tanaka W, Hirano HY, Jackson D. 2013. Grass meristems I: shoot apical meristem maintenance, axillary meristem determinacy and the floral transition. *Plant Cell Physiol* **54**: 302-312.

Peralta IE, Spooner DM. 2005. Morphological characterization and relationships of wild tomatoes (Solanum L. Sect. Lycopersicon). *Monogr Syst Bot, Missouri Bot Gard* **104**: 227-257.

-. 2007. History, Origin and Early Cultivation of Tomato (Solanaceae). in *Genetic Improvement of Solanaceous Crops* (eds. MK Razdan, AK Mattoo), pp. 1-24. Science Publishers, Enfield, NH.

Pnueli L, Carmel-Goren L, Hareven D, Gutfinger T, Alvarez J, Ganal M, Zamir D, Lifschitz E. 1998. The SELF-PRUNING gene of tomato regulates vegetative to reproductive switching of sympodial meristems and is the ortholog of CEN and TFL1. *Development* **125**: 1979-1989.

Prusinkiewicz P, Erasmus Y, Lane B, Harder LD, Coen E. 2007. Evolution and development of inflorescence architectures. *Science* **316**: 1452-1456.

Ranc N, Munos S, Santoni S, Causse M. 2008. A clarified position for Solanum lycopersicum var. cerasiforme in the evolutionary history of tomatoes (solanaceae). *BMC plant biology* **8**: 130.

Schoof H, Lenhard M, Haecker A, Mayer KF, Jurgens G, Laux T. 2000. The stem cell population of Arabidopsis shoot meristems in maintained by a regulatory loop between the CLAVATA and WUSCHEL genes. *Cell* **100**: 635-644.

SGN. 2013. The Sol-100 Sequencing Project. http://solgenomicsnet/organism/sol100/view.

Stahl Y, Simon R. 2010. Plant primary meristems: shared functions and regulatory mechanisms. *Curr Opin Plant Biol* **13**: 53-58.

Sun HJ, Uchii S, Watanabe S, Ezura H. 2006. A highly efficient transformation protocol for Micro-Tom, a model cultivar for tomato functional genomics. *Plant Cell Physiol* **47**: 426-431.

Van Eck J, Kirk DD, Walmsley AM. 2006. Tomato (Lycopersicum esculentum). *Methods Mol Biol* **343**: 459-473.

Weberling F. 1989. Morphology of the inflorescence. in *Morphology of flowers and inflorescences*. Cambridge University Press, Cambridge, UK.

Wyatt R. 1982. Inflorescence architecture: how flower number, arrangement, and phenology affect pollination and fruit set. *American Journal of Botany* **69**: 585-594.

Xu X, Pan S, Cheng S, Zhang B, Mu D, Ni P, Zhang G, Yang S, Li R, Wang J et al. 2011. Genome sequence and analysis of the tuber crop potato. *Nature* **475**: 189-195.

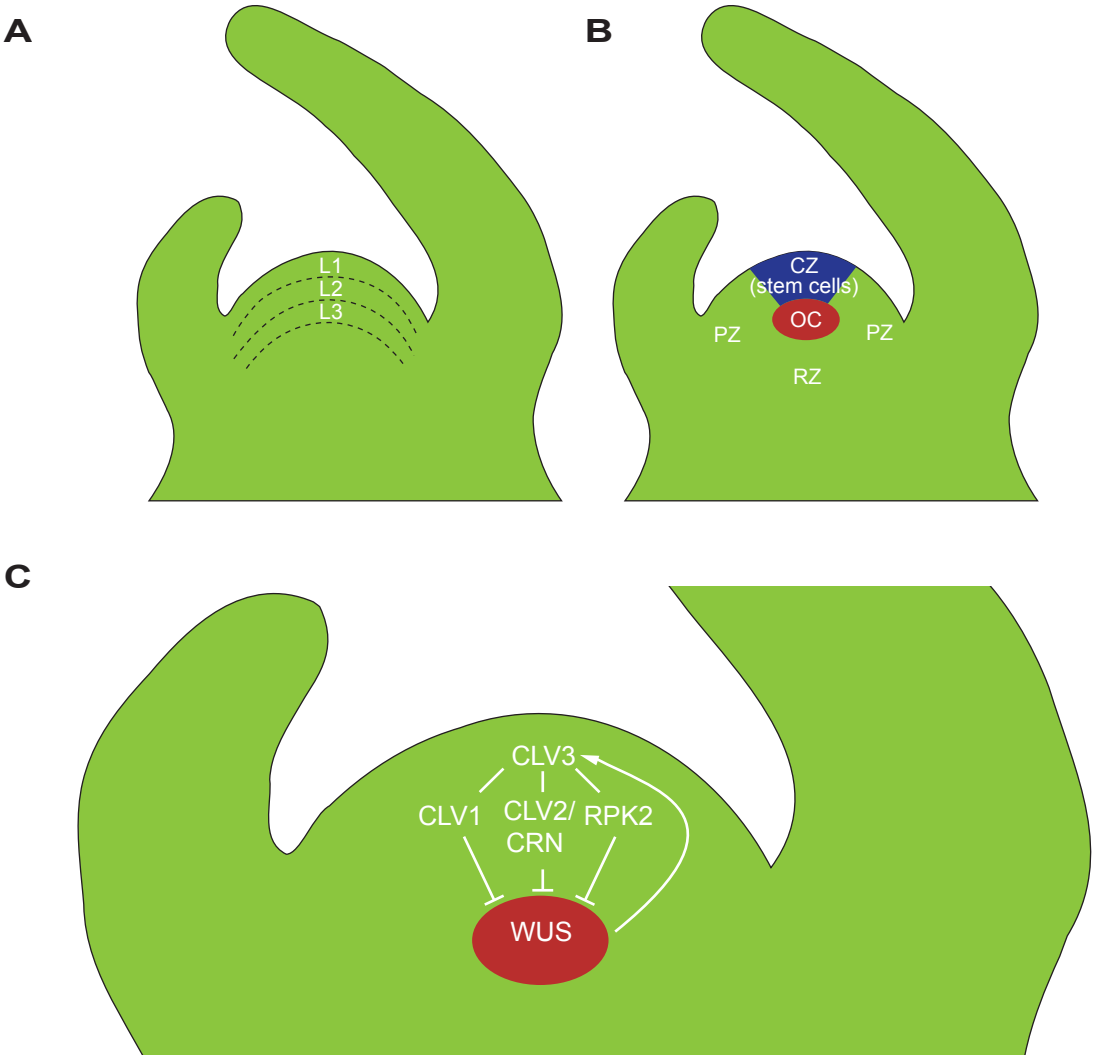


Figure 1.1. Shoot apical meristem organization and the CLV-WUS pathway. (A-C) Renderings of the shoot apical meristem (SAM). (A) Three clonally distinct cell layers are formed in the SAM. The layers (L) 1-3 are marked. (B) Sub-domains of the SAM. The central zone (CZ) is where the stem cells reside (highlighted in blue). Cells in the CZ divide slowly to replenish themselves and generate cells that are pushed out into the peripheral zone (PZ) and the rib meristem (RM). Cells in these zones divide more rapidly and are canalized into lateral organ growth and shoot growth, respectively. Below the CZ is a stem cell promoting region called the organizing center (OC) that is marked by expression of the homeobox transcription factor *WUSCHEL* (*WUS*) (highlighted in red). (C) The CLV-WUS stem cell maintenance pathway. CLV3 peptide is perceived by multiple receptor complexes, which triggers signaling to restrict *WUS* to the OC. In turn, WUS feeds back to promote *CLV3* expression in the CZ to canalize the signaling pathway.

Figure 1.2. Changing meristem dynamics are reflected in the leaves. Leaves from the primary shoot, axillary shoots associated with each leaf, and the first sympodial shoot were dissected from a maturing wild type tomato plant. The primary shoot leaves (bottom) are arranged from left to right from the first leaf formed to the last leaf formed before flowering. Several leaves (on average 8, shown here 10) are made during the transition to flowering. As the plant matures, exceedingly complex leaves are made. Axillary shoots (middle) have the potential to be released from the axil of each leaf. The primary shoot terminates in the formation of the first flower of the first inflorescence (see Figure 1.3), however, growth is renewed by a specialized axillary meristem called the sympodial shoot meristem (SYM), which forms the sympodial shoot. On average, each SYM generates three leaves and terminates in the formation of the first flower of the second inflorescence. Note that the first leaf in the SYM is more complex than the first leaf of the primary shoot. This reflects the molecular age of the SYM, which is intermediate to the age of the PSM and an inflorescence meristem (see text).

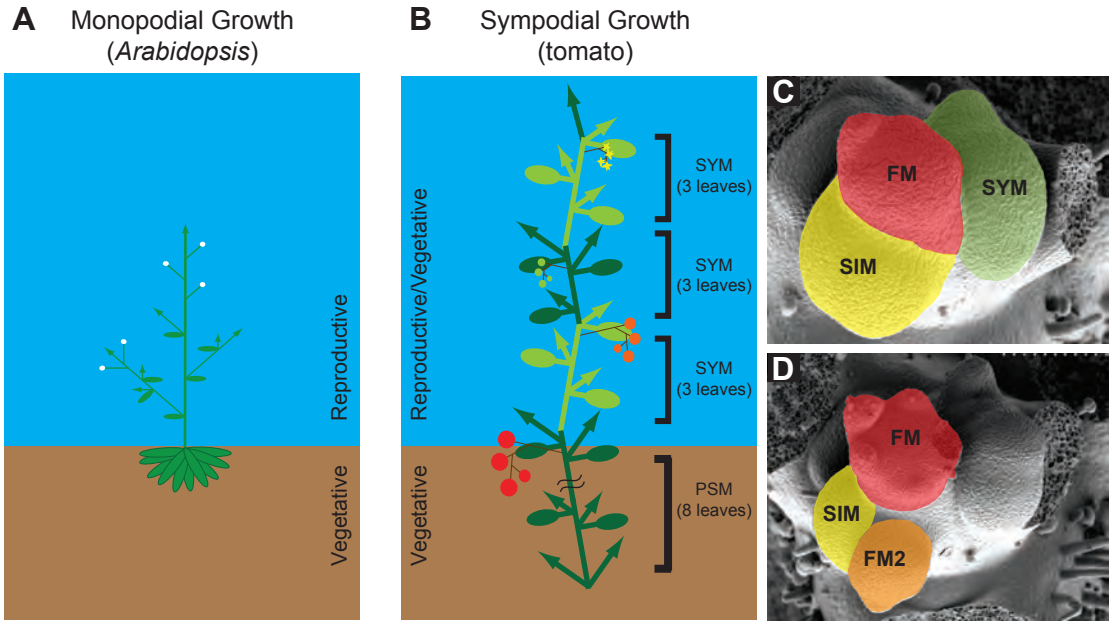


Figure 1.3. Monopodial versus sympodial growth habit. (A and B) Renderings of an *Arabidopsis* plant, which has a monopodial growth habit (A) and a tomato plant, which has a sympodial growth habit (B). (A) Monopodial plants have determinate growth. Growth begins with a vegetative shoot apical meristem, which produces vegetative growth and can only transition once to reproductive growth. Once the switch to reproductive growth is made, only flowers are made on the main shoot before it terminates growth. (B) Sympodial plants like tomato have an indeterminate growth habit. The primary shoot meristem (PSM) generates approximately 8 leaves and then terminates in the first flower (FM). (C) Before PSM termination, stem cells are reserved to form two additional meristems, the sympodial shoot meristem (SYM) and the sympodial inflorescence meristem (SIM), which renew apical vegetative growth and lateral growth within the inflorescence, respectfully. (D) The inflorescence also experiences sympodial growth. Before the SIM terminates in the second FM (FM2), an additional SIM is formed. This reiteration continues to generate a linear truss of approximately 8-10 flowers on each inflorescence. (B) SYM reiteration occurs until the death of the plant, generating the vine-like tomato plant. At the axil of each leaf, an axillary meristem is formed that has the potential to form axillary shoots (green arrows). At each SYM reiteration, one inflorescence is formed and as the plant grows, varying stages of inflorescence maturity can be seen; the last inflorescence formed will be making flowers (yellow stars), while previous inflorescences have progressed to making fruits (green, orange and red balls are colored to denote maturation status from immature to ripe fruit).

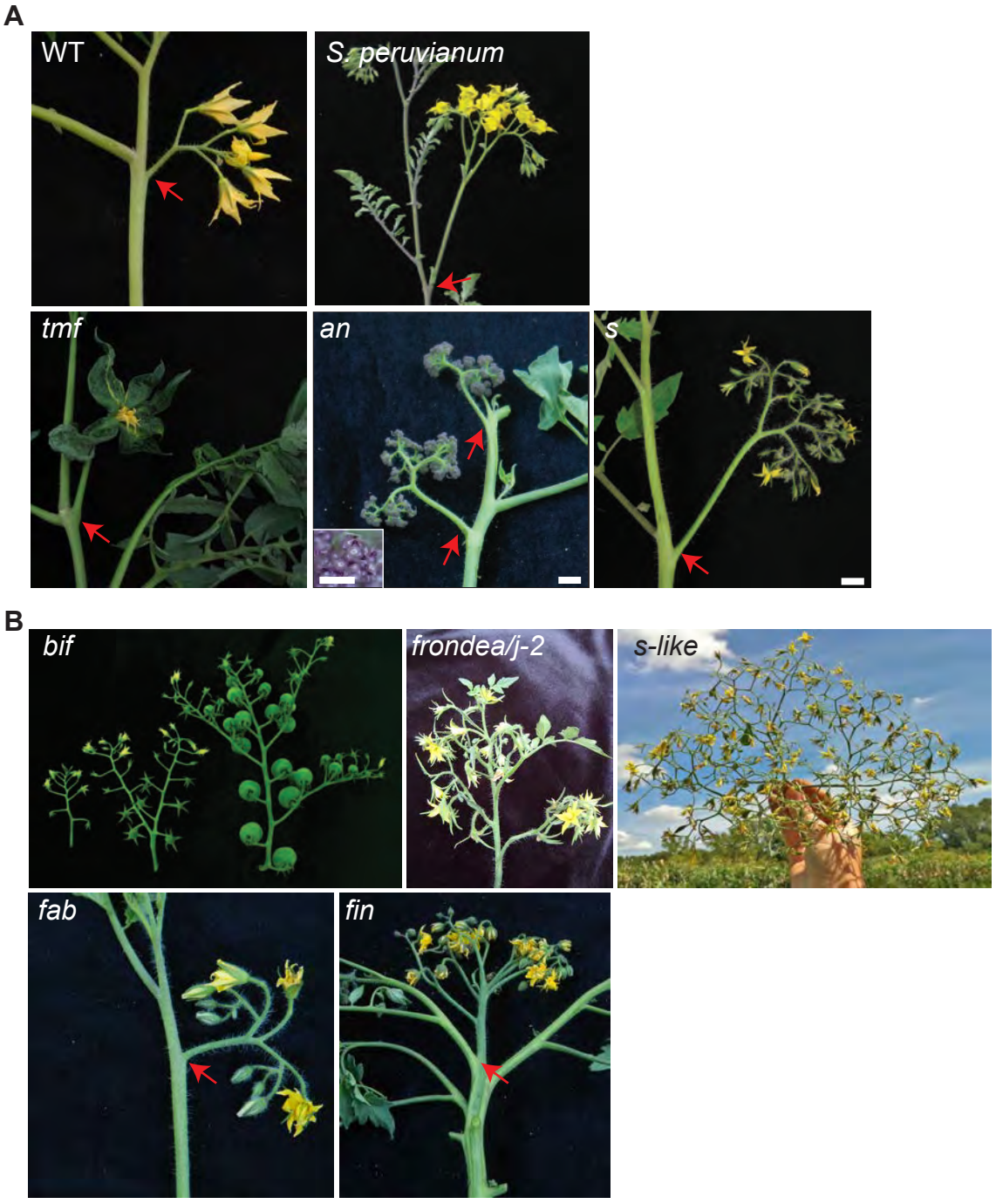


Figure 1.4. Tomato inflorescence branching variants. (A) Images of wild type domesticated tomato (WT) with a standard simple truss, the mildly branched wild species *S. peruvianum*, and inflorescence mutants *terminating flower* (*tmf*), *anantha* (*an*) and *compound inflorescence* (*s*), for which the molecular basis is known (see text). (B) Inflorescence branching variants investigated in this study. *bifurcating inflorescence* (*bif*), *frondea/jointless-2*, *s-like* and *fab* are investigated in Chapter 3. Further characterization of *fab* and an additional mutant with similar phenotypes *fasciated inflorescence* (*fin*) are presented in Chapter 4.

2 Exploration of single-gene overdominance in *Arabidopsis thaliana*

Chapter Contributions:

Work presented in this chapter was performed in collaboration with Dr. Ke Jiang (KJ). Portions of the introductory and discussion text included here were adapted from a chapter we wrote together for the book *Polyploid and Hybrid Genomics* (Appendix A) with minor modifications and updates. With regard to the results sections presented herein, I designed the experiments, KJ assisted with phenotyping, and I analyzed and wrote the results. Experiments on the classic example of *erecta* mutant heterozygosity as a case of single locus heterosis are unpublished. The text from the *ft/+ tfl* experiment was based on a recently published manuscript (Appendix B).

Publications Associated with this Chapter:

Liberatore KL, Jiang K, Zamir D and Lippman ZB. 2013. Heterosis: the case for single-gene overdominance. in Hybrid and Polyploid Genomics (eds. ZJ Chen, JA Birchler). Wiley-Blackwell. (see Appendix A)

Jiang K, **Liberatore KL**, Park SJ, Alvarez JP, Lippman ZB. 2013. Tomato yield heterosis is triggered by a dosage sensitivity of the florigen pathway that fine-tunes shoot architecture. *PLoS Genet.* 9(12):e1004043. (see Appendix B)

Introduction

For centuries, naturalists as well as plant and animal breeders have noted that prolonged inbreeding in normally outcrossing populations leads to progressive accumulation of inferior traits such as smaller, less vigorous, sickly and often malformed offspring (Darwin 1868; Charlesworth and Willis 2009). The genetic and molecular basis of this "inbreeding depression" is still not completely understood (Crow 2008). However, a widely supported hypothesis is that the accumulation of spontaneously formed deleterious recessive mutations are unmasked upon inbreeding, culminating in often maladaptive phenotypes (Charlesworth and Willis 2009). The evolutionary implications of inbreeding depression are widespread, and this topic remains a focal point of population genetics research. Yet, it is the surprising and mysterious antithesis to inbreeding depression known as "hybrid vigor" (also known as heterosis) that has captured the imaginations of breeders and scientists alike for more than a century. Hybrid vigor - the phenotypic superiority and improved fitness among progeny resulting from crossing genetically distinct parents - was first described by Charles Darwin and later refined by the maize geneticists George Shull and Edmund East. Perhaps the most renowned demonstration of hybrid vigor is the mule -- a stronger and more fit, albeit sterile, animal resulting from mating a male donkey with a female horse. Countless additional examples of both plant and animal hybrid vigor have been noted over the last century leading to the suggestion that increased heterozygosity in hybrid organisms provides a "magical" genetic and physiological advantage in growth and fecundity extending beyond the simple masking (i.e. complementation) of deleterious alleles. Thus, over the last century, and especially within the last decade, great efforts have been devoted to deciphering the genetic and molecular underpinnings of hybrid vigor in diverse organisms ranging from yeast to humans.

Hybrids are found throughout nature and in agriculture, and examples of hybridization span the gamut from whole-genome heterozygosity between distinct species to hybridization between nearly isogenic breeding lines that differ only in a small chromosomal segment (i.e. an introgression)

harboring just a few dozen or even a single gene. Although hybridization between genetically distinct parents creates new allelic combinations genome-wide, drastic phenotypic changes in offspring are not always observed. Studies throughout the 20th century involving crosses between different *Drosophila* species, however, revealed several examples where hybrid phenotypes extended beyond the bounds of parents for fitness traits such as viability (Dobzhansky 1950), fertility (Fry et al. 1998), growth rate (Houle 1989), and even cold tolerance (Jefferson et al. 1974). Such "transgressive variation" can be negative, in which progeny are less fit than their parents, or positive, in which progeny exceed the fitness of their progenitors (hybrid vigor/heterosis).

The mechanisms underlying these hybridization-induced transgressive phenotypes are being actively investigated at both the genetic and molecular levels, and studies have shown that several cases of reduced vigor and fitness trace back to simple negative epistatic interactions in hybrids. The most well documented cases have been investigated for nearly a century in the context of speciation and hybrid incompatibility in *Drosophila* (Bateson 1909; Muller 1940; Muller 1942; Wallace and Dobzhansky 1962; Dobzhansky and Spassky 1968). First Bateson, and then Dobzhansky and Muller, noted that crossing different *Drosophila* species often produced sterile or lethal progeny, and surprisingly, genetic analysis indicated as few as two interacting loci were involved. This led to the Bateson-Dobzhansky-Muller incompatibility (BDMI) model, which states that at least two loci, having evolved independently into new allelic forms either through sympatric (overlapping geographic distributions) or allopatric (non-overlapping geographic distributions) speciation, show negative epistasis and reduced fitness upon "meeting" each other again in hybrids (Wallace and Dobzhansky 1962; Dobzhansky and Spassky 1968; Orr 1996). Recently, the molecular identities of several BDMI genes have been discovered and represent a range of molecular functions from chromatin binding factors (Brideau et al. 2006) to components of nuclear pore complexes (Tang and Presgraves 2009). Suggestive that BDMI is widespread in nature and a driver of speciation is the rediscovery and molecular dissection of hybrid necrosis in plants (Bomblies et al. 2007). By crossing hundreds of *Arabidopsis thaliana* accessions in a large hybridization matrix scheme, it was

discovered that 2% of hybrids exhibit an autoimmune-like response at natural growth temperatures, which is triggered by epistatic interactions between two disease resistance genes (Bomblies et al. 2007; Bomblies and Weigel 2007). Additional examples of BDMI in *Arabidopsis* have been found, including embryonic lethality and stunted root growth, and, interestingly, these examples involve different classes of genes (Bikard et al. 2009). Thus, similar to *Drosophila*, the molecular dissection of BDMI in plants suggests that several pathways and networks involved in growth and fitness can drive negative epistasis. Importantly, BDMI interactions have also been found through environment-driven laboratory evolution in yeast, highlighting the spontaneous, rapid, and dramatic impacts that negative epistasis can have on populations and the process of speciation (Anderson et al. 2010). Strikingly, a recent example in *Arabidopsis* showed that heterozygosity at just a single locus harboring tandemly-repeated receptor-like kinase genes can cause negative epistasis (Smith et al. 2011). It should be emphasized that for all of these examples, genetic context and growth condition is key (Wallace and Dobzhansky 1962; Dobzhansky and Spassky 1968; Bomblies et al. 2007; Bikard et al. 2009; Anderson et al. 2010). For instance, hybrid necrosis in *Arabidopsis* is rescued at high temperatures, suggesting that BDMI is driven through interactions between genes and environmental selection pressures. With such dramatic phenotypic consequences resulting from negative epistatic interactions in hybrids involving just one or two genes, the obvious question has been whether the opposite side of the hybridization coin, hybrid vigor (also referred to as heterosis), might be based on similarly simple genetic mechanisms.

The first controlled experiments to understand the positive attributes of hybridization were performed in plants. Although Gregor Mendel and plant breeders of the early 19th century noted anecdotally the increased vigor of hybrid plants, Darwin completed the first set of extensive experimentation on the subject (Mendel 1865; Darwin 1876). Darwin observed increased vigor in hybrid progeny resulting from crosses of different parental lines across many genera of plants, although the specific species and breeding lines, as well as the environment, impacted the magnitude of vigor. Although the benefits of hybridization were evident to plant and animal breeders of the time,

and a remarkable allusion to transgressive variation and hybrid vigor extends as far back as biblical times (Gen. 30:31-43), this phenomenon received little attention until its rediscovery by maize geneticists in the early 1900s, which set the foundation for deciphering the genetic basis of what is now formally known as "heterosis." While George Shull and Edmund East both reported severely diminished vigor in maize inbred lines over several generations of inbreeding, Shull first reported recovery and dramatically improved vigor in hybrids (East 1908; Shull 1908). Over the following decade, the utilization of heterosis became widespread in maize breeding, and the debate over two prominent theories for its genetic basis, "dominance" and "overdominance", had begun (East 1908; Shull 1908; Bruce 1910; Jones 1917; Singleton 1941).

The genetic basis of heterosis has eluded researchers for over a century despite extensive genetic and molecular experimentation. Of all approaches, classical and quantitative genetics have provided some of the greatest progress in deciphering heterosis and addressing the models of dominance and overdominance (East 1908; Shull 1908; East 1909; Birchler et al. 2003; Hochholdinger and Hoecker 2007; Lippman and Zamir 2007; Springer and Stupar 2007; Chen 2010). Dominance complementation, which is the most widely accepted model (Lippman and Zamir 2007; Gore et al. 2009), presumes that superior dominant alleles complement non-overlapping deleterious recessive alleles at potentially hundreds or even thousands of loci across the genome in the F1 hybrid. Overdominance, on the other hand, posits that intralocus allelic interactions at just a single heterozygous gene can cause heterosis, or that several overdominant loci of small effect contribute cumulatively to heterosis.

Although few and far between, it is important to note that examples of overdominance tracing back to a single gene have been observed in several organisms, including yeast, plants and animals (Schuler 1954; Mukai and Burdick 1959; Rédei 1962; Efron 1974; Hall and Wills 1987; Grobet et al. 1997; Schuelke et al. 2004; Mosher et al. 2007; Delneri et al. 2008). These scattered reports linking heterozygosity at a single gene to transgressive phenotypes remain the most tantalizing from both a fundamental and applied perspective, because, single gene overdominance could easily be leveraged

for crop improvement. One hypothesis for the molecular nature of single gene overdominance is that a combination of alleles encoding proteins adapted to different conditions, for example isozymes, generates better performance in a wider condition spectrum, and thus an overdominant effect on the hybrid phenotype. For instance, nearly 40 years ago, Efron (1974) described a case of single gene heterosis in maize for two inbred lines homozygous for different alcohol dehydrogenase (ADH) alleles that produced isozymes seemingly adapted to produce maximum enzyme activities in two different tissues: scutellum and pollen. Hybrids between the two lines now heterozygous for the two ADH alleles showed optimized enzyme activities in both tissues, thus expanding the enzyme activity spectrum and producing a more balanced overall metabolic efficiency (Efron 1974). Intriguingly, heterozygosity for two distinct temperature-sensitive alleles of an alcohol dehydrogenase isozyme (ADH1) in yeast confers transgressive alcohol tolerance over either parent (Hall and Wills 1987). Interestingly, as in maize, the overdominance in yeast seems also to result from broader enzyme activity in the heterozygous individuals. Together, these studies imply that an expanded activity spectrum due to the unification of alleles adapted to different conditions might be a simple molecular mechanism to explain at least some cases of single gene overdominance.

However, there are examples of single gene overdominance involving mutant alleles in the heterozygous condition that cannot be explained by the aforementioned molecular mechanism, as only the dose of a gene and gene product is altered in mutant heterozygotes, and there is therefore no manipulation of activity spectrums as for isozymes. For instance, a classical example of mutant single gene overdominance is that of heterozygosity for a mutant allele of hemoglobin. Individuals homozygous for mutated hemoglobin suffer from sickle cell anemia, because the mutated protein causes the formation of abnormal crescent-shaped blood cells that have aggregation problems and are not efficient oxygen carriers; however, individuals who are heterozygous for one mutant copy of this gene do not suffer from this disease because they are also able to make normal blood cells, and remarkably they also have a higher resistance to malaria than individuals who are homozygous wild type for the hemoglobin gene (Allison 1954). Such early examples of mutant single gene heterosis

were criticized as rare and special conditional cases, and thus overdominance was still widely contested. Yet, additional evidence for "mutant overdominance" in *Drosophila* (Mukai and Burdick 1959), an intriguing case involving the classical *erecta* mutant in *Arabidopsis* (Rédei 1962), and an example from our lab demonstrating that manipulating dosage of the flowering hormone florigen can drive heterosis for yield in tomato (Semel et al. 2006; Krieger et al. 2010; Jiang et al. 2013; see Appendix B), bolster the case for mutant overdominance driving heterosis and is suggestive that this effect can arise from diverse genetic pathways. (For a more extensive review of single-gene overdominance see Appendix A).

Many unanswered questions remain regarding the mechanism(s) of overdominance. Is there classical dosage in the expression levels of genes in a heterozygous state that underlie overdominant phenotypes or is there also transgressive gene expression? Is gene expression downstream of such genes altered, and if so, how? Does the specific allele and genetic background alter the transgressive phenotype? What is the association between molecular changes and the phenotype? Do overdominant effects translate across species? With these questions in mind, this chapter explores two cases of mutant single-gene overdominance in *Arabidopsis thaliana.* In the first part, I revisit the classic Rédei example of *erecta* heterozygosity driving heterosis and describe plans to profile the molecular changes in *erecta* heterozygotes in order to improve our understanding of how gene dosage effects downstream gene expression and how this might translate to transgressive phenotypes. However, following multiple quantitative phenotyping experiments, I was unable to consistently reproduce Rédei's results and traced discrepancies to likely technical artifacts in Rédei's original experimental design. In the second part, I investigate florigen mutant heterozygosity in *Arabidopsis* relative to tomato, in which florigen-dependent overdominance is observed, and describe that while a dosage effect from florigen mutant heterozygosity is conserved, it does not translate to heterosis in *Arabidopsis*. I postulate that these differences trace back to the different growth habits of the two species.

Results

Validation of *erecta*/+ single-locus heterosis in the F1 generation, but not the F2 generation

With the goal of gaining better understanding of the molecular basis of single-gene overdominance driving heterosis, I revisited a classic example of single-locus heterosis that was published over 50 years ago in *Arabidopsis* (Rédei 1962) involving the *ERECTA* (*ER*) gene. *ERECTA* encodes a leucine-rich repeat (LRR) receptor-like kinase (RLK) (Torii et al. 1996), a member of a LRR-RLK family that is involved in many aspects of plant development (van Zanten et al. 2009). Loss-of-function mutations in *ER* cause pleiotropic developmental defects including stunted shoot growth, decreased internode elongation in the inflorescence, and shorter, wider siliques compared to wild type (Torii et al. 1996). Rédei isolated mutations in the *ERECTA* (*ER*) gene and described transgressive phenotypes when the mutation was in the heterozygous state. He subsequently performed several generations of backcrossing to its wild type progenitor line to evaluate whether *erecta* heterozygosity was responsible for this phenotype or if additional background mutants contributed to the transgressive variation. However, confirmation of these transgressive phenotypes with additional alleles is lacking and the molecular mechanism by which *er* heterozygosity translates to transgressive phenotypes has not been investigated. Therefore, I sought to first verify this overdominant effect with five independent *er* alleles in three different genetic backgrounds to determine how specific this effect is based on genotype. Upon confirmation of this effect, I aimed to investigate the molecular basis for the transgressive phenotypes by profiling transcriptional changes in *erecta* heterozygotes compared to progenitor lines. Transcriptional profiling would allow us to begin to answer open questions regarding single-gene overdominance, such as whether there is classical dosage in gene expression levels upon loss of a single copy of *ERECTA*, if and how gene expression downstream of *ERECTA* is altered, how specific alleles might alter these transcriptional changes, and how these transcriptional changes may translate to increased vigor in the heterozygous plants.

First, to verify that *er*/+ heterozygosity indeed resulted in transgressive phenotypes compared to the parental lines, we generated *er*/+ F1 heterozygous lines for five independent *er* alleles, including the original X-ray mutant, two T-DNA insertion lines and two EMS alleles, all of which have similar strong phenotypes in the homozygous state. Rédei harvested and measured his plants at the "beginning of maturity i.e. about one month after germination" and found that first generation hybrids were heterotic for several traits, including length of the main stem, yield and plant weight (Rédei 1962). Preliminary measurements on both of the T-DNA *er*/+ F1 heterozygous lines showed a significant increase in main shoot height in comparison to their progenitor lines (Figure 2.1A) consistent with Rédei's results. While I did not carefully phenotype the other three alleles in this preliminary analysis, the main shoot height did appear to be taller in the three additional er/+ F1s in comparison to their parental lines (Figure 2.1B-D). I also measured the size of the rosette (diameter from opposing oldest leaves) and, as a proxy for flowering time, the number of rosette leaves produced prior to flowering, but did not detect a significant difference between wild type and the *er*/+ F1 heterozygotes (not shown). Therefore, I focused the remaining analysis on the height of the main shoot.

A major criticism of the theory of single-gene overdominance driving heterosis is that transgressive phenotypes seen in F1 hybrids are often lost in the F2 generation. This common observation supports the theory of dominance complementation as the major driver of heterosis, which is based on the idea that increased vigor in the F1 hybrids is a result of complementation of deleterious effects genome-wide, but due to meiotic recombination, this effect is lost in the F2 generation. However, generating crosses between *erecta* mutants and their progenitor wild type lines allowed us to create *er*/+ heterozygotes in nearly isogenic backgrounds and therefore, if *erecta* heterozygosity is a true case of single-gene overdominance driving heterosis, this effect should be maintained in the F2 generation and should cosegregate with *erecta* heterozygosity. Preliminary measurements were promising in the F1 generation; however, in the two *er*-T-DNA F2 populations segregating for *er*/+ heterozygotes, the height of *er*/+ plants was indistinguishable from wild type

(Figure 2.2), suggesting that *er*/+ heterozygosity does not cause heterosis, at least for plant height, or that a special intralocus allelic interaction occurs in the F1 generation when the mutant and wild type alleles meet each other for the first time, but this special interaction is lost in the F2 generation after segregation occurs.

Height advantage of *er*/+ F1 heterozygotes is not consistently correlated to seed size and is lost under controlled sowing density

Preliminary phenotyping at an early time point demonstrated that bolt height was significantly higher in the F1 *er* heterozygotes compared to both parents (Figure 2.1A), consistent with Rédei (Rédei 1962). However, in the F2 populations, no significant differences in bolt height were observed (Figure 2.2), suggesting that technical artifacts in the design of the F1 experiments could have skewed the results. One possibility was that the F1 crosses could have larger seeds and therefore have a growth advantage even prior to germination. Although homozygous mutant seeds were smaller than F1 seeds for all mutant alleles in all four different backgrounds, F1 seeds are not consistently larger than wild type (Figure 2.3). While I cannot rule out the possibility that seed size is influencing the vigor in particular F1s, there is not a consistent correlation between seed size and genotype and therefore, seed size cannot be the only contributing factor providing a growth advantage.

I next considered that differences in sowing density between the parents and the F1s could have affected plant vigor and skewed the outcome. Due to the limited number of F1 seeds resulting from each cross, these seeds were always sown carefully as single seeds. However, because there was plenty of seed from the parental lines, these seeds were sown more densely, and once germinated, thinned to single plants. In the F2 population, sowing density was more normalized among genotypes because all seeds were sown by tapping a few seeds per cell and subsequently thinned to single plants after germination and therefore the *er*/+ heterozygous plants were also germinated under crowded conditions. I hypothesized that the heterozygous F1 plants were more vigorous because they were given a growth advantage when sown; specifically, in contrast to the homozygous parental lines, the

F1 *er/+* heterozygotes were not crowding each other and competing for space and nutrients. To determine if sowing density affected plant vigor and influenced our results, I compared two different sowing methods side-by-side; controlled sowing of single seeds versus uncontrolled sowing density followed by thinning to single seedlings (Figure 2.4). As suspected, when I carefully sowed single seeds for all genotypes, the height advantage was lost in the *er/+* heterozygous F1 plants (Figure 2.4B), in contrast to the increased vigor observed in *er/+* heterozygotes that were grown at the same time in the greenhouse, but along side parental lines that were sown at high density and then thinned to single seedlings after germination (Figure 2.4A). Therefore, a growth disadvantage for the homozygous parental lines was introduced by our experimental design, which led us to misinterpret the growth of the F1s as a true case of heterosis, when in fact this does not hold under controlled sowing conditions. My detailed and careful phenotyping of five *erecta* alleles in three different backgrounds is the most thorough analysis of *erecta* heterozygosity and demonstrates that the single-locus heterosis described by Rédei could not be replicated. Sowing density may explain the discrepancy between our results and Rédei's original results. While it is impossible to tell from Rédei's description as to how he handled his seeds, all of his experiments involved sowing F1 seeds from each BC generation, and as in our experiments, these precious seeds were likely sown with greater care than the progenitor lines.

A dosage effect from florigen mutant heterozygosity is conserved in *Arabidopsis*, but does not cause yield heterosis as in tomato

Early evidence for "mutant overdominance" in *Drosophila* (Mukai and Burdick 1959), combined with the classical case involving the *erecta* mutant in *Arabidopsis* (although we were unable to reproduce this effect as described above) (Rédei 1962), motivated work in tomato to investigate whether and to what extent mutant single-gene heterozygosity could translate to yield heterosis (Semel et al. 2006; Krieger et al. 2010). Intriguingly, tomato plants homozygous for mutations in *SINGLE FLOWER TRUSS* (*SFT*), the ortholog of *Arabidopsis FLOWERING LOCUS T* (*FT*), which

encodes the flowering hormone florigen, are enlarged, highly vegetative plants that are late flowering and produce very few inflorescences and flowers (Lifschitz et al. 2006); however, in the heterozygous state, *sft*/+ plants exceed flower and fruit production of its progenitor lines and even of high yielding varieties (Krieger et al. 2010). It must be noted that this effect is conditional on a background mutation in a flowering repressor in the *SFT* gene family called *SELF PRUNING (SP)*, the tomato ortholog of *TERMINAL FLOWER 1* (*TFL1*), mutations in which create short, bushy plants (Pnueli et al. 1998). By nature, tomato plants grow like vines and can be kept flowering for years under the right conditions, such as in the greenhouse. This is because of the sympodial growth habit of tomato (Figure 1.3). In sympodial plants, a determinate SAM terminates in the formation of the first flower of an inflorescence, however, before this termination occurs, two additional specialized determinate meristems are formed, the sympodial inflorescence meristem (SIM) and the sympodial shoot meristem (SYM), which renew growth within the inflorescence and apical growth, respectfully. The SYM produces a few leaves before itself transitioning to flowering and this SYM reiteration continues indefinitely under conditions conducive to growth. However, mutations in *SP* condition this sympodial growth habit by both reducing the number of flowering transitions and the number of leaves produced within each transition on each shoot to produce only a few inflorescences clustered inflorescences before terminating, effectively "self-pruning" the indeterminate tomato vine to a determinate bushy plant. Determinate (*sp* mutant) varieties are popular in the processing tomato industry, which relies on growing large amounts of fruit in the field, because the "self-pruning" facilitates synchronization of fruit maturation thereby maximizing yield with a single mechanical harvest.

While the mechanism underlying *sft*/+ overdominance has been unclear, recent work in the lab shows that epistatic interactions between *SFT* and *SP* is key to the overdominant effect, in which *sft*/+ heterozygosity causes a weak dosage-dependent suppression of the *sp* determinate phenotype (Krieger et al. 2010; Jiang et al. 2013; see Appendix B). Our lab investigated the molecular underpinnings of *sft*/+ overdominance by profiling transcriptomes. *SFT* is expressed in leaves and

acts as a mobile signal. We profiled young leaves and transition stage meristems (TMs) in *sft*/+ *sp*, *sft sp* double mutants and *sp* single mutants, prior to the floral transition to determine if *sft*/+ heterozygosity impacted the transcriptome profile in these tissues. Even with just the loss of a single copy of *SFT*, many genes were differentially expressed when comparing all three genotypes. We next used the digital differentiation (DDI) algorithm to assess the maturation states of the different genotypes. DDI takes "reference" samples (here *sp* and *sft sp*), defines a set of transcripts that peak in each reference sample, and then interrogates the transcript profile of an "unknown" sample (*sft*/+ *sp*) and estimates its maturation state in relation to the references (Efroni et al. 2008). Interestingly, in both leaves and TMs, *sft*/+ *sp* was predicted to have a maturation state intermediate to *sp* and *sft sp*, indicating that dosage at this single gene was sensed in a more global way in the meristem. Furthermore, the dosage in *sft/+sp* individuals causes a delay in the flowering transition on the main and side shoots compared to *sp*, enabling the production of 1-2 additional sympodial units, and therefore 1-2 additional inflorescences, each with ~7-9 flowers on every shoot before termination, and thus *sft*/+ dosage is detected molecularly at each flowering transition and is compounded plant wide, resulting in a significant increase in inflorescence production and increased fruit yield in determinate tomato types (Jiang et al. 2013; see Appendix B).

As florigen is a universal inductive signal for flowering that several flowering pathways converge upon (Shalit et al. 2009; Wigge 2011), we wondered if and how florigen mutant heterozygosity in a different system might affect growth, and specifically whether heterosis would result. We tested this by creating orthologous mutant combinations in *Arabidopsis thaliana*, a monopodial plant in which a single flowering event converts the SAM into a continuously growing inflorescence meristem (IM) that produces flowers laterally, in contrast to the tomato sympodial growth habit in which multiple flowering transitions occur (Figure 1.3). Despite this difference, *Arabidopsis ft* (*sft*) mutants are likewise late flowering (Koornneef et al. 1991) and completely epistatic over the early flowering and precocious termination of inflorescence meristems of *tfl* (*sp*) mutants (Shannon and Meeks-Wagner 1991). To evaluate potential dosage effects of *ft*/+

heterozygosity, we phenotyped progeny from *ft-2/+ tfl1-2* plants, in which the *ft-2* mutation, a strong allele, segregates in the *tfl1* background (Figure 2.5A). We measured flowering time by counting rosette leaves and found a clear dosage effect in *ft-2/+ tfl1* plants compared to *tfl1* single and *ft-2 tfl1* double mutants (Figure 2.5A). We next tested for heterosis by quantifying yield related traits, including plant height, number of axillary shoots, and, as a parallel to tomato yield, the number of siliques, flowers, and flower buds (Figure 2.5B-C and Figure 2.6C-D). Surprisingly, *ft/+ tfl* plants showed semi-dominance for plant height and total yield (Figure 2.5B and C), and similar effects were observed for a moderate second allele of *ft* (Figure 2.6E). Thus, whereas the dosage effect on flowering time from florigen mutant heterozygosity is conserved in the monopodial growth habit of *Arabidopsis*, it does not translate to heterosis.

Discussion

Investigation of two cases of single-gene heterozygosity in *Arabidopsis thaliana* demonstrated the sensitivity of plants to both environmental and intrinsic, genetic context. I revisited a classical case of *erecta* heterozygosity, but was unable to consistently reproduce heterosis for plant height as was described previously (Rédei 1962). Under controlled sowing conditions of F1 heterozygotes (Figure 2.4B) and in the F2 generation (Figure 2.2), the main shoot height of the *erecta*/+ heterozygotes was indistinguishable from wild type matched lines. The methods of the Rédei paper do not specify the sowing conditions and it is possible that Rédei introduced a similar technical artifact by carefully sowing his precious F1 seeds and liberally sowing the progenitor lines and thinning the plants out post-germination as we originally did. It is also possible that *erecta* heterozygosity is a true case of single locus heterosis, but that our growth conditions were different enough from Rédei's that we did not see this same effect, however, with careful phenotyping of five *erecta* alleles in three genetic backgrounds, I was unable to reproduce his results. Two differences in our growth conditions are apparent from Rédei's methods; first, he grew his plants under constant light, and, second, the plants

were grown 9 plants per 5-inch pot as opposed to our planting in flats with smaller cells. It is unclear why Rédei chose to grow his plants under constant illumination and why the *erecta* heterozygotes would perform better than wild type under these conditions. We chose to perform our experiment under more biologically relevant, standard long day growth conditions. With regard to planting density, the overall density of our plants (2 plants per 2.35 in^2 cell) is equivalent to his density in terms of space; however, plastic barriers more physically divided our plants. We assume that Rédei grew one genotype per pot, although it is possible that he grew the *erecta* heterozygotes in competition with the wild type and homozygous mutant progenitor lines within the same pot. Observations since Darwin's time have described increases of vigor in hybrids when grown in competition with their progenitors (Darwin, 1876) and thus, it is possible that that the *er/+* heterozygotes can outperform the parental lines when grown under such conditions.

Amusingly, Rédei noted that "Arabidopsis is actually not well suited for use in studies of quantitative genetic variation because of its extreme sensitivity to various environmental, physiological effects (Rédei 1962)." Indeed, as sessile organisms, plants are at the mercy of the environment in which they germinate and develop and it is well known that day length, light intensity, temperature, soil composition, planting density and water availability are just a few factors that have a profound effect on plant growth and development. In addition to the effect sowing density had on plant growth (Figure 2.4), we noticed significant differences in flowering time (not shown) and final main shoot height (Figures 2.2 and 2.4) in our experiments with *erecta* mutants from one planting to the next (6 weeks difference; early-September versus follow-up experiments in late-October), which was presumably due to changes in environmental effects such as lower light intensity and shorter periods of natural light (although the greenhouse was supplemented with artificial light in both cases). Such inconsistencies in growth conditions and experimental design were major points of critique in early examples of overdominance in Drosophila (Muller 1940; Muller 1942; Dobzhansky 1950; Mukai and Burdick 1959; Wallace and Dobzhansky 1962; Dobzhansky and Spassky 1968) and our results echo these concerns.

Although we were unable to reproduce the *erecta* overdominance results and extend the analysis to the molecular basis of this effect, there are *bona fide* examples of single-gene overdominance driving heterosis in plants, including an example in tomato from our lab involving dosage of a flowering hormone, florigen (Krieger et al. 2010; Jiang et al. 2013; see Appendix B). This work on *sft*/+ heterozygosity in tomato and the experiments I performed in parallel with *Arabidopsis ft*/+ *tfl* mutants has begun to answer some of the questions regarding the molecular basis of overdominance. For instance, we found evidence for transcriptional changes upon loss of a single copy of a gene and demonstrate that such mutant heterozygosity can significantly alter phenotypic output. Moreover, our experiments show that different alleles can modulate the phenotypic output within a single species, yet while similar manipulations of gene dosage in highly conserved genes do not always translate to overdominance between divergent species. In particular, in tomato, we see that loss of a single copy of *SFT* does indeed result in downstream transcriptional changes in all directions and this reduced dosage of florigen is sensed at every floral transition to have a major impact on the phenotypic output, in this case, changes in flowering time and inflorescence development, which translates to increased fruit yield. Although different alleles of *sft* were not tested in this study, artificial microRNA knockdown of *SFT* transcript confirmed the dosage effects of *sft*/+ heterozygotes and suggests that manipulation of florigen levels can be further fine-tuned and harnessed for improving yield. Identification and exploitation of additional alleles of *SFT* or mutations in other genes in the *SFT* pathway could achieve this desired yield boost. The experiment that I performed in *Arabidopsis* provided key insight into the conservation of dosage effects across species. I showed that while a dosage effect is conserved in *Arabidopsis ft*/+ heterozygotes to partially relieve the early flowering effects of *tfl*, this results in plants with intermediate height between the parental lines and does not translate to heterosis (Figures 2.5 and 2.6). This is likely because the sensitivity to florigen is only sensed once on each shoot in *Arabidopsis,* which does not have multiple flowering transitions as in tomato (Figure 1.3). Therefore, the context of this dosage is crucial, and indeed, while florigen dosage may be useful for improving crops beyond tomato, species-specific tuning will be required to

account for differences such as growth habit. Future work is necessary regarding the fine-tuning of alleles in *SFT* and additional florigen network components in tomato and other crops, and, moreover, the search for additional cases outside of the florigen network may still prove fruitful.

Evidence that single-gene overdominance driving heterosis extends beyond florigen heterozygosity in tomato

It could be argued that heterosis caused by *sft* heterozygosity in an *sp* background in tomato is a special case, because this effect is conditional on the sympodial growth habit of tomato, in which the dosage-dependent suppression of *sp* termination by *sft* happens at each flowering transition and is therefore compounded plant-wide to allows for a significant increase in the number of inflorescences and shoots per plant. This is in contrast to the results presented here in the monopodial plant *Arabidopsis*, which has only a single transition to flowering, and therefore, the same florigen dosage does not have the opportunity to translate to yield heterosis. However, both old and new literature in diverse plants, including monocots and eudicots, suggest that the role of flowering and florigen in heterosis is a more general phenomenon and that dosage effects originating from allelic variation in flowering time genes can drive transgressive variation for yield by subtle quantitative modulation of the plant reproductive transition. For example, a *Sorghum bicolor* mutant exhibiting strongly delayed flowering time in a day-length dependent manner also shows overdominance in the heterozygous state as a result of intermediate flowering time (Quinby and Karper 1946). Although the underlying gene has not been identified, the late flowering phenotype of the mutant implies that *FT* could be involved, or at least a component of the florigen network (Quinby and Karper 1946). In addition, a domestication QTL tracing back to a loss-of-function mutation in an *FT* paralog in sunflower also causes single gene overdominance for flower size (Blackman et al. 2010). It should be noted that the dosage effect here may be more complex than in tomato, potentially involving interactions a paralogous *FT* gene and its ancestor. However, importantly, all of the aforementioned examples point towards a single gene, specifically the gene that encodes the vital mobile flowering signal, florigen.

The dosage effect based on florigen and the manipulation of growth, transgressive or not, seems to be universal and likely to occur in all flowering plants. Indeed, in the diverse growth habits and plant systems from which FT-related heterosis has been observed, it may only be transgressive in certain developmental and environmental contexts, similar to many previously reported cases of full genome, IL, and single gene heterosis (Efron 1974; Li et al. 2001; Luo et al. 2001; Welch and Rieseberg 2002; Semel et al. 2006; Mosher et al. 2007; Krieger et al. 2010). Intriguingly, a similar example of heterosis tracing back to two flowering time genes functioning upstream of *FT* called *FRIGIDA (FRI)* and *FLOWERING LOCUS C (FLC)* was recently documented in *Arabidopsis* (Moore and Lukens 2011).

Intriguingly, while the above examples revolve around flowering pathways, cases of single gene overdominance have been identified for other traits. For instance, in maize, fascinating, albeit inconclusive, support for single gene heterosis involving heterozygosity for deleterious mutations impacting multiple traits beyond flowering has been revealed. Dollinger identified a series of recessive mutants in maize that negatively affect diverse aspects of development and growth, and therefore yield, and he crossed these mutants back to their isogenic inbred parents and observed widespread heterotic phenotypes in the F1s (Dollinger 1985). The heterotic effects from creating these mutant heterozygotes affected multiple aspects of growth, which suggested just a single heterozygous gene could have dramatic pleiotropic impacts, resembling in many ways the heterosis caused by *sft*/+ mutant heterozygosity in tomato. Almost all aspects of yield were affected, including flowering time, plant height, ear size, kernel characteristics and total yield. These findings suggested that maize mutations classically defined as recessive may in fact show dosage-effects in the heterozygous condition, lending more support to the hypothesis that a single heterozygous mutation can drive heterosis through pleiotropic dosage-dependent changes on growth. From these several examples involving simpler genetic contexts of inbred lines, it may be reasonable to assume that dosage effects due to mutations are more ubiquitous in wild populations than previously expected due to the widespread masking of deleterious recessive alleles in nature (Charlesworth and Willis 2009).

Indeed, perhaps the genetic and molecular basis of both hybrid vigor and heterosis traces back not only to dominance complementation, but also to dosage effects and pleiotropy.

Remarkably, single gene overdominance extends beyond the plant kingdom to examples in animals as well. One intriguing example involves overdominance for muscle mass in various mammals including cattle, dogs and humans. Originally described in cattle as the "double muscling" phenotype (Grobet et al. 1997), this increased muscle mass has been directly associated with mutations in myostatin genes, and in some conditions these mutations are beneficial. For instance, whippet dogs that are heterozygous for a particular myostatin mutant allele are more muscular and have increased racing performance, while those homozygous for this mutation have excessive muscle that is detrimental to their athleticism (Mosher et al. 2007). A similar muscular disorder is found in humans: myostatin-related muscle hypertrophy (Schuelke et al. 2004). Similar to the dog whippet, humans that are homozygous for a particular mutation in myostatin (MSTN) have double muscle mass, whereas those heterozygous for the mutant allele have muscle mass intermediate to those individuals with two mutant alleles and those lacking the mutation (Schuelke et al. 2004), again suggesting that genetic dosage underlies the molecular mechanism for mutant single gene overdominance. However, not to leave the impression that the myostatin case is unique, examples of overdominance extend beyond this example including a recent discovery involving horn size in wild Soay sheep in which individuals with one allele type of a gene called relaxin-like receptor2 (*RXFP2*) develop large horns, while those with the other allele develop small horns. Sheep with large horns have the greatest reproductive success, while small-horned individuals have an increased survival rate and it is therefore advantageous in respect to overall fitness to harbor both alleles (Johnston et al. 2013). Cumulatively, these emerging examples support a dosage and network-centric view of mutant single-gene overdominance (Birchler et al. 2010), demonstrate that mutant heterozygosity is likely wide-spread in nature, and are suggestive that such cases may be harnessed for agronomic purposes, perhaps in both animals as well as in plants.

Methods

Plant growth conditions, genotyping, and phenotyping for *erecta*/+ experiments

All mutant lines (*er*-T-DNA-1 (CS800010/SALK044110), *er*-T-DNA-2 (CS800017/SALK066455), L*er*-0 (CS24596), *er*-121 (CS3925) and *er*-123 (CS3927)) and their respective parental progenitor lines (Columbia-0 (CS39005), *LER*+ (CS163), Col(*gl1*) (CS28174) and Ws-2 (CS28827)) were acquired from the Arabidopsis Biological Resource Center (ABRC). Homozygous *er* mutant plants were crossed to their wild type progenitor lines. Individual F1 plants from each cross were self-fertilized to generate F2 populations segregating for *er*/+. *Arabidopsis thaliana* plants were grown in the greenhouse under long day (16h light, 8hr dark) conditions in 32-cell flats with two plants per cell. For the preliminary F1 experiments, multiple wild type and homozygous mutant seeds were tapped into the corner of each cell and then thinned to one seedling per corner a few days after germination. As F1 seeds were limited, more care was taken to sow individual seeds to each corner of the cells. In subsequent experiments, to control for sowing density effects, individual seeds were carefully delivered to the corner of each cell irrespective of genotype in order to avoid growth competition during germination. In all experiments, the sown flats were stratified at 4°C for 4 days before transferring to a long day greenhouse maintained at 21°C and supplemented with artificial light. The amount of water delivered to each flat was carefully monitored; each flat was watered the same amount every 2-4 days and excess water was drained ~4 hours post-irrigation. Controlled irrigation helped prevent moss growth on the soil, which may have competed with the growth of our young seedlings and skewed our quantitative measurements. Genotyping assays were developed for two of five populations (the T-DNA mutant populations). The T-DNA insertions were verified by PCR using primers "Lba1" (5'-TGGTTCACGTAGTGGGCCATCG-3') paired with "*er*-SALK066455-R1" (TGTGTGTGAGAAATGGCTCTG) to genotype for *er*-T-DNA-2, or "LBb1.3" (5'- ATTTTGCCGATTTCGGAAC-3') paired with "*er*-SALK044110-R2" (GCAACGTTGCTGGAGATTAAG) to genotype for *er*-T-DNA-1. Fragments were amplified by

PCR using the following parameters: initial denaturation at 95°C for 3 minutes followed by 34 cycles at 95°C for 20 seconds, 60°C for 30 seconds, 72°C for 60 seconds and ending with a final extension at 72°C for 10 minutes. Wild type versus mutant bands were resolved on a 2% agarose gel. Phenotyping was completed in either the F1 or F2 generation, depending on the experiment. Phenotyping was performed at minimum once per experiment as follows: for the preliminary trial, at 3 weeks post-germination, for the subsequent F1 experiments, at ~3 weeks and ~4 weeks post-germination, and for the F2 experiments, at three time-points: 2, 3 and 4 weeks post-germination. In all experiments, the height of each plant was measured along the main shoot of the plant from where the base emerged from the rosette to top of the shoot. In the preliminary experiment, rosette diameter was quantified by measuring from tip to tip of the most mature opposing true leaves. Flowering time was measured as the number of true leaves to bolting (measured for most experiments) and/or the date of the first open flower (measured in all cases). To measure seed size, *Arabidopsis* seeds of all genotypes were measured with a Nikon SMZ1500 microscope. The seeds were measured using the NIS Elements tools software to automatically detect the outline of each seed and calculate the 2D area. For each measured trait, the mean and standard deviation was calculated for each genotype. The means were compared using a Wilcoxon rank sum test.

Plant growth conditions, genotyping, and phenotyping for *ft/+ tfl1* experiment

All mutant lines were acquired from the Arabidopsis Biological Resource Center (ABRC) and originated from EMS mutagenesis in the Landsberg *erecta* (L*er*) background. Homozygous *tfl1-2* mutant plants were crossed to a moderate (*ft-1*) and strong (*ft-2*) allele of *ft.* Individual F1 plants from each cross were self-fertilized to generate F2 populations segregating for both *tfl1-2* and *ft* mutants. Plants homozygous for the *tfl1-2* mutation and heterozygous for the *ft-2* mutation were self-fertilized to generate F3 populations fixed for the *tfl1-2* mutation and segregating for the *ft-2* mutation. *Arabidopsis thaliana* plants were grown in the greenhouse under long day (16h light, 8hr dark) conditions in 32-cell flats with two plants per cell. Individual seeds were delivered to the corner of

each cell to avoid growth competition during germination. The seeds were stratified at 4°C for 4 days before transferring to a long day greenhouse maintained at 21°C and supplemented with artificial light. Tissue was harvested from young rosette leaves and DNA was extracted using a standard CTAB DNA extraction protocol. The *tfl1-2* and *ft-2* mutations were detected using derivative CAPS (dCAPS) assays. A fragment of *TFL1* was amplified by PCR using the primers "tfl1-2 dCAPS-F" 5'-AAACGTCTCACTTCCTTTTCCTC-3' and "tfl1-2 dCAPs-R2" 5'-AAATGAAAAGAAAGAATAAATAAATTAAA**G**GTAC-3' and a fragment of *FT* was amplified using "ft-2 dCAPS-F2" 5'- CCCTGCTACAACTGGAACAACCTTTGGTG-3' and "ft-2 dCAPS-R2" 5'- AAACTCGCGAGTGTTGAAGTTCTGG**G**GC-3'. Both *TFL1* and *FT* fragments were amplified using a touchdown PCR program: initial denaturation at 95°C for 3 minutes, then 10 cycles at 95°C for 20 seconds, 65°C for 30 seconds (decreased by -0.5°C/cycle), 72°C for 30 seconds followed by an additional 30 cycles at 95°C for 20 seconds, 52°C for 30 seconds, 72°C for 30 seconds and ending with a final extension at 72°C for 10 minutes. Underlined nucleotides in the aforementioned sequences introduce a new restriction site in the wild type PCR amplicons. *TFL1* PCR amplicons were digested using *Kpn*I for 3 hours at 37°C, which cuts wild type but not mutant sequences. *FT* PCR amplicons were digested using *Hae*III for 3 hours at 37°C, which cuts wild type but not the *ft-2* mutant sequences. Wild type versus mutant banding patterns were resolved on a 3% half MetaPhor agarose-half regular agarose gel. Phenotyping was completed in the F3 generation, and we compared *tfl1-2 ft-2* double, *tfl1-2 ft-2/+* and *tfl1-2* single mutants. Homozygous single mutants and wild type L*er*-0 were grown at the same time for comparison. Phenotyping and imaging was performed when the plants completed flowering and inflorescence meristems stopped growing (6-8 weeks after germination). The height of each plant was measured along the main shoot of the plant from where the base emerged from the rosette to top of the shoot. The number of rosette leaves, axillary shoots, siliques, open flowers, and floral buds were also recorded as measures of flowering time and yield. For each measured trait, the mean and standard deviation was calculated for each

genotype. The means were compared using a Student's t-test or a Wilcoxon rank sum test when the phenotypic distribution was not normal.

References

Allison A. 1954. Protection afforded by sickle-cell trait against subtertian malarial infection. *British Medical Journal* **1**.

Anderson JB, Funt J, Thompson DA, Prabhu S, Socha A, Sirjusingh C, Dettman JR, Parreiras L, Guttman DS, Regev A et al. 2010. Determinants of divergent adaptation and Dobzhansky-Muller interaction in experimental yeast populations. *Curr Biol* **20**: 1383-1388.

Bateson W. 1909. *Heredity and variation in modern lights*, Cambridge.

Bikard D, Patel D, Le Mette C, Giorgi V, Camilleri C, Bennett MJ, Loudet O. 2009. Divergent evolution of duplicate genes leads to genetic incompatibilities within A. thaliana. *Science* **323**: 623-626.

Birchler JA, Auger DL, Riddle NC. 2003. In search of the molecular basis of heterosis. *Plant Cell* **15**: 2236-2239.

Birchler JA, Yao H, Chudalayandi S, Vaiman D, Veitia RA. 2010. Heterosis. *The Plant cell* **22**: 2105-2112.

Blackman BK, Strasburg JL, Raduski AR, Michaels SD, Rieseberg LH. 2010. The role of recently derived FT paralogs in sunflower domestication. *Curr Biol* **20**: 629-635.

Bomblies K, Lempe J, Epple P, Warthmann N, Lanz C, Dangl JL, Weigel D. 2007. Autoimmune response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in plants. *PLoS Biol* **5**: e236.

Bomblies K, Weigel D. 2007. Hybrid necrosis: autoimmunity as a potential gene-flow barrier in plant species. *Nat Rev Genet* **8**: 382-393.

Brideau NJ, Flores HA, Wang J, Maheshwari S, Wang X, Barbash DA. 2006. Two Dobzhansky-Muller genes interact to cause hybrid lethality in Drosophila. *Science* **314**: 1292-1295.

Bruce AB. 1910. The Mendelian Theory of Heredity and the Augmentation of Vigor. *Science* **32**: 627-628.

Charlesworth D, Willis JH. 2009. The genetics of inbreeding depression. *Nat Rev Genet* **10**: 783-796.

Chen ZJ. 2010. Molecular mechanisms of polyploidy and hybrid vigor. *Trends Plant Sci* **15**: 57-71.

Crow JF. 2008. Mid-century controversies in population genetics. *Annual review of genetics* **42**: 1-16.

Darwin C. 1868. The Variation of Animals and Plants under Domestication. John Murray, London.

-. 1876. *The Effects of Cross and Self-fertilization in the Vegetable Kingdom.* John Murray, London.

Delneri D, Hoyle DC, Gkargkas K, Cross EJ, Rash B, Zeef L, Leong HS, Davey HM, Hayes A, Kell DB et al. 2008. Identification and characterization of high-flux-control genes of yeast through competition analyses in continuous cultures. *Nature genetics* **40**: 113-117.

Dobzhansky T. 1950. Genetics of natural populations. XIX. Origin of heterosis through natural selection in populations of Drosophila pseudoobscura. *Genetics* **35**: 288-302.

Dobzhansky T, Spassky B. 1968. Genetics of natural populations. XL. Heterotic and deleterious effects of recessive lethals in populations of Drosophila pseudoobscura. *Genetics* **59**: 411-425.

Dollinger EJ. 1985. Effects of visible recessive alleles on vigor characteristics in maize hybrid. *Crop Science* **25**.

East EM. 1908. Inbreeding in corn. in *Reports of the Connecticut Agricultural Experiment Station for Years 1907-1908*, pp. 419-428. Connecticut Agricultural Experiment Station.

-. 1909. The distinction between development and heredity in inbreeding. *The American Naturalist* **43**: 173-181.

Efron Y. 1974. Specific differences in maize alcohol dehydrogenase: possible explanation of heterosis at the molecular level. *Nature New Biology* **241**: 41-42.

Efroni I, Blum E, Goldshmidt A, Eshed Y. 2008. A protracted and dynamic maturation schedule underlies Arabidopsis leaf development. *The Plant cell* **20**: 2293-2306.

Fry JD, Nuzhdin SV, Pasyukova EG, Mackay TF. 1998. QTL mapping of genotype-environment interaction for fitness in Drosophila melanogaster. *Genet Res* **71**: 133-141.

Gore MA, Chia JM, Elshire RJ, Sun Q, Ersoz ES, Hurwitz BL, Peiffer JA, McMullen MD, Grills GS, Ross-Ibarra J et al. 2009. A first-generation haplotype map of maize. *Science* **326**: 1115-1117.

Grobet L, Martin LJR, Poncelet D, Pirottin D, Brouwers B, Riquet J, Schoeberlein A, Dunner S, Menissier F, Massabanda J et al. 1997. A deletion in the bovine myostatin gene causes the double-muscled phenotype in cattle. *Nature Genetics* **17**: 71-74.

Hall JG, Wills C. 1987. Conditional overdominance at an alcohol dehydrogenase locus in yeast. *Genetics* **117**: 421-427.

Hochholdinger F, Hoecker N. 2007. Towards the molecular basis of heterosis. *Trends Plant Sci* **12**: 427-432.

Houle D. 1989. Allozyme-associated heterosis in Drosophila melanogaster. *Genetics* **123**: 789-801.

Jefferson MC, Crumpacker DW, Williams JS. 1974. Cold temperature resistance, chromosomal polymorphism and interpopulation heterosis in Drosophila pseudoobscura. *Genetics* **76**: 807-822.

Jiang K, Liberatore KL, Park SJ, Alvarez JP, Lippman ZB. 2013; see Appendix B. Tomato Yield Heterosis is Triggered by a Dosage Sensitivity of the Florigen Pathway that Fine-Tunes Shoot Architecture. *PLoS Genet* **9**: e1004043.

Johnston SE, Gratten J, Berenos C, Pilkington JG, Clutton-Brock TH, Pemberton JM, Slate J. 2013. Life history trade-offs at a single locus maintain sexually selected genetic variation. *Nature* **502**: 93-95.

Jones DF. 1917. Dominance of Linked Factors as a Means of Accounting for Heterosis. *Genetics* **2**: 466-479.

Koornneef M, Hanhart CJ, van der Veen JH. 1991. A genetic and physiological analysis of late flowering mutants in Arabidopsis thaliana. *Mol Gen Genet* **229**: 57-66.

Krieger U, Lippman ZB, Zamir D. 2010. The flowering gene SINGLE FLOWER TRUSS drives heterosis for yield in tomato. *Nat Genet* **42**: 459-463.

Li ZK, Luo LJ, Mei HW, Wang DL, Shu QY, Tabien R, Zhong DB, Ying CS, Stansel JW, Khush GS et al. 2001. Overdominant epistatic loci are the primary genetic basis of inbreeding depression and heterosis in rice. I. Biomass and grain yield. *Genetics* **158**: 1737-1753.

Lifschitz E, Eviatar T, Rozman A, Shalit A, Goldshmidt A, Amsellem Z, Alvarez JP, Eshed Y. 2006. The tomato FT ortholog triggers systemic signals that regulate growth and flowering and substitute for diverse environmental stimuli. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 6398-6403.

Lippman ZB, Zamir D. 2007. Heterosis: revisiting the magic. *Trends Genet* **23**: 60-66.

Luo LJ, Li ZK, Mei HW, Shu QY, Tabien R, Zhong DB, Ying CS, Stansel JW, Khush GS, Paterson AH. 2001. Overdominant epistatic loci are the primary genetic basis of inbreeding depression and heterosis in rice. II. Grain yield components. *Genetics* **158**: 1755-1771.

Mendel G. 1865. Versuche über Plflanzenhybriden. *Abhandlungen*: 3-47.

Moore S, Lukens L. 2011. An Evaluation of Arabidopsis thaliana Hybrid Traits and Their Genetic Control. *G3 (Bethesda)* **1**: 571-579.

Mosher DS, Quignon P, Bustamante CD, Sutter NB, Mellersh CS, Parker HG, Ostrander EA. 2007. A mutation in the myostatin gene increases muscle mass and enhances racing performance in heterozygote dogs. *PLoS Genet* **3**: e79.

Mukai T, Burdick AB. 1959. Single Gene Heterosis Associated with a Second Chromosome Recessive Lethal in Drosophila Melanogaster. *Genetics* **44**: 211-232.

Muller HJ. 1940. *Bearing of the Drosophila work on systematics*. Clarendon Press, Oxford.

-. 1942. Isolating mechanisms, evolution and temperature. *Biol Symp* **6**: 71-125.

Orr HA. 1996. Dobzhansky, Bateson, and the genetics of speciation. *Genetics* **144**: 1331-1335.

Pnueli L, Carmel-Goren L, Hareven D, Gutfinger T, Alvarez J, Ganal M, Zamir D, Lifschitz E. 1998. The SELF-PRUNING gene of tomato regulates vegetative to reproductive switching of sympodial meristems and is the ortholog of CEN and TFL1. *Development* **125**: 1979-1989.

Quinby JR, Karper RE. 1946. Heterosis in sorghum resulting from the heterozygous condition of a single gene that affects duration of growth. *Am J Bot* **33**: 716-721.

Rédei GP. 1962. Single Locus Heterosis. *Zeitschrift für Vererbungslehre*: 164-170.

Schuelke M, Wagner KR, Stolz LE, Hubner C, Riebel T, Komen W, Braun T, Tobin JF, Lee SJ. 2004. Myostatin mutation associated with gross muscle hypertrophy in a child. *N Engl J Med* **350**: 2682-2688.

Schuler JF. 1954. Natural Mutations in Inbred Lines of Maize and Their Heterotic Effect. I. Comparison of Parent, Mutant and Their F(1) Hybrid in a Highly Inbred Background. *Genetics* **39**: 908-922.

Semel Y, Nissenbaum J, Menda N, Zinder M, Krieger U, Issman N, Pleban T, Lippman Z, Gur A, Zamir D. 2006. Overdominant quantitative trait loci for yield and fitness in tomato. *Proc Natl Acad Sci U S A* **103**: 12981-12986.

Shalit A, Rozman A, Goldshmidt A, Alvarez JP, Bowman JL, Eshed Y, Lifschitz E. 2009. The flowering hormone florigen functions as a general systemic regulator of growth and termination. *Proceedings of the National Academy of Sciences of the United States of America* **106**: 8392-8397.

Shannon S, Meeks-Wagner DR. 1991. A Mutation in the Arabidopsis TFL1 Gene Affects Inflorescence Meristem Development. *The Plant cell* **3**: 877-892.

Shull GH. 1908. The composition of a field of maize. *Am Breed Assn Rep*: 269-301.

Singleton WR. 1941. Hybrid vigor and its utilization in sweet corn breeding. *American Naturalist*: 48-60.

Smith LM, Bomblies K, Weigel D. 2011. Complex evolutionary events at a tandem cluster of Arabidopsis thaliana genes resulting in a single-locus genetic incompatibility. *PLoS genetics* **7**: e1002164.

Springer NM, Stupar RM. 2007. Allelic variation and heterosis in maize: how do two halves make more than a whole? *Genome research* **17**: 264-275.

Tang S, Presgraves DC. 2009. Evolution of the Drosophila nuclear pore complex results in multiple hybrid incompatibilities. *Science* **323**: 779-782.

Torii KU, Mitsukawa N, Oosumi T, Matsuura Y, Yokoyama R, Whittier RF, Komeda Y. 1996. The Arabidopsis ERECTA gene encodes a putative receptor protein kinase with extracellular leucine-rich repeats. *The Plant Cell* **8**: 735-746.

van Zanten M, Snoek LB, Proveniers MC, Peeters AJ. 2009. The many functions of ERECTA. *Trends in plant science* **14**: 214-218.

Wallace B, Dobzhansky T. 1962. Experimental proof of balanced genetic loads in Drosophila. *Genetics* **47**: 1027-1042.

Welch ME, Rieseberg LH. 2002. Habitat divergence between a homoploid hybrid sunflower species, Helianthus paradoxus (Asteraceae), and its progenitors. *Am J Bot* **89**: 472-478.

Wigge PA. 2011. FT, a mobile developmental signal in plants. *Curr Biol* **21**: R374-378.

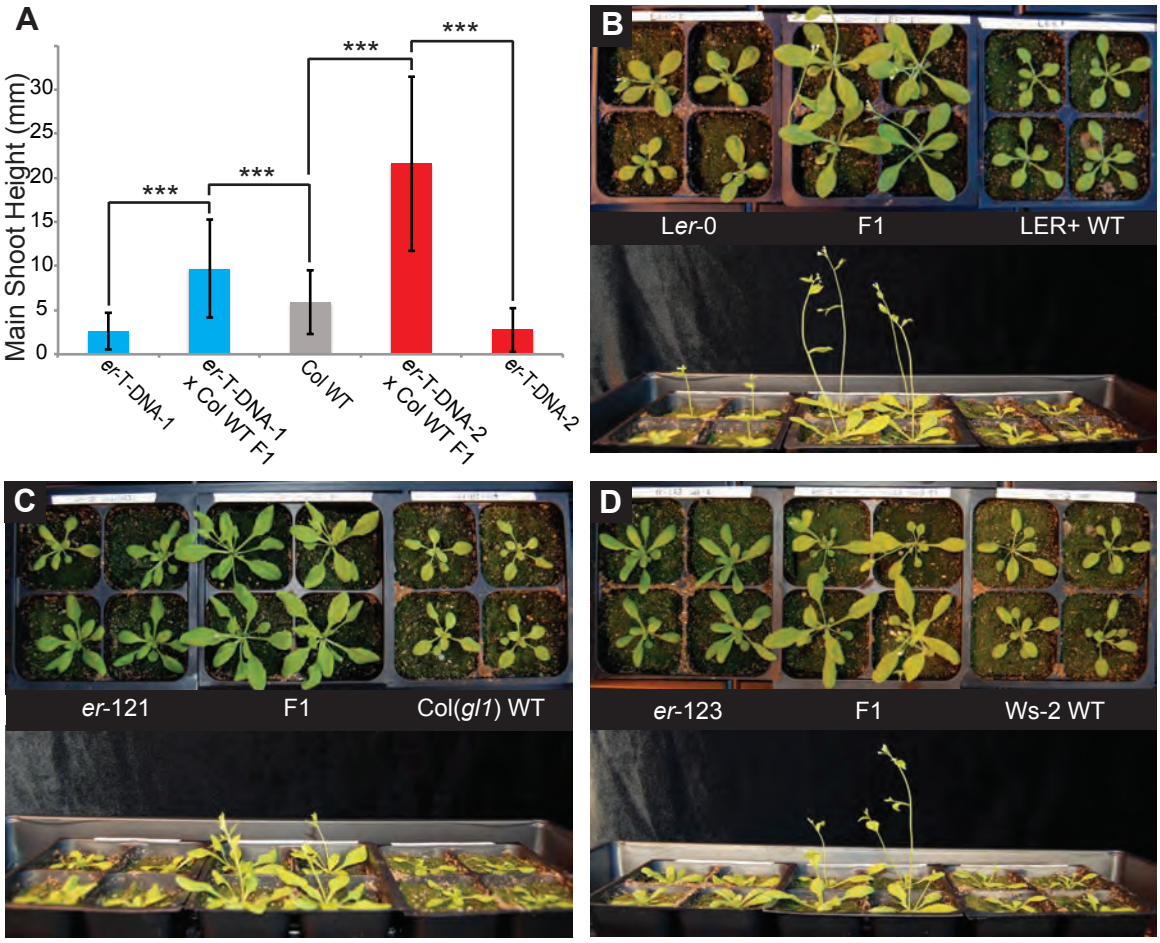


Figure 2.1. Increased bolt height in *er*/+ F1 heterozygotes under dense sowing conditions. (A) Statistic analysis of main shoot height for two independent *er*-T-DNA/+ F1 heterozygotes in comparison to their progenitor lines. Bars indicate average height of the main shoot with standard deviations. Genotypes are shown below. Differences between genotypes were tested by a Wilcoxon rank sum test and significance levels are marked by asterisks (**P < 0.01, n.s. = not significant). Bars indicate average height of the main shoot with standard deviations. Genotypes are shown below. Differences between genotypes were tested by a Wilcoxon rank sum test and significance levels are marked by asterisks (***P < 0.001). (B-D) Representative images from three additional *erecta* lines used in this study. In each panel, the homozygous erecta mutant is shown (left), the F1 (middle) and wild type progenitor (right) is shown from both a top-down view (top) and side view (bottom). While not measured during this preliminary sowing, all of the er/+ F1 heterozygotes appeared to have a taller main shoot height in comparison to the homozygous *er* mutant and wild type parental lines. Anecdotally, the leaf size and rosette diameter also seemed larger in the F1s, however, they were not significantly larger (not shown).

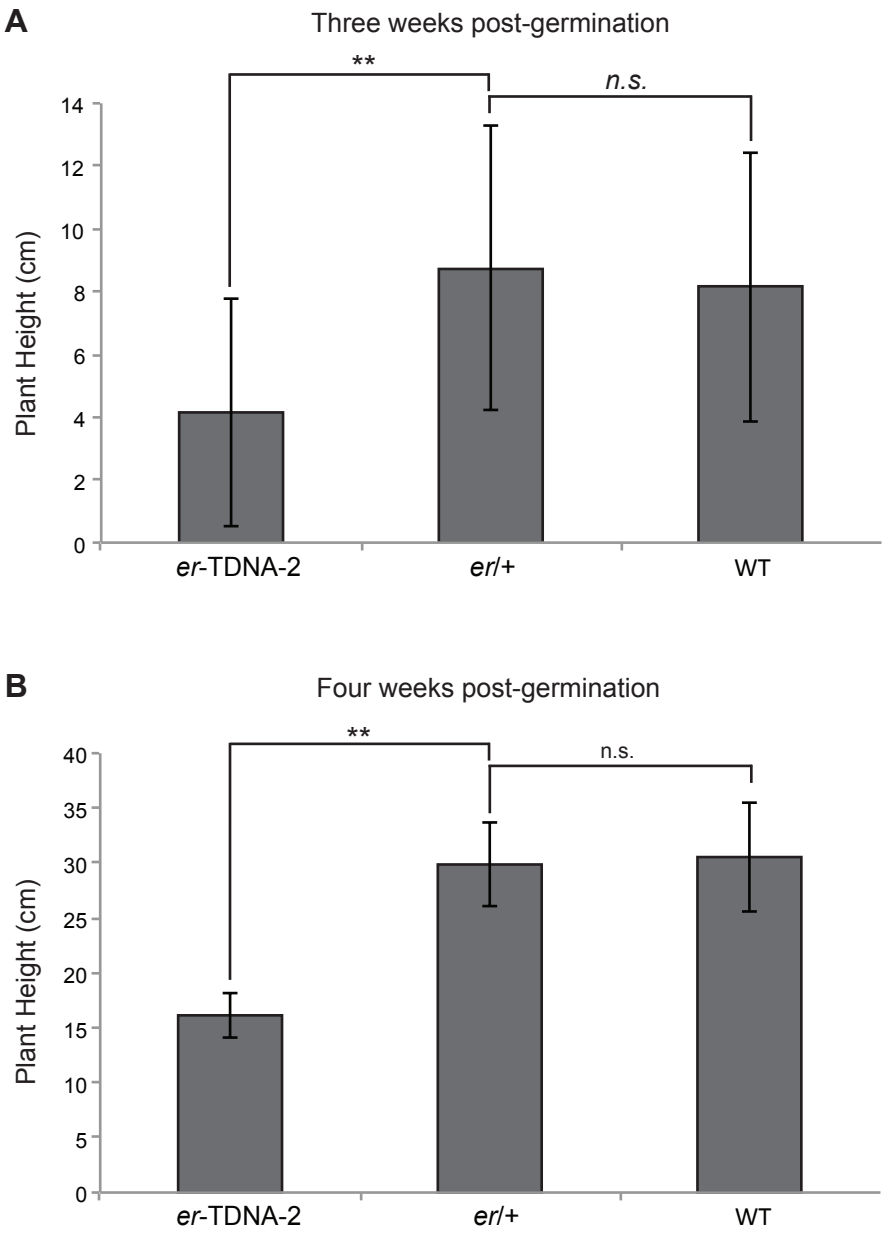


Figure 2.2. Bolt height of *er*/+ F2 heterozygotes are not significantly different from wild type. (A and B) Statistic analysis of main shoot height for er-TDNA-2 F2 families at two different time points. 120 plants per population were phenotyped. Bars indicate average height of the main shoot with standard deviations. Genotypes are shown below. Differences between genotypes were tested by a Wilcoxon rank sum test and significance levels are marked by asterisks (**P < 0.01, n.s. = not significant). (A) An early measurement of plant height (~3 weeks after germination) demonstrates that the *er*/+ heterozygotes segregating from an F2 population are no different from wild type siblings. (B) A later measurement (~4 weeks after germination) towards the end of growth demonstrates that while the homozygous *er*-TDNA-2 mutants remain significantly shorter, the er/+ heterozygotes and wild type siblings remain indistinguishable from one another. Results from the *er*-T-DNA-1 segregating F2 population are not shown, but are indistinguishable from the *er*-T-DNA-2 results shown here.

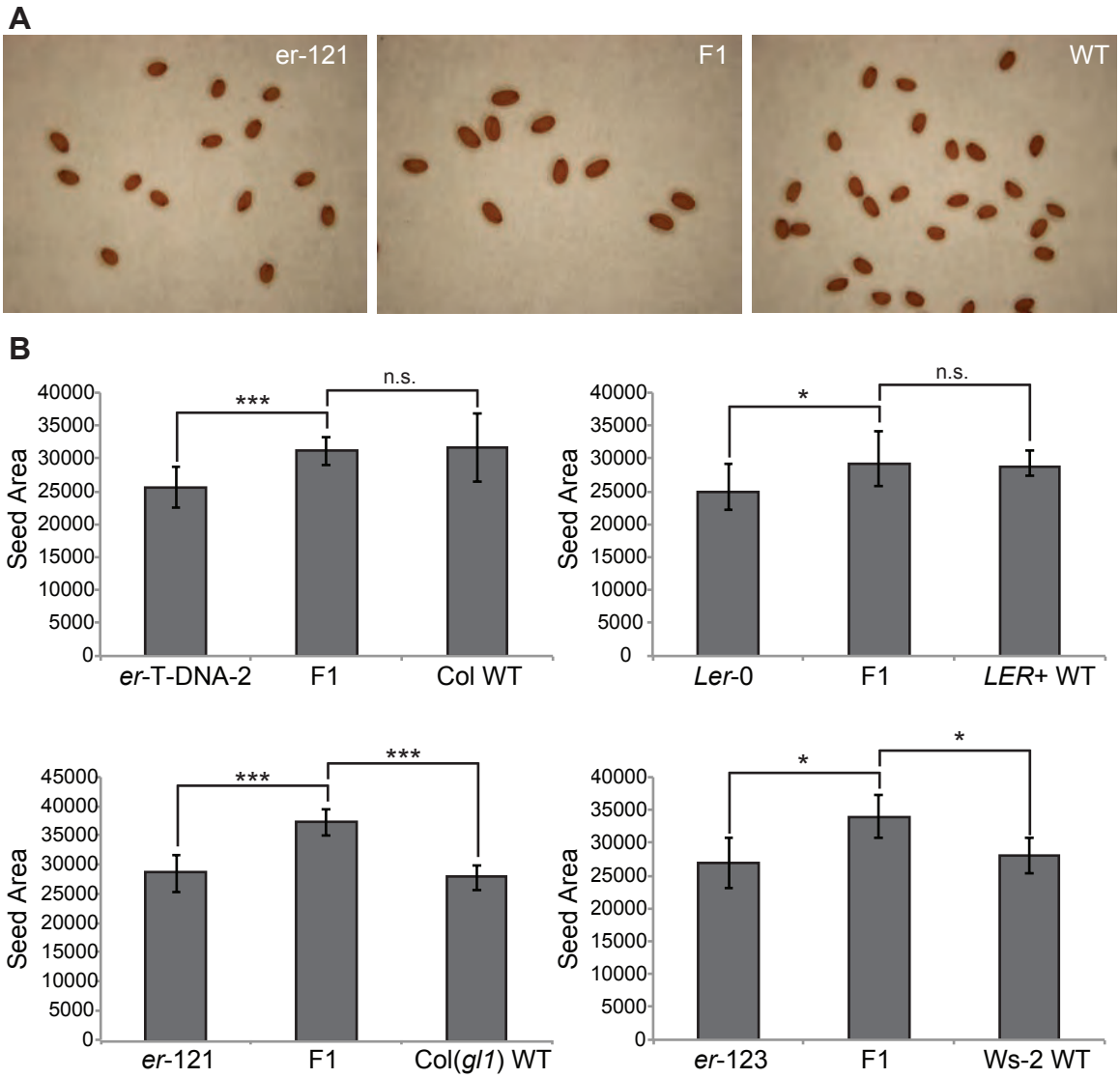


Figure 2.3. *erecta* seed size is not consistently correlated to genotype. (A) Representative images that were used to calculate seed area; homozygous *er*-121 allele (left), *er*-121/+ heterozygous F1 seeds (middle) and the Col(gl1) wild type parental line (right) are shown. (B) Statistical comparisons for seed area in four of the five *er*/+ F1 lines used in this study compared to their progenitor lines show that the F1 seeds are not consistently larger than wild type seeds. Bars indicate average seed area with standard deviations. Genotypes are shown below. At least 10 seeds per genotype for all *er*/+ F1s and their respective homozygous mutant and wildtype progenitor lines were measured. Differences between genotypes were tested by a Wilcoxon rank sum test and significance levels are marked by asterisks (*P < 0.05, ***P < 0.001, n.s. = not significant).

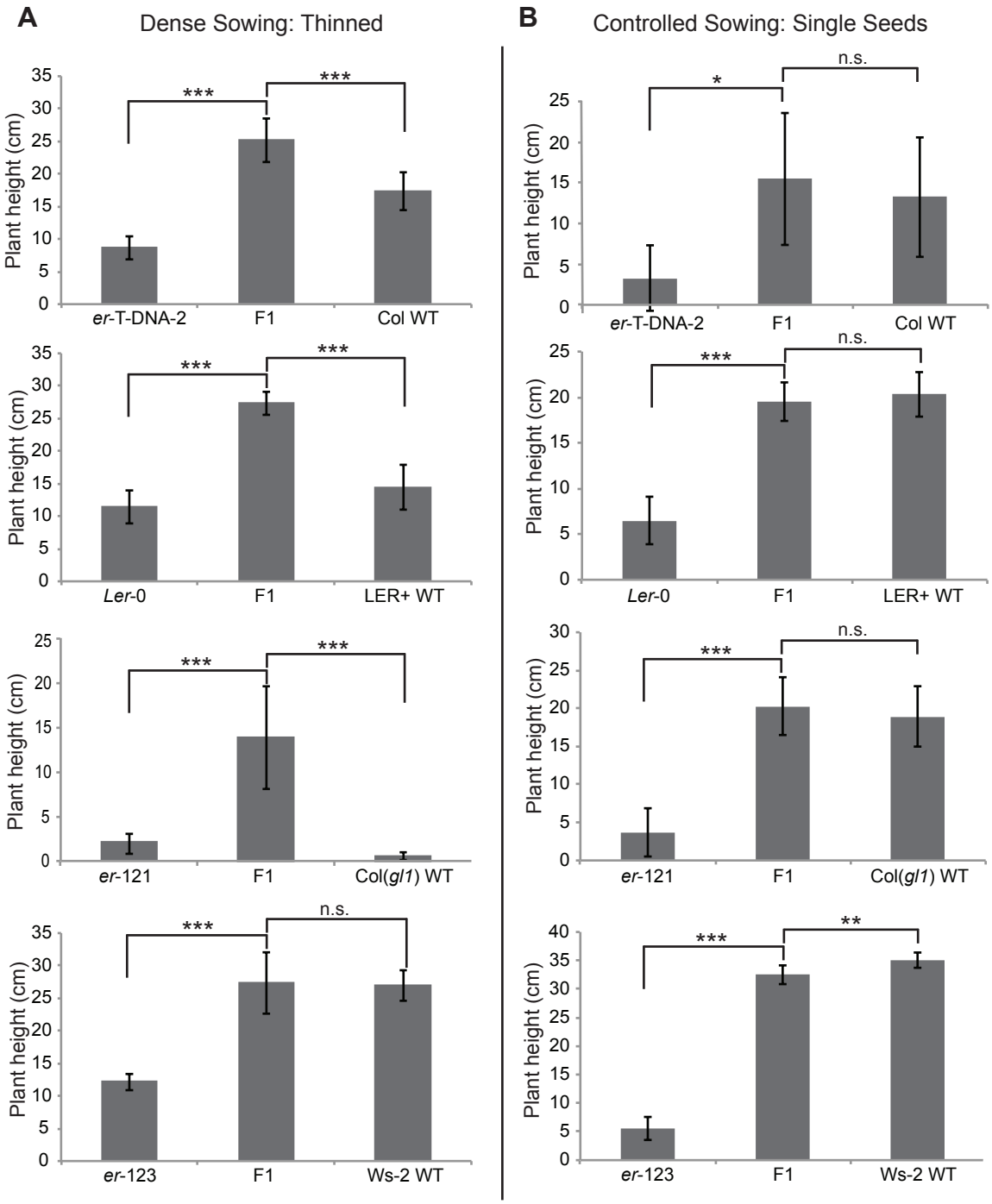


Figure 2.4. Height advantage of er/+ F1 heterozygotes is lost under controlled sowing density. (A and B) Statistic analysis of bolt height for four independent *er*/+ F1 lines under two different sowing conditions. Bars indicate average height of the main shoot with standard deviations. Genotypes are shown below. Differences between genotypes were tested by a Wilcoxon rank sum test and significance levels are marked by asterisks (*P < 0.05, **P < 0.01, *P < 0.001, n.s. = not significant). A height advantage is observed in nearly all *er*/+ F1 plants (except for the *er*-123 allele) sown as single seeds in comparison to their progenitor lines sown under dense conditions and thinned post-germination (A), however, transgressive height of the F1s is lost in all cases when all genotypes are sown as single seeds (B).

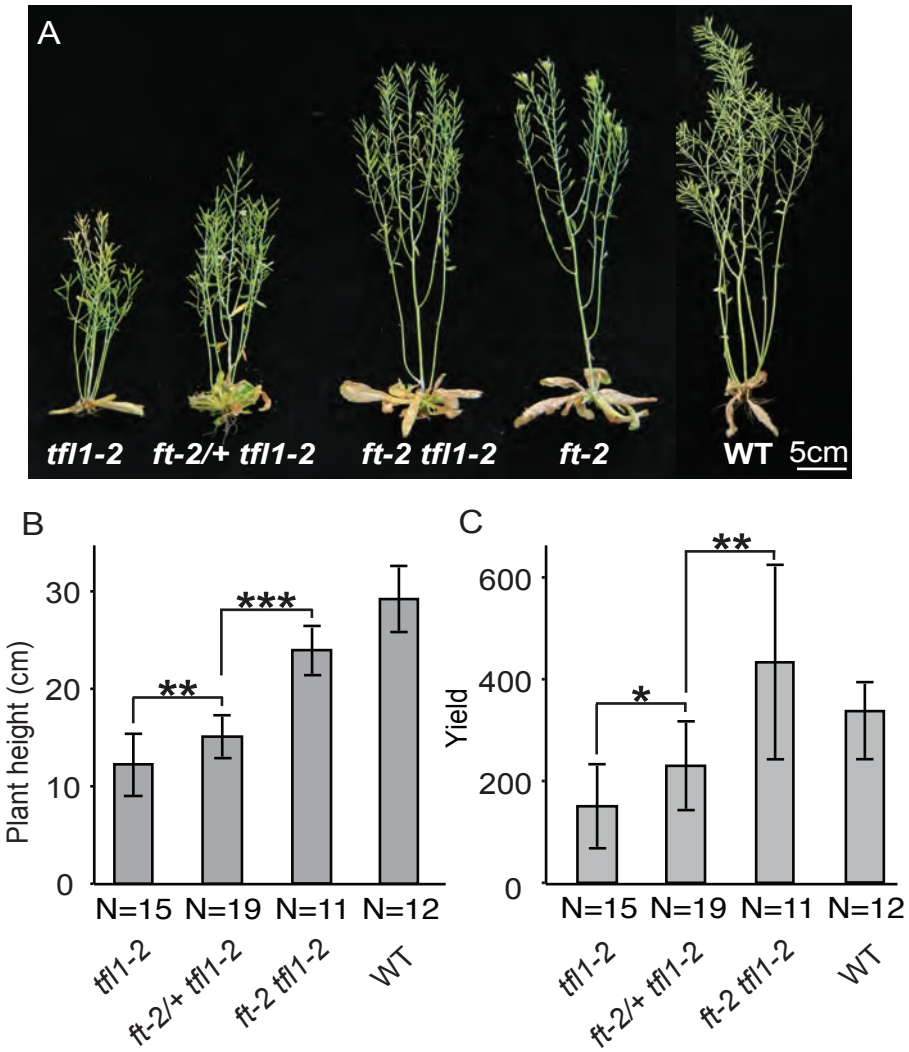
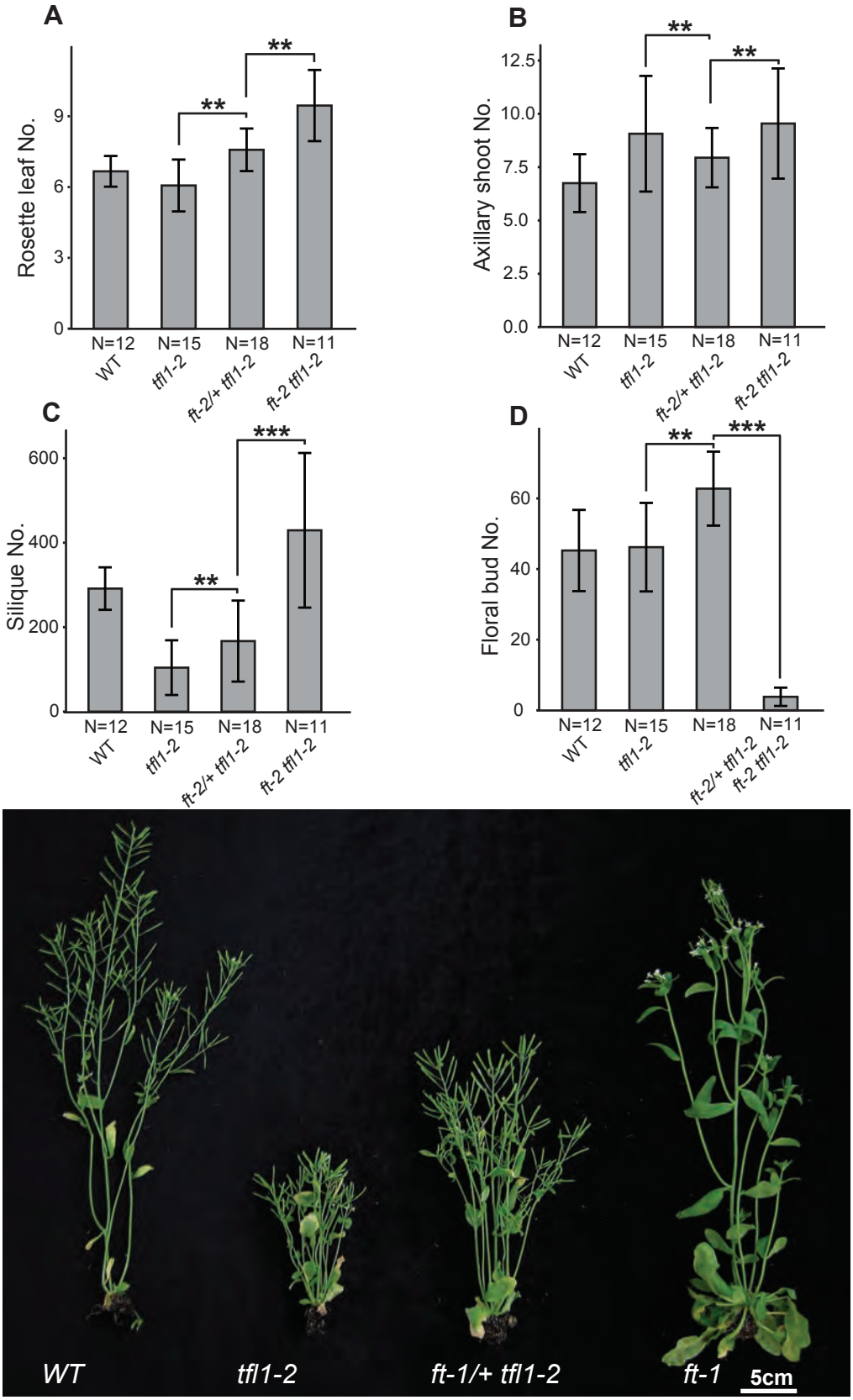


Figure 2.5. Dose-dependent suppression of *tfl1* (*sp*) by *ft/+* (*sft/+*) heterozygosity is conserved in *Arabidopsis thaliana*. (A) Representative plants from left to right of: *tfl1-2* single mutants, *ft-2/+ tfl1-2*, *ft-2 tfl1-2* double mutants, *ft-2* single mutants and wild type *Ler-0* (WT) showing the intermediate height of *ft-2/+ tfl1-2* plants compared to *tfl1-2* and *ft-2 tfl1-2* genotypes. (B-C) Statistical comparisons among all genotypes for plant height and flower/fruit yield showing semi-dominant effects from *ft-2/+*heterozygosity in the *tfl1-2* background. Bars indicate average values with standard deviation. Genotypes and sample size are shown below. Differences between genotypes were tested by a Wilcoxon rank sum test and significance levels are marked by asterisks (*P < 0.05, **P < 0.01, ***P < 0.001). (B**)** *ft-2* heterozygosity in a *tfl1-2* mutant background partially suppresses the early flowering and early termination phenotype of the *tfl1-2* mutation in a semi-dominant manner, resulting in plant height in between *tfl1-2* and *ft-2 tfl1-2* mutant parental lines. (C**)** Unlike tomato, *ft/+* heterozygosity in a *tfl1-2* mutant background does not drive heterosis for yield (number of total siliques and floral buds) in *Arabidopsis*. Rather, yield in the *ft-2/+ tfl1-2* plants is intermediate to *tfl1-2* and *ft-2 tfl1-2* double mutants.

Figure 2.6. Dose-dependent suppression of *Arabidopsis thaliana tfl1* mutant flowering time and yield-associated traits when either strong or moderate mutant alleles of *ft* are heterozygous. (A-D) Statistic analysis of Arabidopsis phenotypes caused by *ft-2/+* heterozygosity in the *tfl1-2* mutant background. Bars indicate average values with standard deviation. Genotypes and sample size are shown below. Statistical significance was tested by a Wilcoxon rank sum test, and significance levels are indicated by asterisks (*P < 0.05, **P < 0.01, ***P < 0.001). (A) Total number of rosette leaves; (B) Total number of axillary shoots; (C) Total number of siliques; (D) Total number of floral buds; (E) Representative plants from left to right of wild type Ler-0 (WT), *tfl1-2* single mutants, *ft-1/+ tfl1-2*, and *ft-1* single mutants. Like for *ft-2*, *ft-1* mutants are completely epistatic over *tfl1-2* mutants, and therefore *ft tfl* double mutants (not shown) are not significantly different from *ft* single mutants (Figure 2.5).


A
Rosette leaf No.
0
3
6
9
**
**
N=12
N=15
N=18
N=11
WT
tfl1-2
ft-2/+ tfl1-2
ft-2 tfl1-2
B
Axillary shoot No.
0.0
2.5
5.0
7.5
10.0
12.5
**
**
C
Silique No.
0
200
400
600
**
***
D
Floral bud No.
0
20
40
60
**
***
WT
tfl1-2
ft-1/+ tfl1-2
ft-1
5cm

3 Characterization of novel inflorescence branching variants in tomato

Introduction

One of the most striking features of flowering plants is the diversity in number and arrangement of flowers born on specialized reproductive branches called inflorescences. Inflorescences can range from simple single-flowered structures to highly complex branched structures with hundreds of flowers. Inflorescence types are generally categorized into a handful of distinct classes such as racemes, cymes and panicles (Weberling 1989; Prusinkiewicz et al. 2007). However, theoretical modeling suggests that while these major classes are most prevalent in nature due to environmental and genetic constraints, these structures are just major aggregation points along a continuum of potential architectures (Prusinkiewicz et al. 2007).

The Solanaceae family provides an interesting opportunity to study the development and evolution of different inflorescence architectures because it includes species which bear inflorescences running the gamut from single-flowered to highly branched multi-flowered structures, and, importantly, the Solanaceae contains many agriculturally important food crops such as potato, tomato, pepper and eggplant, as well as popular ornamentals like petunia. In particular, tomato provides a unique model for investigating inflorescence development, because domesticated tomato and its closely related wild relatives present a range of inflorescence phenotypes from the standard linear truss with approximately 6-8 flowers in domesticated tomato, to simple branching in wild relatives such as in *S. peruvianum*. Mutants identified in domesticated tomato that present a range of phenotypes from single flowers to highly branched structures (Figure 1.4) have been informative (Molinero-Rosales et al. 1999; Lippman et al. 2008; MacAlister et al. 2012), but the genetic cause for a number of inflorescence variants is still unknown. The work presented in this chapter focuses on the characterization and mapping of several previously uncharacterized tomato branching variants ranging from mild to moderately branched.

Inflorescences arise from meristems, which are undifferentiated pools of cells at the growing tips of plants that maintain potential to produce various organ types. Inflorescences can take on different forms, but are in large part directed by meristem determinacy. Monopodial plants such as *Arabidopsis* have an indeterminate growth habit in which the shoot apical meristem (SAM) transitions from vegetative to reproductive growth only once to form an inflorescence meristem (IM). Apical growth continues with an indeterminate IM that forms floral meristems (FMs) laterally along a single axis until growth terminates, resulting in a simple racemose inflorescence structure (Figure 1.3A). Tomato, on the other hand, has a sympodial growth habit in which the primary SAM (PSM) is determinate and terminates by differentiating into the first flower of the first multi-flowered inflorescence (Figure 1.3B). However, prior to this termination, two specialized meristems, that are themselves determinate, are formed to continue growth in unique ways; the sympodial shoot meristem (SYM) forms in the axil of the last leaf made by the PSM and will continue apical, vegetative growth of the plant, whereas the sympodial inflorescence meristem (SIM) forms on the flank of the first FM and will reiterate growth within the inflorescence (Figure 1.3C). Upon transition of each SIM to an FM, there is potential for one additional SIM to form, and this cycle repeats approximately 6-8 times to generate flowers in a zig-zag pattern on a single linear truss typical of domesticated tomato cymose inflorescences (Figure1.3D).

Much remains to be learned about the genetic factors that regulate inflorescence structure in domesticated tomato, and a few recently cloned mutants have been informative. Mutants range from single-flowered, such as the *terminating flower* (*tmf*) mutant (MacAlister et al. 2012), to highly branched structures such as *compound inflorescence* (*s*), *anantha* (*an*) and *falsiflora* (*fa*) mutants (Molinero-Rosales et al. 1999; Lippman et al. 2008); (Figure 1.4). *S* encodes a WUSCHEL homeobox transcription factor, the ortholog of *WUSCHEL HOMEOBOX 9* (*WOX9*) (Lippman et al. 2008). Mutations in *S* cause a delay in meristem maturation, which allows for extra meristem proliferation before committing to floral identity, hence branching in the mature

inflorescence (Lippman et al. 2008; Park et al. 2012). *AN*, the ortholog of the F-box gene *UNUSUAL FLORAL ORGANS* (*UFO*) acts as a transcriptional co-factor of *FA*, the ortholog of the *LEAFY* (*LFY*) transcription factor to activate floral homeotic genes (Molinero-Rosales et al. 1999; Chae et al. 2008; Lippman et al. 2008). Finally, *TMF* belongs to the *ALOG* (*Arabidopsis LIGHT SENSITIVE HYPOCOTYL1*) family encoding a protein of unknown function, and, interestingly, the single-flowered inflorescences observed in the first inflorescence of *terminating flower* (*tmf*) is due to precocious activation of *AN* and *FA* (MacAlister et al. 2012). Together, these genes, amongst others, participate in a "meristem maturation clock" that controls the gradual changes in meristem identity, which allows for variation in inflorescence architecture (Park et al. 2012). Many dynamically expressed genes were uncovered in this study, including *S* and *AN*. *S* normally increases through the vegetative stages and peaks in the SIM and sharply decreases in the FM, whereas *AN* is lowly expressed throughout the transition and comes on strong in the FM. In *s* mutants, *S* expression is delayed and *AN* expression is reduced, allowing for a pause in meristem maturation and proliferation of SIMs. Park et al. (2012) found that *s* causes a substantial delay in maturation during the transition to flowering at two time-points, during the maturation of the PSM and within the SIMs (marked by a delay in *S* expression, and decreased *AN* expression in the FM), which allows for over-proliferation of inflorescence meristems at both stages, and hence, a highly-branched inflorescence develops. However, a more subtle delay in maturation in the wild species *S. peruvianum,* marked by a delay of *S* only in the TM stage, but WT expression in the SIM and slightly reduced *AN* expression in the FM, allows for only one or two additional inflorescence meristems to form resulting a mildly branched inflorescence with one to two bifurcations. Importantly, while retarding maturation allows for increased branching, speeding up maturation could reduce inflorescence complexity, for instance, to single-flowered inflorescences as found on other solanaceous plants like petunia, tobacco and pepper. The precocious activation of *AN* in *tmf* primary shoot meristems, which terminate in a single-flowered inflorescence, provides some evidence for this theory (MacAlister et al. 2012).

The dynamic meristem expression profiles identified by Park et al. (2012) suggest that many other factors contribute to this maturation clock and that characterization of additional inflorescence branching variants in tomato has the potential for uncovering additional components. Identifying such factors is significant on many levels: first, it will improve our fundamental understanding of meristem maturation, second, it may provide insight into the evolution of inflorescence types, and finally, it could also provide targets to harness for agricultural gain.

Of the branched variants for which the genetic basis is known, *s* is the only one that develops viable flowers and produces fruit, at least for the described alleles. However, with regard to improving tomato productivity, the extreme branching of *s* mutants is problematic, because, although additional flowers are produced, much energy is consumed in the making of these flowers resulting in poor fruit set and an overall decrease to plant yield (Lippman et al. 2008). In fact, much of tomato yield improvement has come from increasing fruit size and manipulating plant determinacy as opposed to modulating inflorescence architecture and flower production. During the domestication of tomato, fruit size increased nearly 1000-fold, and a majority of this increase in size traces back to a handful of major effect quantitative trait loci (QTL) (Frary et al. 2000; Tanksley 2004). Tomatoes are grown for two different markets, the fresh market, in which indeterminate varieties are grown primarily in the greenhouse for sale to markets as fresh produce, and the processing tomato market, which grows high-yielding varieties in the field to make processed products (i.e. tomato juice, sauces, catsup, etc.). The most significant improvement in the processing tomato industry came from the implementation of determinate, bushy tomato varieties amenable for high-density growth and mechanical harvesting in the field and, interestingly, this determinacy is a result of a mutation in the flowering repressor *SELF PRUNING* (*SP*), the tomato ortholog of *TERMINAL FLOWER 1* (*TFL1*) (Pnueli et al. 1998). While these determinate *sp* varieties provided an advantage of creating a compact plant with clustered inflorescences that helps synchronize fruit maturation for a once-over mechanical

harvest, a disadvantage is that the early termination of flowering limits the number of inflorescences and hence total fruit yield. However, this determinacy can be partially relieved by manipulating dosage of *SINGLE FLOWER TRUSS* (*SFT*; orthogous to *FLOWERING LOCUS T* (*FT*)), which encodes the flowering hormone florigen, to make a few extra sympodial shoots and inflorescences (Lifschitz et al. 2006; Krieger et al. 2010; Jiang et al. 2013; see Appendix B). While there is potential for further fine-tuning tomato determinacy through manipulations of florigen and members of the florigen network (Lifschitz et al. 2006; Krieger et al. 2010; Jiang et al. 2013; see Appendix B), there is only so much optimization that is possible before plant size becomes too large for processing tomato production. Thus, to provide greater improvements in yield, we need to look beyond fruit size and manipulation of plant growth habit, and one approach is to modestly increase inflorescence branching to slightly boost flower production without compromising fruit set. With this goal in mind, I focused my efforts on characterizing several inflorescence branching variants that are milder than *s*.

This chapter summarizes work on the mapping and characterization of several mild to moderately branched tomato variants. First, I describe mapping and characterization of an interesting subset of variants called *bifurcating inflorescence* (*bif*). *bif* variants are the most weakly branched lines and, interestingly, unlike *s* in which all of the inflorescences are branched, only around 50% of *bif* inflorescences branch and we observed a quantitative range of branching phenotypes in the F2 generation. QTL mapping in *bif* populations uncovered a single major QTL overlapping the chromosomal region containing the *S* gene, mutants of which are highly branched, and, although a causative genetic lesion has not been identified, I suggest that regulatory changes in *S* may underlie the *bif* phenotype. In the second part, I describe mapping of several moderately branched inflorescence mutants called *frondea*, *jointless-2* and *s-like*. While the causative genetic lesions underlying these branching phenotypes have yet to be found, excitingly, none of these mutants map to Chromosome 2, which contains the known branching genes *S* and *AN*. Thus, further pursuit of these mutants should uncover novel genes involved in

regulating inflorescence architecture. Finally, I present results on the mapping and characterization of an interesting mutant that has branched inflorescences, flowers with increased floral organ numbers and enlarged fruits that I called *fasciated and branched* (*fab*). This section ends with the exciting suggestion that a mutation in the tomato ortholog of *CLAVATA1* (*CLV1*) underlies the *fab* mutant phenotype. *CLV1* encodes a LRR receptor kinase known to function in a highly conserved signaling pathway involved in regulating meristem size in other species including *Arabidopsis* (Clark et al. 1997), rice (Suzaki et al. 2004), and maize (Bommert et al. 2005), suggesting a connection between increased meristem size and branching in *fab*. Detailed genetic and molecular characterization of *fab* and a similar mutant, *fasciated inflorescence* (*fin*) are presented in Chapter 4.

Results

QTL mapping and molecular analysis reveals that *COMPOUND INFLORESCENCE* (*S*) may underlie the *bifurcating inflorescence* phenotype

Variation in inflorescence architecture complexity is found in domesticated tomato and its wild tomato ancestors (Peralta and Spooner 2005; Lippman et al. 2008). A collection of over 6000 domesticated tomato varieties called the Tomato Core Collection (TCC) is an important resource for investigating the allelic variation contributing to diverse traits including fruit size, fruit shape, fruit color, plant growth habit, and plant architecture (Lippman et al. 2008). In a survey of the full collection, more than forty unique accessions with highly branched inflorescences were identified from the TCC (Lippman et al. 2008). Genotyping in these lines revealed that nearly all had mutations in *S,* which encodes a WUSCHEL homeobox transcription factor that is the tomato ortholog of *Arabidopsis* WOX9. *S* was recently found to act in

modulating the timing of meristem maturation during the transition to flowering; *s* mutants delay maturation and allow for proliferation of meristems in the inflorescence, and hence, branching (Lippman et al. 2008; Park et al. 2012). However, three additional accessions were isolated from the TCC that presented mild to moderately branched inflorescences, bearing cherry-sized fruits. The molecular basis for this interesting subset of weaker branching variants is still unknown. We called these variants *bifurcating inflorescence* (*bif*), because moderate (approximately 1-4) branching or "bifurcation" events occurred often after a few flowers were produced along a single truss as in wild type (Figure 3.1A). Importantly, allelism tests suggested that *bif* might be allelic to *S* (Figure 3.3A). However, the branching observed in *bif s* F1s was reduced compared to *s* homozygous mutants, suggesting that other factors could contribute to the *bif* phenotype, which warranted further investigation.

Three *bif* lines CC2692 (*bif1*), CC6736 (*bif2*) and CC6757 (*bif3*) were crossed to *S. pimpinellifolium* to generate F2 mapping populations and populations were grown twice and in two different locations. Interestingly, a range of branching phenotypes was observed in the *bif* F2 mapping populations (Figure 3.1A); individual plants had a combination of non-branched "normal" inflorescences, weakly branched inflorescences (single bifurcation) and moderately branched inflorescences (two or more branching events). Due to the quantitative nature of the *bif* phenotype, these populations were treated as quantitative trait locus (QTL) mapping populations. In year one, phenotyping was completed on all three populations, which each had 30 (*bif1*) to 100 (*bif2* and *bif3*) plants each, and on average, more than 50 inflorescences per plant were scored for the branching. I calculated the percentage of branched inflorescences produced per plant and found that the F2 mapping populations contained a distribution of phenotypes from individuals that looked like wild type (less than 5% of the inflorescences were branched) to individuals that resembled homozygous *bif* plants (40-60% of the inflorescences were branched). Interestingly, there were also a few individuals with transgressive branching beyond the homozygous *bif* lines (up to 80% of the inflorescences were branched) (Figure 3.1B). Although not factored in to the

QTL analysis, the degree of branching seemed to increase as the plant aged; primary inflorescences were less often branched than later inflorescences, and the percentage of multiple branched inflorescences consistently increased with age (Figure 3.1C).

Ninety-nine plants from the *bif2* F2 population were genotyped and 60 markers covering 10 of the 12 chromosomes were used for QTL analysis, which revealed a single highly significant QTL peak on chromosome 2 (Figure 3.2A). Unfortunately, regions that lacked polymorphism between the *bif* lines and *S. pimpinellifolium* were revealed throughout the genome, and particularly for large segments of chromosomes 6 and 9. Therefore, I was unable to include these chromosomes in the QTL analysis. To determine if these regions may harbor additional branching QTL, large populations of two *bif* x *S. lycopersicum* var. M82 F2 lines (*bif1* and *bif2*) were grown to continue mapping in these difficult regions. Although these data could not be combined into the QTL analysis, regression analysis was run on chromosome 6 and 9 markers to determine whether there was linkage to the branching phenotype. Simple regressions were also run on the chromosome 6 and 9 markers run on the *S. pimpinellifolium* population that could not be included in the QTL analysis due to low coverage on these chromosomes. Four markers from each of these chromosomes were used to determine if there was any linkage on these chromosomes to the branching phenotype (Figure 3.2C). Regression analysis showed potential linkage at one of the markers on the long arm of chromosome 9, TG348 (Figure 3.2C), but is predicted to explain less than 10% of the phenotypic variation in branching, far less than the chromosome 2 QTL, but still high enough to warrant further investigation. This will need to be verified with additional markers in the same region and in the other populations.

Interestingly, the chromosome 2 QTL spans a region that contains *COMPOUND INFLORESCENCE* (*S*). Due to the similarity in phenotype and preliminary complementation tests suggesting that *bif* and *s* might be allelic (Figure 3.3A), it was not perhaps surprising to find a QTL overlying a locus containing *S*. Interestingly, sequencing identified a non-synonymous SNP in the *S* coding sequence in a non-conserved region C-terminal to the homeodomain (Figure

3.3B); however, while there is a bias towards homozygosity for this SNP in highly branched plants within the F2 population, it did not co-segregate with branching, and some non-branched plants were also homozygous for this SNP (Figure 3.3C). To verify the Chromosome 2 QTL, all three *bif* populations in both backgrounds (crossed to both *S. lycopersicum* var. M82 and *S. pimpinellifolium*) were grown in a second year and in a new location, and phenotyping was completed for all lines as described above. Genotyping was performed for markers flanking the Chromosome 2 QTL identified in the full analysis of *bif2* and for the *bif-S* SNP. Simple regressions confirmed the QTL in all of populations, which is predicted to account for 25-30% of the branching phenotypic variation (Figure 3.2B). I named this QTL *bif*2.1 (as the first branching QTL identified on chromosome 2).

With the goal to identify and focus on large effect QTL, I pressed forward assuming variation at the *S* genomic locus beyond the coding sequence explained the major effect QTL. The mild branching in *bif* is reminiscent of the weak branching observed in wild tomato species such as *S. peruvianum*. Considering the observation that a subtle delay in primary shoot meristem maturation in *S. peruvianum* allows for a few branching events in the inflorescence (Park et al. 2012), I hypothesized that a subtle reduction in *S* expression in *bif* might result in a similar delay in meristem maturation allowing for weak branching. Preliminary expression analysis of *bif-S* was completed using semi-quantitative RT-PCR on *bif* and wild type meristems at four matched stages during meristem maturation. I found that *S* is expressed in *bif* and at the expected developmental time-points. However, importantly, *S* is expressed at lower levels than wild type particularly in the TM stage, and *AN* expression appears to be slightly delayed, consistent with our hypothesis (Figure 3.3D). More detailed and quantitative expression analysis (RNA sequencing) in near isogenic lines (NILs) is necessary to confirm these results. Guided by the hypothesis that there may be regulatory changes in *S*, I also made multiple attempts to amplify the *bif-S* promoter for sequencing, and while not successful at achieving clean results for sequencing, it appears that there may be a slight change in the *S* promoter size between *bif* lines and wild type

(Figure 3.3E), which will require further attention. To further test the hypothesis that *S* underlies *bif*2.1, wild type *S* constructs under its native promoter (S-3kb promoter::*S*gDNA:GFP and S-2.5kb pro::GFP:*S*-CDS) were transformed into *bif2*, *s*-mutant and wild type plants. Unfortunately, these constructs failed to complement *bif2* or *s* mutants, and we did not detect GFP expression in any of these lines (not shown). One explanation is that the promoter was not sufficient to drive proper *S* expression. Alternatively, the fusion proteins were not functional.

While the QTL analysis was informative, only a single major QTL, *bif*2.1, was identified that explained approximately 30% of the phenotypic variation. The small population size and low marker density used in this study limited the resolution of the QTL mapping to large-effect QTL. Smaller effect QTL contribute to the rest of the phenotypic variation, which higher resolution mapping might resolve that are currently masked by the major QTL. Moreover, single-QTL analysis cannot take into account linkage between loci and epistatic interactions, and thus, some smaller-effect QTL are lost because these interactions are not considered. Therefore, I extended my analysis to a two-dimensional two-QTL analysis. In the single-QTL analysis, interval mapping is completed in a one-dimensional genome scan to ask if there is a single QTL in a given interval, ignoring interactions with any other loci. However, in two-dimensional two-QTL analysis, interval mapping is completed with a two-dimensional scan to simultaneously probe for QTL at two separate chromosomal positions. This allows one to ask whether there are genetic interactions, either additive or epistatic, between two loci. As expected, *bif*2.1 was identified as the major-effect QTL by this method; however, three pairs of additional small-effect QTL were detected in the two-dimensional genome scan: additive interactions were found between chromosomes 1 and 5, chromosomes 2 and 5, and chromosomes 2 and 11 (Figure 3.4). With single-QTL mapping, I detected a second peak approaching the threshold for detecting a QTL, suggesting the presence of a small-effect QTL on chromosome 5. However, a significant gap remains between the markers defining the boundaries of this QTL and therefore, this may be an artifact of low marker density. Nonetheless, it is quite interesting that additional small effect QTL

have appeared in this secondary analysis. However, with this low-resolution mapping, the regions underlying the QTL are still large and verification with higher-density mapping or with NILs is necessary. If verified, it will be exciting to complete further fine mapping in these regions to identify the underlying genetic lesions.

While only as single QTL, *bif2.1*, was found, which spans the region containing S. While S is a known gene involved in the regulation of inflorescence branching, my results suggest the possibility that regulatory changes in *S*, which subtly reduce *S* expression during the transition to flowering, allows for more moderate inflorescence branching in comparison to strong *s* mutant alleles. In addition, *bif2.1* only explains approximately 30% of the phenotypic variation, which indicates that additional small-effect QTL can be uncovered with higher resolution mapping. It will be exciting to further investigate both of these possibilities.

frondea/j-2 and *new s-like* do not map to *S*

While *S* may underlie the *bif* variant phenotype, there were several additional moderate inflorescence mutants (*frondea, j-2* and *s-like*) available for mapping, providing an additional opportunity to uncover novel branching genes (Figure 3.5A). Complementation tests revealed that *frondea* and *jointless-2* (*j-2*) were allelic and are herein grouped as one mutant. Complementation tests also indicated that neither *frondea/j-2* nor *s-like* are allelic to *s*. Both *frondea/j-2* and *s-like* phenotypes are more clearly "branched," with all inflorescences branching in the homozygous mutants, unlike *bif* in which there is a quantitative range in branching. Both of these mutants lack a joint at the pedicel of the flower (hence "jointless") and this phenotype has been previously roughly mapped in *j-2* to the centromere of Chromosome 12 (Zhang et al. 2000; Budiman et al. 2004). I found that the jointless phenotype can be separated from the branching phenotype; however, all extremely branched mutants are jointless, suggesting *j-2* may be an enhancer of branching (data not shown). A bulk segregant mapping strategy, using 4 markers per

chromosome in an initial sweep, was used to rough map the mutants. I confirmed that *frondea/j-2* maps to the centromere of Chromosome 12 and also found that *s-like* maps to this region (Figure 3.5B). Thus, the jointless phenotype of *s-like* likely originates from this same locus. Interestingly, neither of these mutants map to *S*, but both carry the SNP found in the *bif-S* coding sequence (Figure 3.5B). This provided further evidence that the SNP in *S* is not the causative lesion underlying the *bif* branching phenotype. Excitingly, as there is no evidence that *S* underlies the *s-like* and *frondea/j-2* branching phenotypes, novel genes involved in inflorescence architecture may be identified from cloning these mutants.

Unfortunately, while a few additional putative loci were found that warranted further investigation, upon deconvoluting the bulk populations into individuals, none of these loci were clearly associated with the branching phenotype (not shown). I considered that the categorization of the mutant class might not have been sufficiently clean. Considering *j-2* as a modifier of branching in these populations, I pre-screened large F2 populations of both *frondea/j-2* and *s-like* (960 plants of each) using markers flanking the *j-2* locus in the hopes of eliminating "jointed" plants. However, the pre-screening did not work as anticipated. Plants with joints on the pedicles and a range of branching phenotypes from weak to extreme were observed. Additionally, many other phenotypes segregated in the background (e.g. leaf size, leaf shape, growth habit). This indicated that further backcrossing to clean up these lines and generation of new F2 mapping populations is necessary before continuing with further mapping of these mutants. The only linked locus that we revealed for both mutants was near the centromere on Chromosome 12 where *j-2* previously mapped (Zhang et al. 2000; Budiman et al. 2004). Fine mapping within this region will undoubtedly be quite difficult, so continuing with a cloning by sequencing approach would probably be more worthwhile. Mapping was difficult on multiple chromosomes due to a lack of polymorphism. As mentioned, the jointless phenotype and branching phenotype can be uncoupled and, therefore, we would also expect to uncover a second locus, perhaps in one of these difficult chromosomal segments, and ultimately the gene underlying the branching

phenotype, using a high resolution sequencing approach.

fasciated and branched (*fab*) maps to the bottom of Chromosome 4

Additional promise for revealing novel genes involved in regulating inflorescence architecture came from an interesting subset of mutants identified in a saturated mutagenesis screen of domesticated tomato (Menda et al. 2004). In addition to mild to moderate inflorescence branching in common with the aforementioned variants, these mutants had additional phenotypes including extra floral organs and enlarged ("fasciated") flowers and fruits. Therefore, we called these mutants *fasciated and branched* (*fab*) and *fasciated inflorescence* (*fin*) (Figure 3.6A). In progeny tests, *fab* segregated as one-quarter mutant, suggesting a single recessive gene. Tissue was collected from 18 mutants and 17 wild type siblings and two pools of DNA were created: wild type and mutant for use in a bulk-segregant analysis mapping approach. Using 4 markers per chromosome, the mutation was roughly mapped to a 3.75Mb region on the bottom of Chromosome 4 (Figure 3.6B) and subsequently verified by deconvolution (Figure 3.6C). Further fine mapping narrowed the interval to 326kb and the strongest candidate gene within this region was a putative LRR-protein kinase predicted to be the tomato ortholog of *Arabidopsis CLAVATA1* (Figure 3.6C). Confirmation that a mutation in *CLAVATA1* underlies the *fab* mutant phenotype is presented in Chapter 4. The second mutant, *fin*, was cloned around the same time by two postdocs in the lab. I followed up with detailed genetic and molecular characterization of both *fab* and *fin*, which is described in Chapter 4.

Discussion

Characterization of several inflorescence branching variants from naturally occurring diversity in domesticated tomato germplasm and mutagenesis populations provided evidence for both known

and potentially novel genes involved in the regulation of inflorescence branching in these weakly branched lines. QTL mapping of moderately branched cherry-type lines identified a single significant QTL on chromosome 2 (Figure 3.2A), *bif*2.1, which spans a chromosomal region containing *S* (*WOX9*). *S* encodes a homeobox transcription factor that when mutated delays the maturation of meristems during the transition to flowering and results in highly branched inflorescences (Lippman et al. 2008; Park et al. 2012). Although a causative lesion in *S* was not identified, *S* remains a strong candidate. It is possible that a regulatory change, perhaps in the promoter, which has been difficult to isolate from the *bif* genotypes, has led to a decrease in *S* expression. While semi-quantitative RT-PCR did not find a change in the timing of *S* expression, *S* expression is lower in the TM stage, and *AN* is also slightly delayed in *bif* compared to WT (Figure 3.4D), similar to *S* and *AN* expression profiles in *S. peruvianum*, which has mild inflorescence branching. Therefore, these expression patterns are consistent with a delay in the progression to flowering that could explain the weak branching in *bif*, but this will need to be verified by more quantitative methods and in NILs. Although this analysis did not uncover new genes involved in regulating inflorescence architecture, the notion that changes in gene expression, and perhaps a subtle decease in the dosage of *S*, could underlie the phenotype is exciting. Consistent with this hypothesis, concurrent experiments in the lab have found that while homozygous *s* mutants produce fewer fruits than wild type due to low fruit set, ~40-50% of the inflorescences in *s*/+ heterozygotes branch just once or twice, reminiscent of *bif* phenotypes, thereby effectively increasing inflorescence number without compromising the ability of the plants to set fruit (KJ and ZBL, personal communication). Thus, similar to the suggestion that dosage levels of the flowering hormone florigen and components of the florigen pathway could be fine-tuned to further modulate inflorescence production (described in chapter 3), these results suggest that fine-tuning of *S* expression levels can provide a subtle boost in inflorescence branching, which may also be optimized to increase fruit yield. Yet, interestingly, the locus that contains *S* only accounts for at most ~30% of the phenotypic variation seen in the *bif* mapping

lines (Figure 3.2B). It is possible that small-effect QTL were not detected in the low-resolution single-QTL mapping. Indeed a handful of small-effect QTL and significant additivity between these small-effect QTL was suggested from the two-dimensional QTL analysis (Figure 3.4). NILs have been made by phenotyping and back-crossing the most highly branched plants in each generation to wild type domesticated tomato. Further characterization in these lines should be more informative.

The second class of genes that I attempted to map, *frondea/j-2* and *s-like* do not map to *S*, suggesting novel genes might underlie the branching phenotypes. Both mutants mapped to the previously defined *j-2* locus on chromosome 12; however, proximity to the centromere has precluded fine-mapping of the gene underlying the jointless phenotype in *j-2* mutants for well over a decade, and a cloning-by-sequencing approach now being carried out in the lab may be a more worthwhile pursuit. The branching can be uncoupled from the jointless phenotype, and I therefore anticipate that a new branching gene will be identified using this strategy.

A recurring challenge when mapping these mutants was that many non-polymorphic regions were uncovered, which are assumed to be due to wild species introgressions present in these backgrounds. The origin of many of the tomato core collection lines are unknown, however, one of the *bif* lines is annotated as "Sweet 100", a popular garden cherry-type variety. Cherry-type tomatoes are classified as *Solanum lycopersicum* var. *cerasiforme* and are admixtures of standard larger-fruited domesticated tomato *Solanum lycopersicum* var. *esculentum* and its closest wild relative *Solanum pimpinellifolium* (Ranc et al. 2008; Blanca et al. 2012). A number of introgressions thought to be from natural hybridization during domestication were identified on chromosome 6 (Labate and Robertson 2012), which is one possible reason why there is a lack of polymorphism in the QTL populations. Another potential source of wild species introgressions is from breeding programs for crop improvement, for instance, a number of introgressions from *S. peruvianum*, possibly from breeding for improved disease resistance, are present on chromosome 9 (Labate and Robertson 2012). Regardless of origin these non-polymorphic regions impeded

genotyping in several areas, which was particularly problematic for achieving sufficient coverage on chromosomes 6 and 9 for QTL mapping of *bif*. Traditional genotyping of QTL populations is rather time consuming and limited by the availability of genetic markers that can distinguish between the parental lines. Furthermore, we have experienced problems with many markers in the cherry-type *bif* populations due to areas that lack polymorphism. Recent utilization of deep sequencing for QTL mapping in rice (QTL-seq) in both recombinant inbred lines (RILs) and F2 populations suggests that bulk-segregant analysis coupled with sequencing (i.e. sequencing pools of the phenotypic extremes) could rapidly detect QTL, and may also help circumvent issues with mapping in regions that seem to lack polymorphism (Takagi et al. 2013). The *bif* segregating populations I have developed provide a nice opportunity to test this method in tomato. For example, by sequencing pools the top 10-20 most highly branched and least-branched individuals from an F2 population, we should at least verify the chromosome 2 QTL and may uncover additional small-effect QTL. This method will provide a much higher density mapping, so should also refine the mapping position for the known QTL. This combined with additional molecular analysis of NILs is crucial to verifying whether mutations in *S* or its regulatory sequences contribute to the *bif* branching phenotype.

Exciting results came from the final mapping project in which I identified the tomato ortholog of *CLAVATA1* as the strongest candidate in the chromosome 4 region to which the *fab* mutant mapped. I reserve further discussion of this mutant for the following results chapter, chapter 4, which is dedicated to detailed genetic and molecular characterization of *fab* in conjunction with a similar mutant, *fasciated inflorescence (fin)*.

Methods

Plant materials, growth conditions, genotyping and phenotyping

The three moderately branched lines were isolated from the tomato core collection (CC2692 (*bif1*), CC6736 (*bif2*) and CC6757 (*bif3*)). F2 populations were generated between inflorescence branching mutants in domesticated tomato backgrounds crossed to the wild species *S. pimpinellifolium* using standard crossing schemes. In the case of *bif* lines, F2 populations were made with *S. lycopersicum* var. M82 sp+ as well. Tomato plants were started in the greenhouse maintained between 65°F (night) to 78°F (day) supplemented with artificial light from high-pressure sodium lamps (50μmol/m^2/sec; 16h light/8h dark) in 72 cell or 96 cell flats, then transferred to the field 4-6 weeks after sowing. For flats that were pre-screened prior to transplanting, young cotyledon tissue was collected into 96-well plates. Otherwise, young apices were collected in the field into eppendorf tubes a few weeks after transplanting. In both cases, DNA was extracted using standard CTAB protocols. A combination of SSR, Indel, CAPS and dCAPS markers designed to detect polymorphisms between domesticated tomato (*S. lycopersicum*) and *S. pimpinellifolium* was used for genotyping. To phenotype *bif* lines, on average 50 or more inflorescences per plant were counted and scored for branching. The percentage of branched inflorescences was calculated for each individual. For some populations, two categories of branching, a single branching event versus multiple branching events, was scored in addition to calculating the total percentage of branched inflorescences. For *frondea/j-2* and *s-like*, plants were scored for both the jointless phenotype (no abscission zone on the fruit pedicle) and for branching in the inflorescence. For *fab*, any plant that had both branched inflorescences and flowers with extra organs were categorized as mutant.

QTL analysis and regression analysis

The *bif* mutants were phenotyped for the percentage of branched inflorescences on each plant and this was used as the quantitative trait for QTL analysis. QTL analysis was performed using R/qtl (Broman et al. 2003) following standard examples (Broman and Sen 2009). Single-QTL mapping was performed using 'scanone' by two methods, Haley-Knott (HK) regression analysis and Expectation-Maximization (EM) analysis, each with 1000 permutations to determine the logarithm of odds (LOD) score threshold. The HK method provides a quick approximation of standard interval mapping using a modified linear-regression algorithm, however it does not deal well with missing genotype information or with large gaps between markers. EM analysis deals is an iterative method of interval mapping and deals with low marker density much better than the HK method. Two-QTL analysis was also performed using 'scantwo' with default settings and 100 permutations to set the LOD threshold. Regression analysis was performed using the statistical software JMP to obtain the R^2 value, which provides an estimate for the percent phenotypic variation explained at the marker.

Transgenic Complementation

Two different *S* transgenic constructs were designed. First, the full genomic wild type *S* gene fused with C-terminal GFP was cloned behind a 3kb native promoter (S-3kb promoter::*S*gDNA:GFP). Second, the S coding sequence fused to N-terminal GFP was cloned behind a 2.5kb native promoter (S-2.5kb pro::GFP:*S*-CDS). Both constructs were transformed into *bif2*, *s*-mutant and wild type plants at the Boyce Thompson Institute using standard transformation techniques. Transgenic plants were transplanted to pots in the greenhouse and phenotyped for branching. The transgene insertion was verified by genotyping for GFP and inflorescences were also checked for GFP expression under the microscope.

References

Blanca J, Canizares J, Cordero L, Pascual L, Diez MJ, Nuez F. 2012. Variation revealed by SNP genotyping and morphology provides insight into the origin of the tomato. *PLoS One* **7**: e48198.

Bommert P, Lunde C, Nardmann J, Vollbrecht E, Running M, Jackson D, Hake S, Werr W. 2005. thick tassel dwarf1 encodes a putative maize ortholog of the Arabidopsis CLAVATA1 leucine-rich repeat receptor-like kinase. *Development* **132**: 1235-1245.

Broman KW, Sen S. 2009. *A guide to QTL mapping with R/qtl*. Springer, Dordrecht.

Broman KW, Wu H, Sen S, Churchill GA. 2003. R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19**: 889-890.

Budiman MA, Chang SB, Lee S, Yang TJ, Zhang HB, de Jong H, Wing RA. 2004. Localization of jointless-2 gene in the centromeric region of tomato chromosome 12 based on high resolution genetic and physical mapping. *Theor Appl Genet* **108**: 190-196.

Chae E, Tan QK, Hill TA, Irish VF. 2008. An Arabidopsis F-box protein acts as a transcriptional co-factor to regulate floral development. *Development* **135**: 1235-1245.

Clark SE, Williams RW, Meyerowitz EM. 1997. The CLAVATA1 gene encodes a putative receptor kinase that controls shoot and floral meristem size in Arabidopsis. *Cell* **89**: 575-585.

Frary A, Nesbitt TC, Grandillo S, Knaap E, Cong B, Liu J, Meller J, Elber R, Alpert KB, Tanksley SD. 2000. fw2.2: a quantitative trait locus key to the evolution of tomato fruit size. *Science* **289**: 85-88.

Jiang K, Liberatore KL, Park SJ, Alvarez JP, Lippman ZB. 2013; see Appendix B. Tomato Yield Heterosis is Triggered by a Dosage Sensitivity of the Florigen Pathway that Fine-Tunes Shoot Architecture. *PLoS Genet* **9**: e1004043.

Krieger U, Lippman ZB, Zamir D. 2010. The flowering gene SINGLE FLOWER TRUSS drives heterosis for yield in tomato. *Nat Genet* **42**: 459-463.

Labate JA, Robertson LD. 2012. Evidence of cryptic introgression in tomato (Solanum lycopersicum L.) based on wild tomato species alleles. *BMC plant biology* **12**: 133.

Lifschitz E, Eviatar T, Rozman A, Shalit A, Goldshmidt A, Amsellem Z, Alvarez JP, Eshed Y. 2006. The tomato FT ortholog triggers systemic signals that regulate growth and flowering and substitute for diverse environmental stimuli. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 6398-6403.

Lippman ZB, Cohen O, Alvarez JP, Abu-Abied M, Pekker I, Paran I, Eshed Y, Zamir D. 2008. The making of a compound inflorescence in tomato and related nightshades. *PLoS biology* **6**: e288.

MacAlister CA, Park SJ, Jiang K, Marcel F, Bendahmane A, Izkovich Y, Eshed Y, Lippman ZB. 2012. Synchronization of the flowering transition by the tomato TERMINATING FLOWER gene. *Nature genetics* **44**: 1393-1398.

Menda N, Semel Y, Peled D, Eshed Y, Zamir D. 2004. In silico screening of a saturated mutation library of tomato. *Plant J* **38**: 861-872.

Molinero-Rosales N, Jamilena M, Zurita S, Gomez P, Capel J, Lozano R. 1999. FALSIFLORA, the tomato orthologue of FLORICAULA and LEAFY, controls flowering time and floral meristem identity. *Plant J* **20**: 685-693.

Park SJ, Jiang K, Schatz MC, Lippman ZB. 2012. Rate of meristem maturation determines inflorescence architecture in tomato. *Proceedings of the National Academy of Sciences of the United States of America* **109**: 639-644.

Peralta IE, Spooner DM. 2005. Morphological characterization and relationships of wild tomatoes (Solanum L. Sect. Lycopersicon). *Monogr Syst Bot, Missouri Bot Gard* **104**: 227-257.

Pnueli L, Carmel-Goren L, Hareven D, Gutfinger T, Alvarez J, Ganal M, Zamir D, Lifschitz E. 1998. The SELF-PRUNING gene of tomato regulates vegetative to reproductive switching of sympodial meristems and is the ortholog of CEN and TFL1. *Development* **125**: 1979-1989.

Prusinkiewicz P, Erasmus Y, Lane B, Harder LD, Coen E. 2007. Evolution and development of inflorescence architectures. *Science* **316**: 1452-1456.

Ranc N, Munos S, Santoni S, Causse M. 2008. A clarified position for Solanum lycopersicum var. cerasiforme in the evolutionary history of tomatoes (solanaceae). *BMC plant biology* **8**: 130.

Suzaki T, Sato M, Ashikari M, Miyoshi M, Nagato Y, Hirano HY. 2004. The gene FLORAL ORGAN NUMBER1 regulates floral meristem size in rice and encodes a leucine-rich repeat receptor kinase orthologous to Arabidopsis CLAVATA1. *Development* **131**: 5649-5657.

Takagi H, Abe A, Yoshida K, Kosugi S, Natsume S, Mitsuoka C, Uemura A, Utsushi H, Tamiru M, Takuno S et al. 2013. QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *Plant J* **74**: 174-183.

Tanksley SD. 2004. The genetic, developmental, and molecular bases of fruit size and shape variation in tomato. *The Plant cell* **16 Suppl**: S181-189.

Weberling F. 1989. Morphology of the inflorescence. in *Morphology of flowers and inflorescences*. Cambridge University Press, Cambridge, UK.

Zhang H-B, Budiman MA, Wing RA. 2000. Genetic mapping of jointless-2 to tomato chromosome 12 using RFLP and RAPD markers. *Theor Appl Genet* **100**: 1183-1189.

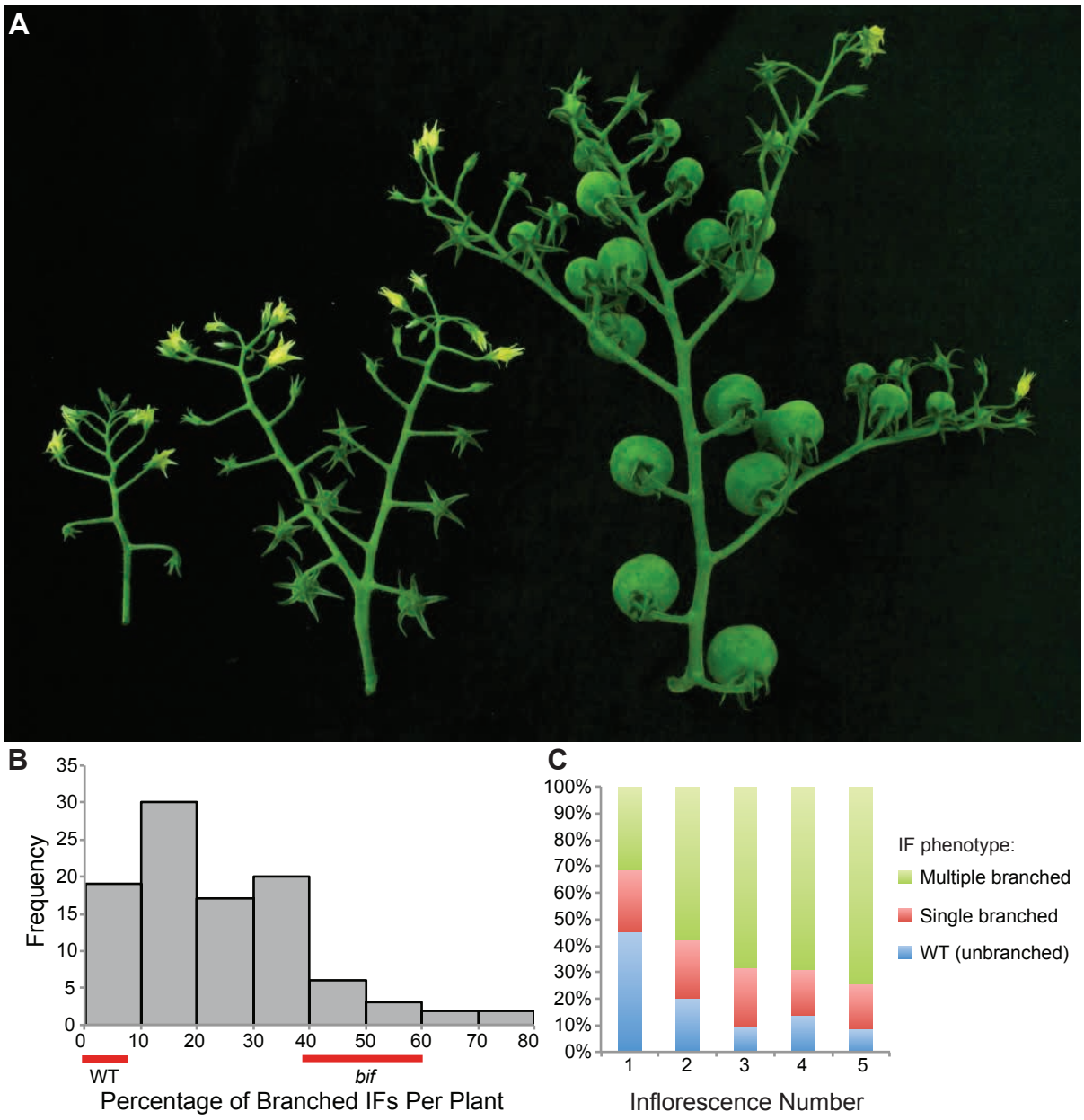
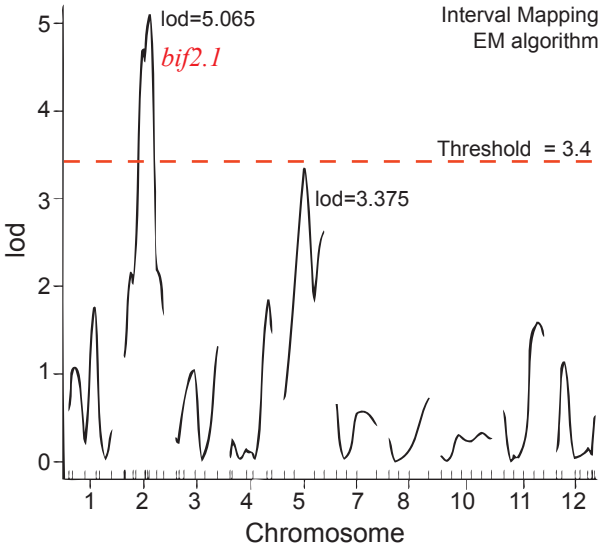


Figure 3.1. A QTL on Chromosome 2 accounts for <30% of the phenotypic variation observed in *bifurcating inflorescence (bif)* populations. (A) An image of inflorescences taken from a single *bif* plant illustrates that inflorescences have a range of phenotypes and may appear "normal" (far left), bifurcate a single time (middle) or have several branching events (far right). (B) The distribution of branching frequency in *bif* F2 plants. *bif* was crossed to *S. pimpinellifolium* to generate an F2 mapping population. Individuals were phenotyped for the percentage of branched inflorescences per plant. The F2 plants have a range of inflorescence phenotypes from plants with mostly "normal" inflorescences, like wild type tomato (<5% of inflorescences are branched), to those with mostly branched inflorescences, like homozygous *bif* (40-60% of inflorescences branched); both of these ranges are underlined in red. Occasionally, F2 plants have branching that exceeds the levels of branching in homozygous *bif* variants (>60%). (C) A summary of branching phenotypes by age. The first 5 inflorescences on at least 5 shoots including the main shoot were scored as wild type (unbranched), single branched or multiple branched. On average, the severity of branching is much less in the first inflorescence formed on each shoot and the proportion of multi-branched inflorescences increases over time.

A



B

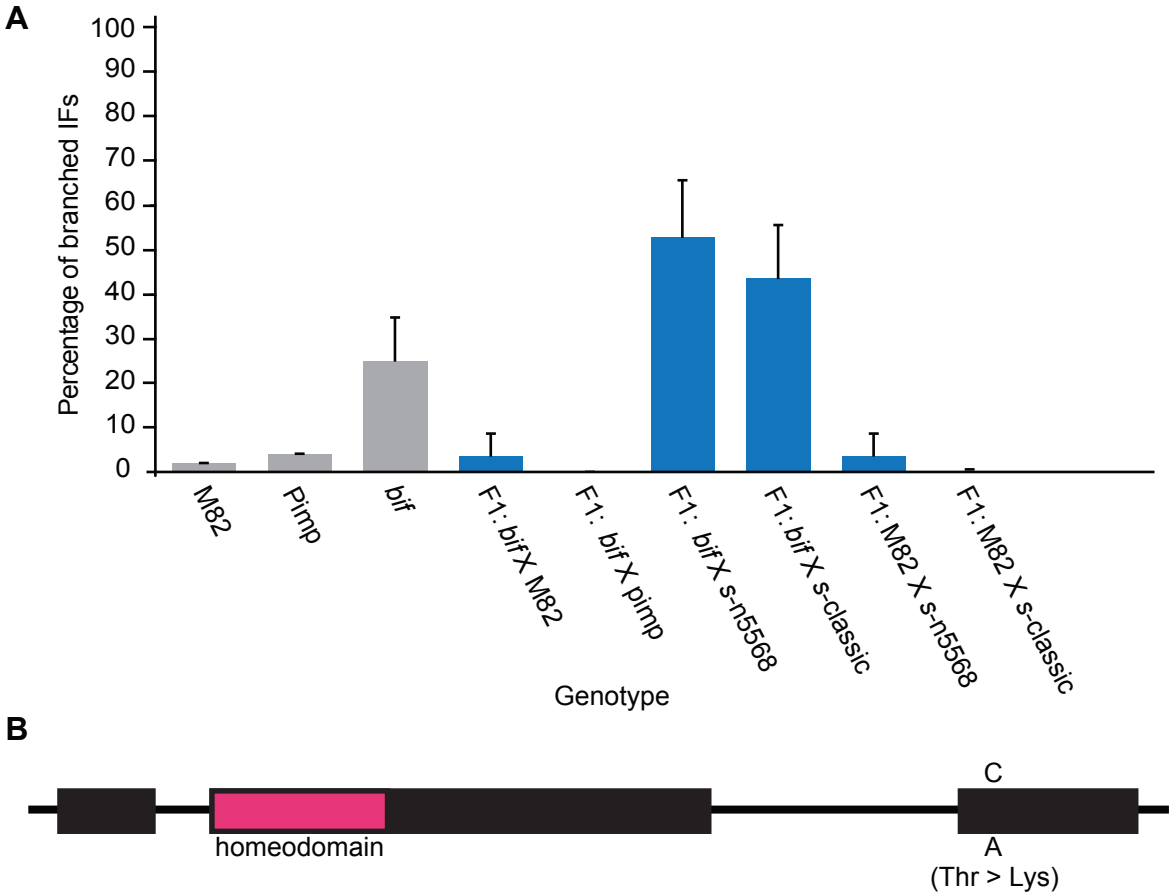
	Percent Variation Explained	
Population	Year 1	Year 2
bif1 x Pimp F2	32.2	6.6
bif1 x M82 F2	N/A	12.1
bif2 x Pimp F2	17.2	28.6
bif2 x M82 F2	N/A	21.6
bif3 x Pimp F2	31.2	N/A

C

Marker	Population	Percent Variation Explained
Chr06_02	*bif2* x *S.pimp* F2	2.8
Chr06_15	*bif2* x M82 F2	1.1
Chr06_5g62530	*bif2* x M82 F2	4.7
Chr06_26	*bif2* x *S.pimp* F2	0.03
Chr09_cosi52	*bif2* x M82 F2	4.4
Chr09_SSR070	*bif2* x *S.pimp* F2	0.4
Chr09_TG348	*bif2* x M82 F2	9.9
Chr09_T0156	*bif2* x *S.pimp* F2	3.5

Figure 3.2. QTL analysis reveals a single highly significant QTL on Chromosome 2. (A) Single-QTL Analysis. A total of 99 plants were genotyped in the *bif2* x *S. pimpinellifolium* F2 mapping population with 60 markers across 10 out of 12 chromosomes. Chromosomes 6 and 9 had large areas of low polymorphism and there was not sufficient marker coverage to include them in this analysis. Standard interval mapping using the EM algorithm was performed with 1000 permutations, which set the LOD cutoff at 3.4. A single major QTL was revealed on Chromosome 2, *bif2.1*, overlapping an interval containing the *COMPOUND INFLORESCENCE (S)* gene. (B) Confirmation of the Chromosome 2 QTL in all populations tested. Regression analysis was performed using the markers flanking the Chromosome 2 QTL and a marker designed for a SNP within *S (bif-S)*. The QTL was confirmed in all populations and accounts for upwards of 30% of the phenotypic variation in branching. The *bif2 x S. pimpinellifolium* F2 population (highlighted in red) was used for the genome-wide scan for QTL. (C) Simple regressions on additional markers on chromosomes 6 and 9. Because there was a lack of polymorphism in the *bif2* x *S. pimpinellifolium* F2 population, simple regressions were run using a combination of the *S. pimpinellifolium* population and the *bif2* x M82 (*S. lycopersicum*) F2 population.

Figure 3.3. Genetic and molecular analysis suggests that *S* may underlie the *bif* phenotype. (A) Allelism tests between the *bif* variant and *s* mutant lines. On average >25% of the inflorescences on *bif* homozygous plants are branched (and reach nearly 60% in some lines and conditions, not shown). Nearly 100% of *s* inflorescences are branched (not shown). *bif* and *s* failed to complement each other for branching; in F1s between *bif* and two separate *s* mutant alleles, more than 40% of the inflorescences were branched. (B) Sequencing revealed a non-synonymous SNP in the C-terminal portion of *S* in *bif* variants *(bif-S*). (C) The *bif-S* SNP is not completely correlated to the *bif* branching phenotype. Results from genotyping for the *bif-S* SNP are shown for the extreme top-10 most highly branched plants (*bif*-like) and bottom-10 least branched plants (wild type-like) from the *bif2* x *S. pimpinellifolium* population. The genotypes are labeled as follows: M = M82 wild type, P = *S. pimpinellifolium*, *bif* = *bif* polymorphism, H = heterozygous. While there is a bias towards homozygosity for the *bif-S* SNP in highly branched plants, non-branched wildtype looking plants also carry homozygosity for this SNP, suggesting that this may not be the causative allele. (D) Semi-quantitative RT-PCR on tomato meristems demonstrates that the level of *COMPOUND INFLORESCENCE (S)* expression and timing of *ANANTHA (AN)* epxression may be altered in *bif*. Four different meristem types were hand-dissected for both wild type and *bif*; late vegetative meristems (LVM), transition meristem (TS), sympodial inflorescence meristem (SIM) and the first floral meristem (FM). The expression level of *S* normally peaks in the TM in wild type. It appears that the expression level is decreased in the TM. *AN* and *SELF PRUNING* (*SP*) were examined as markers of the FM and TM stages to check the fidelity of harvesting the different meristem types. Peaks of expression are seen in the FM for *AN* and in the TM for *SP*, suggesting relatively clean collection of the different meristem types, however *AN* may be slightly delayed compared to WT. *UBIQUITIN* (*UBI*) was run as a loading control. (E) The *bif-S* promoter may be altered from WT. Several attempts were made to amplify the *bif-S* promoter for sequencing. There was trouble specifically amplifying the promoter sequence, from all three *bif* lines (*bif2* and *bif3* shown) however, I did note a possible shift in the size of the promoter, suggesting that in addition to the *bif-S* SNP there may be changes in the *S* regulatory sequences.


A
Percentage of branched IFs
100
90
80
70
60
50
40
30
20
10
0
M82
Pimp
bif
F1: bif X M82
F1: bif X pimp
F1: bif X s-n5568
F1: bif X s-classic
F1: M82 X s-n5568
F1: M82 X s-classic
Genotype
B
homeodomain
C
A
(Thr > Lys)
C
Controls
Bottom 10 branching (wt-like)
Top 10 branching (bif-like)
M P bif H M M M bif M H M H H M bif bif bif bif H bif H H bif bif
XmnI dCAPs:
bif "extremes"
D
wildtype
bif
g LVM TM SIM FM LVM TM SIM FM
S
(3-prime)
AN
(FM marker)
SP
(TM marker)
UBI
E
WT WT bif2 bif3 bif2

1	2	3	4	5	6	7	8	9	10
Chr.	**pos1f**	**pos2f**	**lod.full**	**lod.fv1**	**lod.int**	**pos1a**	**pos2a**	**lod.add**	**lod.av1**
c1:c5	192.0	32.5	11.5	8.17	5.11	30.0	57.5	6.43	3.06
c2:c5	65.0	57.5	11.0	5.95	2.35	65.0	60.0	8.61	3.60
c2:c11	65.0	50.0	12.8	7.81	5.49	52.5	102.5	7.33	2.32

Figure 3.4. Two-QTL analysis reveals evidence for interacting small-effect QTL. A two-dimensional two-QTL scan was performed for the *bif2* mapping population. Two-QTL analysis is an extension of interval mapping that simultaneously asks if there are QTL at two given positions in the genome. This analysis allows one to scan for the probability of interacting QTL and test for additive QTL. Two-QTL analysis was run and 100 permutations were used to set threshold values. Those values that exceed these cutoffs are highlighted in the table above. Two models are considered, the full model (lod.full) and the additive model (lod.add). The full model predicts that there are precisely two QTL if you allow for interactions between the QTL – meaning that the effect at one locus is linked to the genotype at the second locus. The additive model predicts that there are precisely two QTL, but no interaction – meaning that the effect at each locus is the same independent of what the genotype is at the other locus. The first column lists the chromosome pairs (Chr.) resulting from the two-QTL analysis. Columns 2-6 consider the full model and columns 7-10 consider the additive model. Columns 2 and 3 (pos1f and pos2f) give the positions of the QTL on each chromosome for the full model. Columns 7 and 8 (pos1a and pos2a) give the positions of the QTL on each chromosome for the additive model. Using the full model (column 4: lod.full) or the additive model (column 9: lod.add) we see that three QTL pairs are predicted: ch1:ch5, ch2:ch5, ch2:ch11. The ch1:c5 pair and the ch2:ch11 pair have the most significant support and this is true when considering that there is an interaction between the loci (column 5: lod.fv1).

A

B

variant	map to *S* locus on Chr. 2?	*bif* polymorph in *S*?	map to j-2 locus on Chr. 11?
bif	yes	yes	no
fr/j-2	no	*fr* (yes); *j-2* (no)	yes
s-like	no	yes	yes

Figure 3.5. *s-like* and *frondea/jointless-2* do not map to Chromosome 2. (A) Phenotypes of *s-like* (left) and *frondea/jointless-2(j-2)* (right). *frondea* and *j-2* are allelic to one another (not shown). Both *s-like* and *frondea/j-2* have branched inflorescences and lack the fruit abscission zone normally found in the pedicle of the fruit (hence the name "jointless"). A close-up of a *frondea/j-2* flower is inset to show the bract-like structure subtending the flower often seen in *frondea*. Additional leaves are often seen within the *frondea/j-2* inflorescence (main image). (B) Summary of mapping for *bif*, *fr/j-2* and *s-like* variants. As described in the text, *bif* maps to Chromosome 2 on a locus containing *S* and a polymorphism was found in *bif-S*. *frondea/j-2* and *s-like* do not map to *S*, but do both map to the *j-2* locus on Chromosome 11. Interestingly, although neither *frondea/j-2* or *s-like* map to *S*, *frondea* and *s-like* both carry the *bif-S* polymorphism, providing additional evidence that this SNP is not the causative lesion underlying the *bif* Chromosome 2 QTL, *bif2.1*.

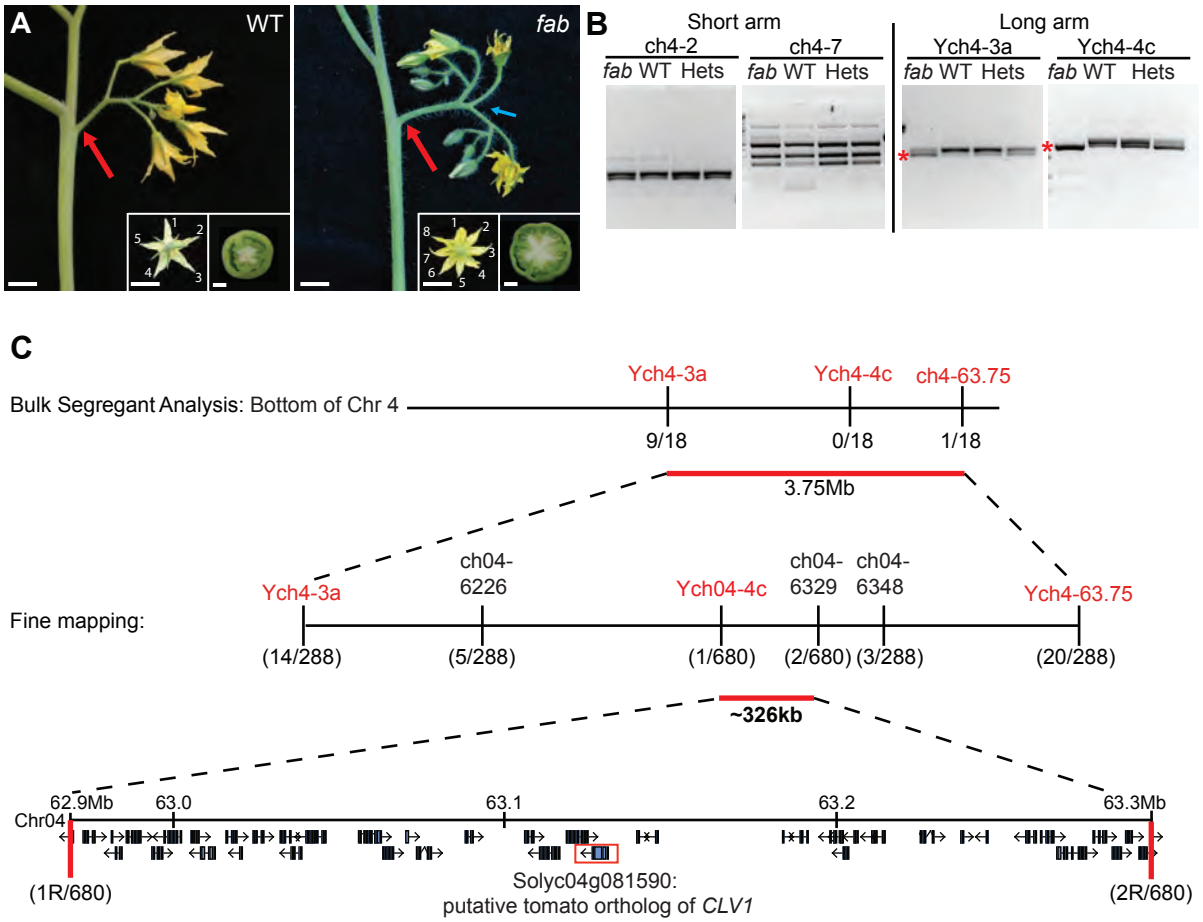


Figure 3.6. *fasciated and branched* maps to the bottom of Chromosome 4 in a region that includes the putative tomato ortholog of *CLAVATA1*. (A) The *fasciated and branched* (*fab*) mutant (right) compared to wild type (left) images illustrates the branched inflorescences and increased number of floral organs characteristic of the mutant. The base of the inflorescence is indicated with a red arrow and the branch point within the inflorescence is indicated with a blue arrow. Images of flowers and a cross-section of an immature fruit are inset in the bottom right corners of each panel. Petal numbers are labeled. Scale bars = 1 cm. (B) Bulk-segregant mapping results for Chromosome 4 markers. The *fab* mutant (domesticated tomato var. M82 background) was crossed to *S. pimpinellifolium* to generate F2 mapping populations. Pools of 18 *fab* mutant plants and 10 wild type sibling plants pulled from the F2 population were interrogated at four markers per chromosome (two per chromosome arm) alongside two different F1s heterozygous plants (Hets) for rough-mapping. Bias towards the domesticated tomato genotype was only detected on the bottom of Chromosome 4 in the *fab* mutant pool indicating possible linkage to the phenotype. (C) Fine-mapping of the *fab* mutation. Deconvolution of the fab mutant pool confirmed a skew towards the domesticated tomato genotype at the bottom of Chromosome 4. Further recombinant screening narrowed the region to ~326kb containing 45 genes. Upon interrogating a list of the annotated genes within this region, the putative tomato ortholog of CLAVATA1 (Solyc04g081590) was identified as a candidate based on the shared mutant phenotype of increased floral organ number observed in *Arabidopsis clv1* mutants.

4 Control of meristem size and shoot architecture by the tomato *FASCIATED AND BRANCHED* and *FASCIATED INFLORESCENCE* genes

Chapter Contributions:

Work presented in this chapter was performed in collaboration with Drs. Cora A. MacAlister (CAM), Ke Jiang (KJ) and Zachary B. Lippman (ZBL) and is in preparation for publication. ZBL performed rough-mapping of *fin*. CAM and KJ completed the mapping and cloning of *fin* by deep cDNA sequencing (mRNA-seq) to identify mutations in the mapping interval and CAM performed Sanger sequencing of five *fin* alleles to verify the mutations. CAM performed semi-quantitative RT-PCR for the tomato *FIN* family members described in Figures 4.9 and 4.16. In addition, CAM made the FIN-YFP fusion that I used for localization studies. CAM made the phylogenetic trees for Figures 4.3 and 4.7. KJ helped with the bioinformatic analysis for the *fab* and *fin* meristem transcriptome profiling. Except as otherwise stated above, I performed the rest of the experiments including all of the work related to mapping and cloning *fab*, the detailed phenotyping of *fab, fin, fab fin* and the rest of the double mutant analyses, the extensive molecular characterization related to both *fab* and *fin* and analyzed the data. I made the figures and wrote the text.

Publications Associated with this Chapter:

Liberatore K.L., MacAlister C.A., Jiang K. and Lippman Z.B. *in preparation*. Control of meristem size and shoot architecture by the tomato *FASCIATED AND BRANCHED* and *FASCIATED INFLORESCENCE* genes.

Abstract

Plant reproductive success and crop yield is largely dependent on the production of inflorescences, flowers, and fruits, all of which originate from pluripotent cells in growing tips called meristems. We report the genetic and molecular characterization of the tomato mutants *fasciated and branched (fab)* and *fasciated inflorescence (fin)* that form branched inflorescences with enlarged flowers and fruits due to defects in the tomato ortholog of *Arabidopsis CLAVATA1* (*CLV1*) and a membrane protein of unknown function, respectively. Both *FAB* and *FIN* are expressed broadly throughout plant development, and mutations in both genes cause expansion of stem cell promoting cells marked by larger meristems and broader domains of expression of the *WUSCHEL* homeobox gene. Despite these similarities, double mutant analysis suggested *FIN* likely functions separately from the classical *CLV-WUS* pathway. Paradoxically, *CLV3* and an additional *CLE* gene are highly overexpressed in *fin* from embryonic stages through the transition to flowering and yet these mutant meristems are grossly enlarged, suggesting that *fin* fails to perceive CLV3p and perhaps additional CLE peptide signals. We show FIN localization and movement in puncta to and from the plasma membrane and hypothesize that FIN may shuttle proteins, including CLV3p receptor proteins to the plasma membrane. Thus, *FIN* adds a new layer of complexity to the mechanisms underlying stem cell control in plants, and provides a novel target for manipulating crop productivity, especially in fruit crops like tomato.

Introduction

Nearly all plant development occurs post-embryonically through the growth and differentiation of self-maintained stem cell niches at growing points of the root and shoot called meristems (reviewed in Aichinger et al., 2012; Barton, 2010; Ha et al., 2010; Stahl and Simon, 2010). Aboveground growth following germination is initiated by an embryonic shoot apical meristem (SAM) that is the progenitor of all vegetative and reproductive shoots and organs, including stems, leaves, inflorescences, and flowers. During the transition to flowering, meristem size gradually increases but remains constrained to finely balance stem cell renewal with commitment of founder cells to lateral organ formation, which is essential for maintaining the ability to continually develop new organs.

Deep knowledge on the control of meristem size has come from 25 years of research in *Arabidopsis thaliana* (reviewed in Aichinger et al., 2012; Barton, 2010; Ha et al., 2010). In particular, genetic and molecular characterization of mutants showing dramatic changes in meristem size exposed major factors and pathways controlling stem cell initiation and maintenance. At one extreme, strong loss-of-function alleles of the *SHOOTMERISTEMLESS (STM)* and *WUSCHEL (WUS)* genes, both genes encoding homeodomain proteins, completely lack a SAM. However, whereas *STM* encodes a class I knotted-like homeodomain transcription factor essential for establishing the SAM during embryogenesis and maintaining stem cell identity thereafter (Long et al., 1996), *WUS* is the founding member of the *WUSCHEL HOMEOBOX* (*WOX*) family of transcription factor genes that functions in parallel to *STM* to promote stem cells throughout development in a classical feedback loop involving the *CLAVATA1, 2,* and *3* genes (Mayer et al., 1998). *CLAVATA1* (*CLV1*) and *CLV2* encode leucine-rich repeat (LRR) membrane-associated receptor proteins that contain and lack an intracellular kinase domain, respectively (Clark et al., 1997; Jeong et al., 1999), whereas *CLV3* encodes a small, secreted peptide ligand (CLV3p) that is bound extracellularly by homomers of CLV1, RPK2/TOAD2, or heteromers of CLV2 with a pseudokinase CORYNE (CRN) (Bleckmann et al., 2010; Clark et al., 1997; Kinoshita et al., 2010; Miwa et al., 2008; Muller et al., 2008; Ogawa et al., 2008). This perception of CLV3p triggers a signaling

cascade involving POLTERGEIST (POL) and PLL1 phosphatase intermediates, which restricts *WUS* expression to a small internal domain underlying the stem cell niche called the organizing center (OC) (Song et al., 2006). Consequently, *CLV* mutations lead to an expanded *WUS* expression domain that causes stem cell overproliferation and dramatically enlarged, or "fasciated", meristems. In a negative feedback loop, the WUS protein promotes expression of *CLV3* non-cell autonomously in cells above the OC called the central zone (CZ) to canalize meristem size (Brand et al., 2000; Schoof et al., 2000).

Recent studies on meristem size in maize and rice has revealed the *CLV-WUS* pathway is highly conserved, particularly the functions of CLV homologs (reviewed in Pautler et al., 2013). For instance, the classical fasciated maize mutants *thick tassel dwarf1* (*td1*) and *fasciated ear2* (*fea2*) are mutated in *CLV1* and *CLV2*, respectfully (Bommert et al., 2005; Taguchi-Shiobara et al., 2001). Both of these mutants produce greatly enlarged tassels with increased spikelet density and ears with increased kernel row numbers. Similarly, the rice mutant *floral organ number1* (*fon1*) - defective in the ortholog of *CLV1* – likewise shows enlarged floral meristems and increased floral organ numbers (Suzaki et al., 2004). A *CLV3* ortholog in rice, FON2, along with another closely related CLE protein, FON2 SPARE1 (FOS1), have also been found to functionally regulate stem cell maintenance (Suzaki et al., 2009).

Yet, despite *CLV* pathway conservation across the monocot-dicot divide, emerging evidence points to a more complex role for *CLV* genes in development. For example, a maize Gα protein CT2, which interacts with the CLV2 ortholog FEA2 and this complex is proposed to transmit CLV signaling in a pathway separate from TD1/CLV1 signaling to restrict SAM size (Bommert et al., 2013a). Likewise, in *Arabidopsis,* additional signaling complexes have been found functioning in the root – in addition to its role in the SAM, CLV1 has recently been shown to work in the regulation of root stem cell maintenance in complex with the non-LRR receptor kinase ARABIDOPSIS CRINKLY4 (ACR4) through the perception of a CLV3 homolog, CLE40, to restrict *WOX5* expression in the root quiescent center (QC), a region overlying the root stem cell niche (Stahl et al., 2013; Stahl et al., 2009). Interestingly, in legumes, the closest homologs of *CLV1* in *Medicago*, *Lotus* and pea and *CLV2* in *Lotus* have been identified for defects in root nodulation, but curiously, mutations in these genes lack meristem defects (Krusell et al.,

2011; Schnabel et al., 2005; Searle et al., 2003). Similar to the recent findings in the *Arabidopsis* root, it is suggested that restriction of *WOX5* expression via CLE signaling is important for regulating cell proliferation within indeterminate root nodules (Osipova et al., 2012). Combined, these emerging examples suggest that WOX/WUS-CLV signaling is common across species and in the regulation of different meristem types; however, there is still a deficit in our understanding of the functional combinations of receptors and signaling molecules, particularly beyond the SAMs of *Arabidopsis*, maize and rice.

Beyond *Arabidopsis* and the grasses, little is known about the genetic factors controlling meristem size and how changes in meristem size influence overall plant growth. However, growing evidence suggests that subtle changes in meristem size can have a profound effect on plant development and crop yield. For instance, during maize domestication there was selection for increased kernel row number (KRN) in the ear, which is linked to increased meristem size and interestingly, a QTL for KRN mapped to mutations in *FEA2*, the maize ortholog of *CLV2* (Bommert et al., 2013b). Similarly, evidence suggests that meristem size was also selected for in tomato domestication, providing the basis for extreme fruit size that has contributed to yield increases over the last century (Cong et al., 2008; Frary et al., 2000; Munos et al., 2011; Tanksley, 2004). For instance, while wild tomato species have 5 sepals, 5 petals, 5 stamens and small currant to cherry-sized fruits with 2 locules (Peralta and Spooner, 2005), the first botanical descriptions of tomato show enlarged flowers with extra organs, enlarged and flattened (fasciated) fruits, and mild inflorescence branching (Mattioli, 1544) resembling the phenotypes of modern large-fruited varieties, suggesting these traits were selected for very early in cultivation. The nearly 1000-fold increase in fruit size in modern cultivars in comparison to wild tomato is attributed to a small number of genes controlling cell cycle and organ number and extreme fruit size is associated with two QTL for increased locule number, *fasciated (fas)* and *locule number* (*lc)* (Lippman and Tanksley, 2001; Tanksley, 2004). Interestingly, a wide survey of domesticated varieties revealed that *fas* maps to regulatory changes in a YABBY-like transcription factor (Cong et al., 2008) and explains nearly all variation for extreme fruit size in modern cultivars (Rodriguez et al., 2011), while *lc* maps to two linked SNPs downstream of

tomato *WUSCHEL* and is responsible for a moderate increase in locule number (Munos et al., 2011). Mutations in a related Arabidopsis *YABBY*, *FILAMENTOUS FLOWER* (*FIL*), causes an increase in meristem size and expansion of the *WUS* and *CLV3* expression domains, and while *YABBY* genes do not have a known function in the CLV-WUS pathway, these phenotypes suggest an indirect connection between the control of meristem size by YABBYs and CLV-WUS signaling that is related to changes in meristem organization (Goldshmidt et al., 2008; Sarojam et al., 2010).

Interestingly, many of these large-fruited varieties also have mild branching within the inflorescence (KLL and ZBL, unpublished data - survey of tomato CC), in which the typical single-trussed inflorescence with 6-10 flowers is converted to a branched structure with a dozen or more flowers. The increase in flower number is directly linked to an increase in fruit number and seed production, yet little is known about the factors that control inflorescence architecture. Therefore, identifying the genetic mechanisms that control inflorescence architecture is fundamentally important for understanding reproductive success and may be useful for manipulating crop yield. In a survey of domesticated tomato varieties with highly compound inflorescences, most variation in inflorescence branching traced back to mutations in *COMPOUND INFLORESCENCE* (*S*), which encodes a WUS-related homeobox transcription factor that is the ortholog of *Arabidopsis WOX9* (Lippman et al., 2008). Similar to *s*, a second mutant called *anantha (an)*, which is defective in an F-box protein, also bears highly branched inflorescences, and *S* and *AN* have been shown to act in sequence to modulate the timing of meristem maturation from an inflorescence to a flower (Lippman et al., 2008; Park et al., 2012). In this study, we describe two new mutants, *fasciated and branched* (*fab*) and *fasciated inflorescence* (*fin*) that have enlarged meristems, fasciated flowers and fruits, and branching within the inflorescence. Double mutant analysis with *FAS*, *S* and *AN* suggests that *FAB* and *FIN* work in yet undefined pathways in tomato to control both increased meristem and fruit size, and inflorescence architecture.

Results

Isolation and characterization of tomato inflorescence mutants showing both branching and fasciation

Sympodial plants like tomato are able to make complex inflorescence structures due to renewed growth from axillary meristems that form at the flanks of floral meristems. In tomato, a primary shoot apical meristem (PSM) first goes through a vegetative phase to produce a number of leaves (~8) prior to transitioning to reproductive growth, in which the PSM will turn into an inflorescence meristem (IM) and terminate in the formation of the first flower. Prior to PSM termination, however, two specialized sympodial axillary meristems are formed, the sympodial shoot meristem (SYM) and the sympodial inflorescence meristem (SIM), which will enable the plant produce additional apical shoot growth and a multi-flowered inflorescence, respectfully. The SYM forms in the axil of the last leaf of the PSM renewing apical growth and producing a few leaves, before itself transitioning to flowering. SYMs are reiterated in this fashion at each flowering event and along each shoot to create a bushy tomato plant. Also prior to PSM or SYM termination in a FM, stem cells are reserved to generate SIMs, which will themselves undergo a transition to form a FM and terminate in the formation of a flower; however, as each SIM transitions to an FM, another SIM is formed on its flank, and this SIM reiteration happens several times to generate a multi-flowered inflorescence.

Most tomato cultivars produce inflorescences with 6-10 flowers arranged in zig-zag pattern along a linear truss, however in some mutants (and wild tomato species) the typical linear truss is converted into a branched structure that contains additional flowers (Lippman et al., 2008). Several such mutants were uncovered in a saturated tomato mutagenesis screen (Menda et al., 2004) and most mutants with highly branched structures are a result of mutations in two genes that control the timing of meristem maturation, *S* and *AN* (Lippman et al., 2008; Park et al., 2012). However, another subset of mutants, *fasciated and branched* (*fab*) and *fasciated inflorescence* (*fin*) develop both mild to moderately branched inflorescences

and enlarged and flattened (fasciated) fruits. Both *fab* and *fin* have similarly branched inflorescences with mild branching (~1-3 branching events) in *fab* inflorescences and moderate branching (~3+ branching events) in *fin* inflorescences (Figure 4.1A-D). Floral organ numbers are increased in all four whorls and both mutants develop enlarged and flattened (fasciated) fruits that convert wild type roma-sized tomatoes to a beefsteak tomato (Figure 4.1E-H). For all examined traits, *fin* mutants are more severe and the phenotype is manifested earlier in development during vegetative growth, resulting in a fasciated vegetative shoot that produces ectopic sympodial shoots, which are not observed in *fab* mutants and a much more grossly enlarged first flower (Figure 4.1D).

The *fab* and *fin* inflorescence branching phenotypes suggest an elaboration of SIMs, but to determine if this was due to a defect in meristem maturation as in *s* and *an* mutants (Park et al., 2012), we performed a double mutant analysis. All mutant combinations were additive or synergistic (Figure 4.2A), suggesting that *FAB* and *FIN* do not act in the known meristem maturation pathways. In addition, double mutant analysis between *fin* and *fas*, a mutant that explains nearly all variation in domesticated tomato fruit size (Cong et al., 2008), also resulted in extreme, synergistic phenotypes (Figure 4.2B). Thus both *FAB* and *FIN* function in pathways not yet described in tomato to regulate both meristem size and inflorescence branching. To pinpoint the developmental origins of the striking morphological phenotypes of *fab* and *fin*, we therefore took advantage of the large and readily accessible tomato SAM to measure morphological changes over time (Park et al., 2012). Soon after germination, the tomato vegetative SAM can be observed as a small flat dome that gradually grows taller and wider during the reproductive transition leading to the first flower of the first multi-flowered inflorescence (Lippman et al., 2008; Park et al., 2012). To determine differences in *fab*, *fin*, and WT meristem sizes, we measured meristems at the early vegetative meristem stage (EVM/8dag; 5^{th} leaf initiated) and the transition meristem stage (TM/15dag; 8^{th} and last leaf initiated prior to flowering) (Figure 4.1I-P) (Park et al., 2012). Both *fab* and *fin* were larger than WT at the EVM stage (Figure 4.1I-L), with more extreme enlargement observed in *fin* that became more pronounced in the TM stage (Figure 4.1M-P). Notably, whereas *fin* meristems grow both taller and wider throughout the transition to flowering, *fab* excess growth is mostly from an

increased meristem height. Moreover, *fin* loses control of normal meristem morphology before the transition to flowering and many ectopic leaves form as the meristem becomes greatly enlarged and disorganized, while *fab* meristem morphology remains relatively normal aside from mild enlargement (Figure 4.1Q-T). The early and more dramatic increase in meristem size in *fin* likely explains the combined vegetative shoot, inflorescence and flower fasciation not observed in *fab* mutants (Figure 4.1A-D). Thus, these combined mutant phenotypes indicate that branching and flower and fruit fasciation in both *fab* and *fin* mutants is due to an increase in meristem size and the early shoot fasciation in *fin* suggests a role for *FIN* very early in vegetative development to restrict overproliferation of the stem cell population.

FAB encodes tomato CLAVATA1

To identify the genes underlying the mutant phenotypes, we first focused on *fab,* which was localized to the bottom of chromosome 4. Fine-mapping positioned *fab* to a 326kb interval encompassing 45 genes, including Solyc04g081590 (Figure 4.3A), which encodes the closest homolog of *CLV1* (Figure 4.3B). Sequencing revealed a missense mutation that converts a highly conserved Alanine within the kinase domain to a Valine (Figure 4.3A), and remarkably, this substitution is identical to the classical *clv1-9* dominant negative mutation from *Arabidopsis* (Dievart et al., 2003). We therefore returned to an M82 isogenic population segregating for *fab* and performed detailed quantitative phenotyping of heterozygous mutants compared to WT and homozygotes and observed a weak but significant effect on sepal, petal, and stamen number (Figure 4.4). Although *fab*/+ heterozygotes showed no inflorescence branching and carpel number was unchanged, these effects were consistent with a dominant negative allele of Solyc04g081590 underlying the *fab* mutant phenotypes.

To confirm that we identified the causative mutation we constitutively expressed the mutant genomic sequence of *fab* under the cauliflower mosaic virus *35S* promoter (*35S::gfab*) in WT plants compared to functional *FAB* (*35S::gFAB*) in WT and *fab.* Importantly, and as predicted for a mutation

acting as a dominate negative, we obtained several first generation transformed (T1) plants showing weak flower fasciation phenotypes with hemizygosity for the mutant allele, whereas *35S::gFAB* completely failed to rescue *fab* phenotypes (Figure 4.3C). Surprisingly, two *35S::gFAB* transformants in the *fab* background exhibited stronger fasciation than in *fab* mutants alone (Figure 4.5A). In these lines, we did not detect reduced expression of *FAB* by semi-quantitative RT-PCR (Figure 4.5B), suggesting that the transgene was not causing co-suppression. Another possibility is that overexpressing the transgene contributes to increased poisoning of FAB protein complexes, perhaps explaining the more severe phenotype. Regardless, the dominant negative effects of *fab* combined with the transgenic experiments demonstrate that *fab* is mutated in *CLV1*, providing the first evidence that the *CLV* pathway regulates meristem maintenance in tomato.

FIN encodes a membrane-localized protein of unknown function that acts separately from the canonical CLV-WUS pathway to control meristem size

Due to the similar phenotypes observed in *fab* and *fin*, we questioned whether *FIN* works with *FAB* to control meristem size via the CLV-WUS signaling pathway in tomato. The *fab fin* double mutants develop grossly enlarged and fasciated meristems that produce an excessive amount of leaves on the primary shoot, which stalls at the transition to flowering (Figure 4.6). Simultaneously, basal axillary meristems are released from dormancy (Figure 4.6) to give rise to shoots that can produce flowers and, rarely, immature fruits that lack seeds (data not shown). This synergistic interaction suggests that *FAB* and *FIN* function in separate pathways to restrict meristem size and thus, *FIN* defines a novel meristem maintenance pathway working in parallel to the canonical *CLV-WUS* pathway.

As with *fab*, a bulk-segregant analysis enabled us to roughly position *fin* to a 1Mb region of chromosome 11, which includes 71 annotated genes (Figure 4.7A). Since a lack of recombination prevented further fine-mapping, we extracted meristem-enriched mRNA from four of our five independent alleles (*fin-e4489, fin-e4643, fin-e9501* and *fin-n2326*) and performed deep cDNA

sequencing (mRNA-Seq) to identify mutations in the mapping interval. A total of 42M Illumina paired-end 100bp (PE100) sequencing reads were aligned to the 1Mb interval and two point mutations were detected in Solyc11g064850, encoding a protein of unknown function. Combined PCR and Sanger sequencing revealed mutations in Solyc11g064850 for all five *fin* alleles: *fin-e4489*, showed a nonsense mutation causing a truncated protein, *fin-e4632* had a missense mutation resulting in a Proline to Serine change, and we were unable to amplify Solyc11g064850 along with varying lengths of flanking DNA from *fin-e9501*, *fin-n2326*, *fin-n5644*, suggesting complete deletions, which RT-PCR confirmed (Figure 4.7A, and data not shown). The *fin* phenotype is also highly reminiscent of *polyopha-2*, a mutant in the wild species, *S. pimpinellifolium* (Stubbe, 1961). On sequencing *FIN* in this mutant, we identified a one bp insertion relative to the wild type sequence disrupting the latter two-thirds of the protein sequence. Solyc11g064850 encodes a 373 amino acid protein belonging to a small, highly conserved, plant-specific gene family. Tomato harbors three homologs of *FIN* and three *FIN-Like* genes were found in *Arabidopsis* (Figure 4.7B). FIN is predicted to have an N-terminal signal peptide in the first 25 amino acids and a single transmembrane domain from amino acids 13-35 (Tusnady and Simon, 1998, 2001).

Consistent with bioinformatic prediction of a transmembrane domain, one of the three *Arabidopsis* FINL proteins (At5g25265) was identified in three independent proteomics studies as membrane-associated proteins (Jaquinod et al., 2007; Marmagne et al., 2007; Mitra et al., 2009). To test whether tomato FIN likewise localized to the membrane, we fused the yellow fluorescent protein (YFP) to FIN and transiently expressed FIN-YFP under the 35S promoter by agrobacterium-mediated infection of tobacco leaves and particle bombardment of onion cells. In both experiments, FIN-YFP was detected at the cell periphery in a punctate pattern, and plasmolysis and counterstaining of the onion cells confirmed membrane localization (Figure 4.7C). Interestingly, we consistently observe movement of these puncta through the intracellular space and what appear to be fusions between individual puncta or with the plasma membrane (Figure 4.8), suggestive of FIN trafficking in vesicles. Although we cannot rule out the possibility that FIN-YFP is aggregating from overexpression, these data are consistent with the predicted (Figure 4.7C) and empirical (Jaquinod et al., 2007; Marmagne et al., 2007; Mitra et al., 2009) membrane

localization of FIN proteins. Intriguingly, although *FIN* and its family members are highly conserved throughout the plant kingdom, the only previously known function for a *FIN* gene is in the regulation of root nodulation in *Medicago truncatula* (Schnabel et al., 2011). Although our results suggest that *FAB* and *FIN* work separately to control meristem size in tomato, curiously, mutations in the closest *Medicago* homolog of *CLV1, SUNN,* also has a role restricted to nodulation and no phenotypes in the inflorescence or flowers (Schnabel et al., 2005), suggesting that both genes may have been coopted into diverse developmental processes that vary by species or that meristem regulation and nodule formation have deep similarity.

FIN restricts *WUS* expression from early in vegetative development to control meristem size

Our genetic analyses suggested that *FIN* functions in parallel to *FAB* and the *CLV-WUS* pathway to control stem cell proliferation. Therefore, to begin to dissect the molecular differences in *FAB* and *FIN* action in controlling meristem size, we explored their expression patterns and dynamics throughout development. RT-PCR on a panel of tomato tissue types revealed that both *FAB* and *FIN* are expressed broadly (Figure 4.9A). Our measurement of meristem sizes revealed enlarged vegetative meristems within 8 days after germination, suggesting that both genes act early in development (Figure 4.1I-L). We therefore took advantage of the tomato meristem maturation transcriptome atlas (Park et al., 2012), to look at the expression of each gene as the meristem matured from vegetative stages (early, mid and late vegetative meristems; EVM/8dag, MVM/10dag and LVM/13dag) through the transition meristem stage (TM/15dag) and upon reaching reproductive stages (FM and SIM/18dag). Both *FAB* and *FIN* are expressed in all meristem types, although *FAB* is expressed at higher levels and the genes have different expression dynamics (Figure 4.9B). *FAB* expression decreases through the vegetative and transition stages and peaks in the FM (Figure 4.9B), which is consistent with *fab* phenotypes manifesting primarily within the inflorescence (Figure 4.1B). *FIN* dynamics are more subtle with relatively stable expression

through the vegetative stages and a peak at the TM, however, expression drops upon the transition to flowering in both the SIM and FM (Figure 4.9B), consistent with *fin* phenotypes manifesting primarily in the vegetative shoot and the first flower of the primary inflorescence (Figure 4.1D).

In *Arabidopsis*, CLV signaling restricts *WUS* expression to the organizing center (OC) - a small population of cells underlying the stem cell niche (Schoof et al., 2000). When *CLV* signaling is disrupted, *WUS* expression expands concomitant with the increase in meristem size in both vegetative and floral meristems (Schoof et al., 2000). To determine if *fab* mutants show similar *WUS* expansion, we probed WT and *fab* mutant meristems for *WUS*. As expected in WT, *WUS* was restricted to the OC of vegetative and floral meristems and we noted only a slightly expanded *WUS* domain in *fab* EVMs that grew substantially larger in floral meristems, consistent with branching and fasciation phenotypes in *fab* being restricted to inflorescences (Figure 4.9C).

The dramatic stem fasciation and increased vegetative meristem size in *fin* mutants and *FIN* expression dynamics described above suggested *FIN* plays a major role in regulating stem cell proliferation as early as the early vegetative phase, potentially through *WUS*. Consistent with this hypothesis, we observed a striking expansion of *WUS* in *fin* vegetative meristems, extending laterally towards the flanks of the meristem (Figure 4.9C). Moreover, multiple intensely stained foci of *WUS* expression were observed in *fin* TMs (Figure 4.9C), likely marking the sites of ectopic SYM and SIM formation that provide the basis for the dramatic overproliferation of vegetative and inflorescence branches during the primary transition to flowering (Figure 4.1D).

Given the dramatic expansion of *WUS* in *fin*, we asked whether the *FAB* expression pattern was also altered. Reminiscent of *Arabidopsis CLV1* expression (Clark et al., 1997), in wild type *FAB* is excluded from the L1 and L2 layers and extends throughout the rest of the meristematic zone from early vegetative stages through the transition to flowering (Figure 4.9C and Figure 4.10A). In developing flowers, again similar to *Arabidopsis* (Clark et al., 1997), but in contrast to maize and rice (Bommert et al., 2005; Suzaki et al., 2004), *FAB* expression is restricted to the floral meristem and is not detected in developing lateral organs (Figure 4.10A). *FAB* expression did not change in *fab* or *fin* mutants suggesting

there is no direct feedback from the increased *WUS* expression on *FAB* expression in either mutant. We also tested whether *FIN* was spatially regulated; however, both full and partial probes failed to reveal a clear signal in multiple attempts (Figure 4.10B). Thus, *FIN* expression could be uniform and diffuse; however, signal was close to background, suggesting a lack of detection sensitivity, which could be a reflection of low transcript levels (Figure 4.9B and Figure 4.10B).

Expression profiling reveals that the CLV3-WUS balance is not maintained in *fin* meristems

That the *WUS* expression domain expands in *fin* does not necessarily mean that *FIN* works directly in a signaling pathway to regulate *WUS* expression, nor does it demonstrate whether the number of stem cells *per se* is increased in *fin* mutants. To further investigate the relationship of *FIN* to the *CLV* pathway, and specifically how *FIN* might function separately from *FAB* in stem cell proliferation, we performed RNA-seq on WT, *fab,* and *fin* vegetative apices at the EVM stage when meristem morphologies are most similar (Figure 4.1I-K). Genome-wide, if *FIN* functions in parallel to *FAB* and the *CLV* pathway, it might be expected that a majority of differentially expressed (DE) genes comparing *fin* mutants to WT would differ from those comparing *fab* to WT. However, as both *FAB* and *FIN* act to regulate meristem size, we also expected to see some overlap in the types of DE genes – these genes could be universally important for meristem growth and maintenance, such as genes involved in the cell growth and division. Thus, we first evaluated global expression changes by asking how many DE genes are unique to each mutant versus how many overlap between the two mutants. We found 1194 DE genes in *fin* and 679 DE genes in *fab* (log-fold change cutoff of ±1 in expression) compared to wild type. While a majority of DE genes were unique to either *fin* or *fab,* there was also a substantial overlap of 30-40% in common to both mutants (Figure 4.11A). Using MapMan classification (Thimm et al., 2004; Usadel et al., 2006), which groups genes by broad biological processes, we determined the distribution of DE genes across functional categories and used a Fisher's exact test to determine whether there was an over-representation of DE

genes assigned to particular biological processes. We found very few categories were significantly over-represented in *fab*, which suggests that at the vegetative stage, there is not much difference between the transcriptional identity of *fab* and wild type meristems. In both mutants, a few functional groups were over-represented, although there were not striking differences between the two mutants (Figure 4.11B, Figure 4.12 and Supporting Information).

While MapMan analysis provides an overview of the global functional expression changes occurring in the mutants compared to WT, it does not consider individual genes, direction of change in expression (under or overexpressed), or the level of expression change. Taking a more direct view of the data, we looked at the top 25 most highly upregulated and downregulated genes in both *fab* and *fin* mutants (Table 4.1 and Table 4.2). Among the DE genes we identified *CLV3* was highly upregulated in *fin* mutants (logFC >5 or ~40-fold increase according to normalized counts; Table 4.2 and Table 4.4). Although not in the top 25 overexpressed genes, it is also increased in *fab* mutants, but to a much lesser extent (logFC ~2.3, normalized fold change ~5; Table 4.1 and Table 4.3). We confirmed this upregulation in both mutants by qRT-PCR (Figure 4.11C). Interestingly, RNA sequencing and confirmation by qRT-PCR revealed an additional putative *CLE* that is highly overexpressed in *fin* (Table 4.4 and Figure 4.11C). This putative *CLE* has a conserved motif similar to the *Arabidopsis CLV3* functional dodecapeptide (Figure 4.13). It is yet unclear whether a mature CLE peptide is generated from this protein and whether it functions as a signaling peptide in the shoot apical meristem, however, it does suggest that *FIN* has an effect on CLE signaling that extends beyond *CLV3*.

Elevated *CLE* levels could be a consequence of increased meristem size and an expansion of the stem cell population; however, at the vegetative stage, *fab* meristems are about twice the size of WT, and *fin* meristems are only about three times larger (Figure 4.1I-K), and, therefore, the dramatic increase of *CLV3* and the putative *CLE* transcript levels in *fin* mutants cannot be explained by the increase in meristem size alone. However, the expansion of the *WUS* expression domain we see in *fin* (Figure 4.9C) makes it is difficult to draw conclusions about *CLV3* and *WUS* expression since changes in expression of these genes are known to influence one another. Therefore, using genes identified in *Arabidopsis* that

were found to specifically mark the *WUS* and *CLV3* expressing cells, but that do not fluctuate in expression levels depending on *WUS* or *CLV3* expression (Aggarwal et al., 2010), we can uncouple meristem size from *CLE* expression levels. We performed qRT-PCR on tomato homologs of three of these genes, and found that these additional meristem markers are not dramatically changed compared to wild type (Figure 4.14). This suggests that the size of the stem cell niche is proportional to overall meristem size and that while *WUS* levels are increased relative to the size of the meristem, the increases in expression of *CLV3* and the putative *CLE* exceed the changes in expression that would be expected based on the changes in meristem size alone.

Although *CLV3* expression may vary within a tolerable range (Muller et al., 2006), elevated *CLV3* levels typically cause a rapid decrease in *WUS* expression and development of smaller meristems, or in extreme cases meristem termination (Brand et al., 2000). This is in contrast to *fin* mutants, which have grossly enlarged meristems and an increase in *WUS* expression (~2 to 3-fold higher; Table 4.1 and Figure 4.11C). *WUS* levels can recover and are sometimes higher than normal after elevated *CLV3* levels, suggesting that other factors perceive meristem size and can independently regulate *WUS* expression or control meristem size in a *WUS*-independent manner, but when high levels of CLV3 persist, the meristems still respond to increased CLE levels and are reduced in size (Muller et al., 2006).

In *Arabidopsis*, *CLV3* transcription is localized in a wedge of expression overlapping the stem cell niche and mature CLV3 peptide is exported to the apoplast and perceived by CLV1 and other receptors to trigger the signaling cascade to restrict *WUS* expression, which feeds back to regulate *CLV3* expression (Brand et al., 2000; Schoof et al., 2000). To determine if the increase in *CLV3* expression is localized or if the expression domain is expanded in *fin* mutants, we performed *in situ* hybridization in vegetative meristems. *CLV3* is barely detectable in WT vegetative meristems, however, strong signal is rapidly detected in *fin* mutants (Figure 4.15). In WT meristems, *CLV3* is expressed in just a few cells overlying the *WUS* domain, but unlike in *Arabidopsis* (Brand et al., 2000), it was not detected in the L1 or L2 layers, which could be a result of low transcript levels in WT. However, in both *fab* and *fin* mutants, the expression domain is expanded laterally towards the flanks of the meristem and extends upwards into

the L2 and possibly the L1 layers (Figure 4.15). Interestingly, but as expected given the hypothesis that FAB and FIN work in parallel, qRT-PCR revealed an even greater increase in the levels of *WUS*, *CLV3* and *CLE* expression in *fab fin* double mutants than in either *fab* or *fin* mutants alone (Figure 4.11C). Altogether, while *FIN* mode of action will require further investigation, our expression profiling in combination with molecular and genetic analyses, suggests that *FIN* functions separately from *FAB* and the LRR kinases in the CLV-WUS pathway to control meristem maintenance, but has a direct or indirect role in the perception of CLV3 and potentially additional CLEs by *WUS*.

Discussion

By studying two novel tomato mutants showing a combination of inflorescence branching and shoot and flower fasciation, we have found that the role of the *CLV-WUS* pathway in meristem maintenance is conserved in tomato. Surprisingly, though, we simultaneously revealed a novel pathway defined by the *FIN* gene, encoding a putative membrane protein functioning in conjunction with *CLV-WUS* to restrict stem cells from overproliferating. Delving deeper into understanding the mechanisms that control meristem size and plant architecture in different plant species and different meristem types continues to uncover additional layers of complexity to the combinations of signaling molecules, receptors complexes and downstream effectors that control specific meristem types throughout development.

The link between *FAB* and *FIN* and their underlying pathways remains unclear, but a clue may lie in the finding that their closest homologs in *Medicago* control root nodulation. Mutations in the *SUNN* and *RDN* genes (*FAB* and *FIN*, respectively) show supernodulation phenotypes without defects in meristems, inflorescence architecture, flower production or fruit size (Schnabel et al., 2005; Schnabel et al., 2011), suggesting the developmental roles of the *FAB* and *FIN* pathways may be species-specific. However, *Medicago* develops indeterminate nodules, which maintain meristem identity and therefore it is not far-fetched to speculate that *FAB* and *FIN* orthologs have similar functions to control the proliferation

of stem cells within indeterminate root nodules as they do in the tomato SAM. Furthermore, hypernodulation phenotypes in Lotus (Krusell et al., 2002; Nishimura et al., 2002), pea (Krusell et al., 2002), and soybean (Searle et al., 2003), also trace back to mutations in the closest orthologs of *CLV1* and like *SUNN* do not have flower or inflorescence defects, suggesting that cooption of *CLV1* into a nodule signaling pathway is widespread in legumes and that redundant CLV1-like receptors act in the shoot. Recent work in *Medicago* has demonstrated that *WOX5* expression is upregulated in developing nodules and becomes restricted to a limited number of cells near the apex of the nodule meristem and proper expression patterns are dependent on *SUNN/CLV1* (Osipova et al., 2012). Similarly, a functional relationship between *CLV1* and *WOX5* was recently uncovered in *Arabidopsis* root stem cell maintenance (Stahl et al., 2013), that in fact there is a widespread ancestral role for *CLV1* in both the root and shoot.

Even so, it is unclear as to why *CLV3* levels and the levels of an additional putative *CLE* are highly overexpressed in *fin,* and yet, paradoxically, the meristem is grossly enlarged in these mutants instead of being reduced in size. The *Medicago FIN* ortholog is expressed in the vasculature and is proposed to produce or transmit a long-distance signal from the roots to the shoots (Schnabel et al., 2011). In this study, we found that FIN proteins are localized to the membrane and move through the intracellular space, likely in vesicles (Figure 4.7C and Figure 4.8) and that *CLV3* and an additional *CLE* are upregulated in *fin* mutants. It is possible that the *CLV3* peptides are not properly localized to the apoplast to be perceived by CLV receptor complexes or that the receptor complexes are not formed or properly localized to perceive the signal in *fin* mutants. One intriguing possibility is that FIN is responsible for bringing RLKs, including FAB/CLV1 to the membrane. Without the proper localization of CLE receptor complexes, CLEs continue to be produced, but cannot be perceived by the cell to trigger the appropriate response. Interestingly, while *FIN* is expressed broadly through all tissue types (Figure 4.9), expression of two of the four tomato *FIN* genes is restricted to the flower and both are specifically enriched in pollen (Figure 4.16) and therefore, there is an exciting possibility that FINs are also involved in specialized signaling pathways during gametogenesis. Further analysis of the *FIN* family members in tomato and other systems is necessary to explain the intersecting roles of *CLV* and *FIN* in the regulation

of meristem size and other aspects of plant growth and development. Importantly, increases in meristem size in both *fab* and *fin* is directly related to increased flower production and larger fruits and thus, further characterization of the *FAB* and *FIN* pathways has broad implications for improving crop yield.

Methods

Plant Materials and Genotyping Markers

Six branched and fasciated mutants were isolated from a saturated mutagenesis screen (Menda et al., 2004) and formed two complementation groups with one ethyl methanesulfonate (EMS) induced allele (*fab-e0497*) for a mutant that we designated *fasciated and branched (fab)* and 5 alleles (3 EMS induced and 2 fast neutron (FN) induced: *fin-e4489*, *fin-e4643, fin-e9501, fin-n2326*, *fin-n5644*) for a mutant designated *fasciated inflorescence (fin)*. Both *fab-e0497* and *fin-e4489* (designated *fin* reference allele) were backcrossed to the M82 parental line at least three times to eliminate background mutations. Standard complementation tests demonstrated neither mutant was allelic to the only other characterized fasciated mutant, *fas* or to the inflorescence branching mutants *s* and *an.* To generate F2 mapping populations, both *fab-e0497* and *fin-e4489* were crossed to *S. pimpinellifolium* and F1 plants were subsequently self-fertilized. Double mutants (*fab fin-e4489, fab s, fab an, fin s, fin an, fin fas*) were generated using standard crossing schemes and confirmed by genotyping. Markers for recombinant mapping and genotyping were generated using SNPs and InDels identified between *S. lycopersicum* and *S. pimpinellifolium* (Consortium, 2012). *FAB* and *FIN* genotyping markers are described in Table 4.5.

RNA *in situ* hybridization

Full-length coding sequences of tomato *WUS* (Solyc02g083950), *FAB* (Solyc04g081590) and *FIN* (Solyc11g064850) were amplified from M82 cDNA using PhusionTaq (Invitrogen), sticky-A ends were added by incubating with Taq (NEB) and then cloned into StrataClone pSC-A vector (Agilent Technologies). Plasmids were linearized and T7 or T3 RNA polymerase was used for *in vitro* transcription (Roche). In addition to the full-length probe, a 259bp N-terminal fragment of *FIN* (starting with ATG) was amplified from M82 cDNA using PhusionTaq and blunt-end cloned into pENTR/D-TOPO (Invitrogen). Plasmids were linearized and T7 RNA polymerase was used for *in vitro* transcription (Roche). Tissue fixation and RNA *in situ* hybridizations were performed as described (Coen et al., 1990; Jackson, 1991), with slight modifications for tomato (fixation described below) and without hydrolysis of the probes. Shoot apical meristems were collected and fixed at 8 d.a.g. (EVM, 5 leaves initiated), 12 d.a.g (TM, 8 leaves initiated) and 14 d.a.g. (FM, 8 leaves initiated). Meristematic tissue was dissected and fixed in 4% paraformaldehyde with 0.3% Triton-X under vacuum for 20 minutes at 400 mm Hg.

Generation of Stable Transgenic Plants

The full genomic sequence of *FAB* including UTRs was amplified from wild type cv. M82 and mutant *fab* plants using High-Fidelity PhusionTaq Polymerase (NEB) sub-cloned into pENTR/D-TOPO (Invitrogen), transferred into pMDC32 (Curtis and Grossniklaus, 2003) by LR reaction, then transformed into *Agrobacterium tumificans* strain GV3101 for expression in plants. Tomato transformations were carried out at the Cornell Boyce Thompson Institute using standard techniques.

Transient transfection

FIN-YFP fusions were cloned by introducing the full FIN coding sequence without the stop codon into pH35GY and pH35YG vectors (Funakoshi). Gold particles were incubated with 1µg of plasmid DNA then introduced to onion cells by particle bombardment using a gene gun. Onion slices were incubated

under lights in petri dishes and cells were checked for expression of the protein fusions 48-72 hours after infiltration under a Nikon SMZ1500 microscope. Epidermal peals were counterstained to mark the plasma membrane in 50µm FM4-64 (Life Technologies) for 30min-1hr and in 1X Calcoflour for 5min to mark the cell wall. Plasmolysis was induced by incubation in sucrose solution for 5 minutes. Cells were imaged immediately after plasmolysis by fluorescent confocal microscopy. For expression in *Nicotiana benthamiana* leaves were transformed by agrobacterium-mediated transformation using standard procedures and imaged by confocal microscopy, but without counterstaining.

Microscopy

Hand-dissected tomato meristems were captured by manual z-series on a Nikon SMZ1500 microscope and the photographs were aligned and merged using Nikon NIS Elements to generate focused images. Meristem height and width was measured using the NIS Elements tools. For Scanning Electron Microscopy, meristems were hand-dissected and passed through a dehydrating ethanol series, dried using a critical point dryer (Tousimis), coated in gold particles and imaged on a Hitachi S-3500N microscope. Confocal images were taken using a Zeiss 710 LSM confocal microscope.

mRNA Library Construction, Illumina Sequencing and Data Analysis

RNA was extracted from 30-50 meristems per biological replicate using a PicoPure RNA Extraction kit (Arcturus), from which 1-5 µg of total RNA was used to enrich mRNA for mRNA-Seq library construction using the ScriptSeq™ v2 RNA library preparation kit (Epicentre) with barcodes. The quantity and size distribution of each individual mRNA-Seq library was detected with a High Sensitivity DNA Chip on a Bioanalyzer 2100 (Agilent). Five barcoded libraries were pooled together with relatively equal concentration for one lane of Illumina paired-end 100bp sequencing on an Illumina® HiSeq sequencing machine. All mRNA-seq reads were trimmed to 50bp and mapped with paired-end relationships maintained using Bowtie (Langmead et al., 2009) to the tomato reference CDS (Consortium,

2012). Alignments were sorted and indexed by SAMtools (Li et al., 2009) and uniquely mapped reads were counted to determine the raw counts for each library. Read counts were normalized to library size and differential expression tests were conducted using the edgeR package (Robinson et al., 2010). Differentially Expressed (DE) genes were identified (Table 4.1 and Table 4.2) using a threshold of two-fold change and a P-value <= 0.05.

Quantitative PCR

RNA was extracted from 15-20 meristems per replicate using a PicoPure RNA Extraction kit (Arcturus) followed by cDNA synthesis with Superscript III reverse transcriptase (Life Sciences). qPCR was performed on technical replicates using a SYBR Green Real-Time PCR Master Mix (Life Technology) and expression levels were normalized to levels of *UBIQUITIN*. Primers used for qPCR are listed in Table 4.5.

Phylogenetic Analysis

Multiple sequence alignments were generated using Clustal. Using the PHYLIP packages, one hundred bootstrapped data sets were generated "seqboot" and analyzed in "protpars" with 10 jumbles per data set. The consensus tree was calculated using "consense". For the FAB and FIN families, an unrooted maximum parsimony phylogenetic tree was drawn using "drawgram". For the CLEs, a neighbor-joining tree including was drawn (with node length preserved to denote the number of amino acid changes) using the "phenogram" function from the University of Indiana Phylodendron phylogenetic tree printer.

Supporting Information

MapMan Analysis

When looking at global categories alone, two major classes, "biotic stress" and "transport", were significantly altered in both *fab* and *fin* mutants. The most significantly overrepresented category for both

mutants was “transport”, and particularly enriched in both mutants is in the “p- and v-Type ATPase” category, which includes proteins involved in transport of ions, heavy metals and lipids. In *fin* mutants, “metal transport” and “ABC transport” genes are enriched. ABC transporters are known to be important for many developmental processes including organ development and reaction to abiotic stresses (Kang et al., 2011). Likewise, genes in the “abiotic stress” category are also enriched in *fin* mutants. *fin* mutant meristems expand rapidly compared to wild type and develop ectopic leaves and a highly fasciated shoot, which likely increases the mechanical stress suffered by these young mutant meristems. In *fab* mutants, which do not undergo the same dramatic overproliferation and fasciation, these stress-related categories are not enriched. However, intriguingly, aside from the ATPase transporters, the only transport categories that are slightly enriched are the “amino acid transport” and “peptide and oligopeptide transport” categories. The transport of peptides is particularly interesting as perception of the CLV3 peptide by CLV1 and other receptor protein complexes is essential for proper meristem maintenance.

Several categories involved in the “regulation of transcription” are also highly over represented in *fin* mutants. Among these, the most highly overrepresented are YABBY-like transcription factors and Homeobox transcription factors. YABBY transcription factors are involved in defining boundaries between the meristem and organ primordia (Eshed et al., 2001; Goldshmidt et al., 2008; Sarojam et al., 2010) and homeobox transcription factors are universally important in development, especially the roles of WUSCHEL-related homeobox containing transcription factors (WOXs) for tissue patterning and meristem maintenance (Breuninger et al., 2008; Haecker, 2004; Wu et al., 2007). Overexpression of both classes is consistent with the *fin* meristem showing extreme enlargement with ectopic organ formation early in development. In *fab*, we do not see overrepresentation of either of these categories; rather, the CONSTANS-like (CO-like) transcription factors, many members of which are implicated in regulating flowering time, is the only transcription factor family significantly changed in *fab* mutants.

Also notable in *fin* mutants are changes in expression of genes involved in cell wall modifications, particularly cell wall-associated proteins. Interestingly, several of these proteins are in the top downregulated genes in *fin* mutants (Table S2, Table S4), including a number of pectinesterases,

which are known to help facilitate modifications at the cell wall and cellulose synthesis genes are also altered. Curiously, the ortholog of *Arabidopsis TAPETUM DETERMINANT1*, which was identified in a male sterile screen and is known to interact with the LRR kinase EMS1 (Jia et al., 2008), is not normally expressed in the tomato vegetative meristem, but was detected in *fin* mutants. Similarly, *GAMETE EXPRESSED PROTEIN1* a transmembrane domain protein that has a role in gametophyte development in *Arabidopsis* (Alandete-Saez et al., 2011) is also ectopically expressed. Finally, the tomato homolog of *DVL18*, a gene that encodes a small polypeptide, which when overexpressed in *Arabidopsis* resulted in short plants with clustered inflorescences (Wen et al., 2004), is ectopically expressed in the *fin* vegetative meristem.

References

Aggarwal, P., Yadav, R.K., Reddy, G.V., 2010. Identification of novel markers for stem-cell niche of Arabidopsis shoot apex. Gene Expr Patterns 10, 259-264.

Aichinger, E., Kornet, N., Friedrich, T., Laux, T., 2012. Plant stem cell niches. Annu Rev Plant Biol 63, 615-636.

Alandete-Saez, M., Ron, M., Leiboff, S., McCormick, S., 2011. Arabidopsis thaliana GEX1 has dual functions in gametophyte development and early embryogenesis. Plant J 68, 620-632.

Barton, M.K., 2010. Twenty years on: the inner workings of the shoot apical meristem, a developmental dynamo. Developmental biology 341, 95-113.

Bleckmann, A., Weidtkamp-Peters, S., Seidel, C.A., Simon, R., 2010. Stem cell signaling in Arabidopsis requires CRN to localize CLV2 to the plasma membrane. Plant Physiol 152, 166-176.

Bommert, P., Je, B.I., Goldshmidt, A., Jackson, D., 2013a. The maize Galpha gene COMPACT PLANT2 functions in CLAVATA signalling to control shoot meristem size. Nature.

Bommert, P., Lunde, C., Nardmann, J., Vollbrecht, E., Running, M., Jackson, D., Hake, S., Werr, W., 2005. thick tassel dwarf1 encodes a putative maize ortholog of the Arabidopsis CLAVATA1 leucine-rich repeat receptor-like kinase. Development 132, 1235-1245.

Bommert, P., Nagasawa, N.S., Jackson, D., 2013b. Quantitative variation in maize kernel row number is controlled by the FASCIATED EAR2 locus. Nature genetics 45, 334-337.

Brand, U., Fletcher, J.C., Hobe, M., Meyerowitz, E.M., Simon, R., 2000. Dependence of Stem Cell Fate in Arabidopsis on a Feedback Loop Regulated by CLV3 Activity. Science 289, 617-619.

Breuninger, H., Rikirsch, E., Hermann, M., Ueda, M., Laux, T., 2008. Differential expression of WOX genes mediates apical-basal axis formation in the Arabidopsis embryo. Dev Cell 14, 867-876.

Clark, S.E., Williams, R.W., Meyerowitz, E.M., 1997. The CLAVATA1 gene encodes a putative receptor kinase that controls shoot and floral meristem size in Arabidopsis. Cell 89, 575-585.

Coen, E.S., Romero, J.M., Doyle, S., Elliott, R., Murphy, G., Carpenter, R., 1990. floricaula: a homeotic gene required for flower development in antirrhinum majus. Cell 63, 1311-1322.

Cong, B., Barrero, L.S., Tanksley, S.D., 2008. Regulatory change in YABBY-like transcription factor led to evolution of extreme fruit size during tomato domestication. Nature genetics 40, 800-804.

Consortium, T.T.G., 2012. The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485, 635-641.

Curtis, M.D., Grossniklaus, U., 2003. A gateway cloning vector set for high-throughput functional analysis of genes in planta. Plant Physiol 133, 462-469.

Dievart, A., Dalal, M., Tax, F.E., Lacey, A.D., Huttly, A., Li, J., Clark, S.E., 2003. CLAVATA1 Dominant-Negative Alleles Reveal Functional Overlap between Multiple Receptor Kinases That Regulate Meristem and Organ Development. The Plant Cell 15, 1198-1211.

Eshed, Y., Baum, S.F., Perea, J.V., Bowman, J.L., 2001. Establishment of polarity in lateral organs of plants. Curr Biol 11, 1251-1260.

Frary, A., Nesbitt, T.C., Grandillo, S., Knaap, E., Cong, B., Liu, J., Meller, J., Elber, R., Alpert, K.B., Tanksley, S.D., 2000. fw2.2: a quantitative trait locus key to the evolution of tomato fruit size. Science 289, 85-88.

Goldshmidt, A., Alvarez, J.P., Bowman, J.L., Eshed, Y., 2008. Signals derived from YABBY gene activities in organ primordia regulate growth and partitioning of Arabidopsis shoot apical meristems. The Plant cell 20, 1217-1230.

Ha, C.M., Jun, J.H., Fletcher, J.C., 2010. Shoot Apical Meristem Form and Function. 91, 103-140.

Haecker, A., 2004. Expression dynamics of WOX genes mark cell fate decisions during early embryonic patterning in Arabidopsis thaliana. Development 131, 657-668.

Ito, Y., Nakanomyo, I., Motose, H., Iwamoto, K., Sawa, S., Dohmae, N., Fukuda, H., 2006. Dodeca-CLE peptides as suppressors of plant stem cell differentiation. Science 313, 842-845.

Jackson, D.P., 1991. In situ hybridization in plants, in: DJ Bowles, S.G., and M McPherson (Ed.), Molecular Plant Pathology: A Practical Approach. Oxford University Press, England.

Jaquinod, M., Villiers, F., Kieffer-Jaquinod, S., Hugouvieux, V., Bruley, C., Garin, J., Bourguignon, J., 2007. A proteomics dissection of Arabidopsis thaliana vacuoles isolated from cell culture. Mol Cell Proteomics 6, 394-412.

Jeong, S., Trotochaud, A.E., Clark, S.E., 1999. The Arabidopsis CLAVATA2 gene encodes a receptor-like protein required for the stability of the CLAVATA1 receptor-like kinase. The Plant cell 11, 1925-1934.

Jia, G., Liu, X., Owen, H.A., Zhao, D., 2008. Signaling of cell fate determination by the TPD1 small protein and EMS1 receptor kinase. Proceedings of the National Academy of Sciences of the United States of America 105, 2220-2225.

Jun, J.H., Fiume, E., Fletcher, J.C., 2008. The CLE family of plant polypeptide signaling molecules. Cell Mol Life Sci 65, 743-755.

Kang, J., Park, J., Choi, H., Burla, B., Kretzschmar, T., Lee, Y., Martinoia, E., 2011. Plant ABC Transporters. Arabidopsis Book 9, e0153.

Kinoshita, A., Betsuyaku, S., Osakabe, Y., Mizuno, S., Nagawa, S., Stahl, Y., Simon, R., Yamaguchi-Shinozaki, K., Fukuda, H., Sawa, S., 2010. RPK2 is an essential receptor-like kinase that transmits the CLV3 signal in Arabidopsis. Development 137, 3911-3920.

Krusell, L., Madsen, L.H., Sato, S., Aubert, G., Genua, A., Szczyglowski, K., Duc, G., Kaneko, T., Tabata, S., de Bruijn, F., Pajuelo, E., Sandal, N., Stougaard, J., 2002. Shoot control of root development and nodulation is mediated by a receptor-like kinase. Nature 420, 422-426.

Krusell, L., Sato, N., Fukuhara, I., Koch, B.E., Grossmann, C., Okamoto, S., Oka-Kira, E., Otsubo, Y., Aubert, G., Nakagawa, T., Sato, S., Tabata, S., Duc, G., Parniske, M., Wang, T.L., Kawaguchi, M., Stougaard, J., 2011. The Clavata2 genes of pea and Lotus japonicus affect autoregulation of nodulation. Plant J 65, 861-871.

Langmead, B., Trapnell, C., Pop, M., Salzberg, S.L., 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10, R25.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078-2079.

Lippman, Z., Tanksley, S.D., 2001. Dissecting the genetic pathway to extreme fruit size in tomato using a cross between the small-fruited wild species Lycopersicon pimpinellifolium and L. esculentum var. Giant Heirloom. Genetics 158, 413-422.

Lippman, Z.B., Cohen, O., Alvarez, J.P., Abu-Abied, M., Pekker, I., Paran, I., Eshed, Y., Zamir, D., 2008. The making of a compound inflorescence in tomato and related nightshades. PLoS biology 6, e288.

Long, J.A., Moan, E.I., Medford, J.I., Barton, M.K., 1996. A member of the KNOTTED class of homeodomain proteins encoded by the STM gene of Arabidopsis. Nature 379, 66-69.

Marmagne, A., Ferro, M., Meinnel, T., Bruley, C., Kuhn, L., Garin, J., Barbier-Brygoo, H., Ephritikhine, G., 2007. A high content in lipid-modified peripheral proteins and integral receptor kinases features in the arabidopsis plasma membrane proteome. Mol Cell Proteomics 6, 1980-1996.

Mattioli, P.A., 1544. Discorsi.

Mayer, K.F., Schoof, H., Haecker, A., Lenhard, M., Jurgens, G., Laux, T., 1998. Role of WUSCHEL in regulating stem cell fate in the Arabidopsis shoot meristem. Cell 95, 805-815.

Menda, N., Semel, Y., Peled, D., Eshed, Y., Zamir, D., 2004. In silico screening of a saturated mutation library of tomato. Plant J 38, 861-872.

Mitra, S.K., Walters, B.T., Clouse, S.D., Goshe, M.B., 2009. An efficient organic solvent based extraction method for the proteomic analysis of Arabidopsis plasma membranes. J Proteome Res 8, 2752-2767.

Miwa, H., Betsuyaku, S., Iwamoto, K., Kinoshita, A., Fukuda, H., Sawa, S., 2008. The receptor-like kinase SOL2 mediates CLE signaling in Arabidopsis. Plant Cell Physiol 49, 1752-1757.

Muller, R., Bleckmann, A., Simon, R., 2008. The receptor kinase CORYNE of Arabidopsis transmits the stem cell-limiting signal CLAVATA3 independently of CLAVATA1. The Plant cell 20, 934-946.

Muller, R., Borghi, L., Kwiatkowska, D., Laufs, P., Simon, R., 2006. Dynamic and compensatory responses of Arabidopsis shoot and floral meristems to CLV3 signaling. The Plant cell 18, 1188-1198.

Munos, S., Ranc, N., Botton, E., Berard, A., Rolland, S., Duffe, P., Carretero, Y., Le Paslier, M.C., Delalande, C., Bouzayen, M., Brunel, D., Causse, M., 2011. Increase in tomato locule number is controlled by two single-nucleotide polymorphisms located near WUSCHEL. Plant Physiol 156, 2244-2254.

Nishimura, R., Hayashi, M., Wu, G.J., Kouchi, H., Imaizumi-Anraku, H., Murakami, Y., Kawasaki, S., Akao, S., Ohmori, M., Nagasawa, M., Harada, K., Kawaguchi, M., 2002. HAR1 mediates systemic regulation of symbiotic organ development. Nature 420, 426-429.

Ogawa, M., Shinohara, H., Sakagami, Y., Matsubayashi, Y., 2008. Arabidopsis CLV3 peptide directly binds CLV1 ectodomain. Science 319, 294.

Osipova, M.A., Mortier, V., Demchenko, K.N., Tsyganov, V.E., Tikhonovich, I.A., Lutova, L.A., Dolgikh, E.A., Goormachtig, S., 2012. Wuschel-related homeobox5 gene expression and interaction of CLE peptides with components of the systemic control add two pieces to the puzzle of autoregulation of nodulation. Plant Physiol 158, 1329-1341.

Park, S.J., Jiang, K., Schatz, M.C., Lippman, Z.B., 2012. Rate of meristem maturation determines inflorescence architecture in tomato. Proceedings of the National Academy of Sciences of the United States of America 109, 639-644.

Pautler, M., Tanaka, W., Hirano, H.Y., Jackson, D., 2013. Grass meristems I: shoot apical meristem maintenance, axillary meristem determinacy and the floral transition. Plant Cell Physiol 54, 302-312.

Peralta, I.E., Spooner, D.M., 2005. Morphological characterization and relationships of wild tomatoes (Solanum L. Sect. Lycopersicon). Monogr Syst Bot, Missouri Bot Gard 104, 227-257.

Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139-140.

Rodriguez, G.R., Munos, S., Anderson, C., Sim, S.C., Michel, A., Causse, M., Gardener, B.B., Francis, D., van der Knaap, E., 2011. Distribution of SUN, OVATE, LC, and FAS in the tomato germplasm and the relationship to fruit shape diversity. Plant Physiol 156, 275-285.

Sarojam, R., Sappl, P.G., Goldshmidt, A., Efroni, I., Floyd, S.K., Eshed, Y., Bowman, J.L., 2010. Differentiating Arabidopsis shoots from leaves by combined YABBY activities. The Plant Cell 22, 2113-2130.

Schnabel, E., Journet, E.P., de Carvalho-Niebel, F., Duc, G., Frugoli, J., 2005. The Medicago truncatula SUNN gene encodes a CLV1-like leucine-rich repeat receptor kinase that regulates nodule number and root length. Plant Mol Biol 58, 809-822.

Schnabel, E.L., Kassaw, T.K., Smith, L.S., Marsh, J.F., Oldroyd, G.E., Long, S.R., Frugoli, J.A., 2011. The ROOT DETERMINED NODULATION1 gene regulates nodule number in roots of Medicago truncatula and defines a highly conserved, uncharacterized plant gene family. Plant Physiol 157, 328-340.

Schoof, H., Lenhard, M., Haecker, A., Mayer, K.F., Jurgens, G., Laux, T., 2000. The stem cell population of Arabidopsis shoot meristems in maintained by a regulatory loop between the CLAVATA and WUSCHEL genes. Cell 100, 635-644.

Searle, I.R., Men, A.E., Laniya, T.S., Buzas, D.M., Iturbe-Ormaetxe, I., Carroll, B.J., Gresshoff, P.M., 2003. Long-distance signaling in nodulation directed by a CLAVATA1-like receptor kinase. Science 299, 109-112.

Song, S.K., Lee, M.M., Clark, S.E., 2006. POL and PLL1 phosphatases are CLAVATA1 signaling intermediates required for Arabidopsis shoot and floral stem cells. Development 133, 4691-4698.

Stahl, Y., Grabowski, S., Bleckmann, A., Kuhnemuth, R., Weidtkamp-Peters, S., Pinto, K.G., Kirschner, G.K., Schmid, J.B., Wink, R.H., Hulsewede, A., Felekyan, S., Seidel, C.A., Simon, R., 2013. Moderation of Arabidopsis root stemness by CLAVATA1 and ARABIDOPSIS CRINKLY4 receptor kinase complexes. Curr Biol 23, 362-371.

Stahl, Y., Simon, R., 2010. Plant primary meristems: shared functions and regulatory mechanisms. Curr Opin Plant Biol 13, 53-58.

Stahl, Y., Wink, R.H., Ingram, G.C., Simon, R., 2009. A signaling module controlling the stem cell niche in Arabidopsis root meristems. Curr Biol 19, 909-914.

Stubbe, H., 1961. Mutanten der Wildtomate Lycopersicon pimpinellifolium II.

Suzaki, T., Ohneda, M., Toriba, T., Yoshida, A., Hirano, H.Y., 2009. FON2 SPARE1 redundantly regulates floral meristem maintenance with FLORAL ORGAN NUMBER2 in rice. PLoS genetics 5, e1000693.

Suzaki, T., Sato, M., Ashikari, M., Miyoshi, M., Nagato, Y., Hirano, H.Y., 2004. The gene FLORAL ORGAN NUMBER1 regulates floral meristem size in rice and encodes a leucine-rich repeat receptor kinase orthologous to Arabidopsis CLAVATA1. Development 131, 5649-5657.

Taguchi-Shiobara, F., Yuan, Z., Hake, S., Jackson, D., 2001. The fasciated ear2 gene encodes a leucine-rich repeat receptor-like protein that regulates shoot meristem proliferation in maize. Genes Dev 15, 2755-2766.

Tanksley, S.D., 2004. The genetic, developmental, and molecular bases of fruit size and shape variation in tomato. The Plant cell 16 Suppl, S181-189.

Thimm, O., Blasing, O., Gibon, Y., Nagel, A., Meyer, S., Kruger, P., Selbig, J., Muller, L.A., Rhee, S.Y., Stitt, M., 2004. MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. Plant J 37, 914-939.

Tusnady, G.E., Simon, I., 1998. Principles governing amino acid composition of integral membrane proteins: application to topology prediction. J Mol Biol 283, 489-506.

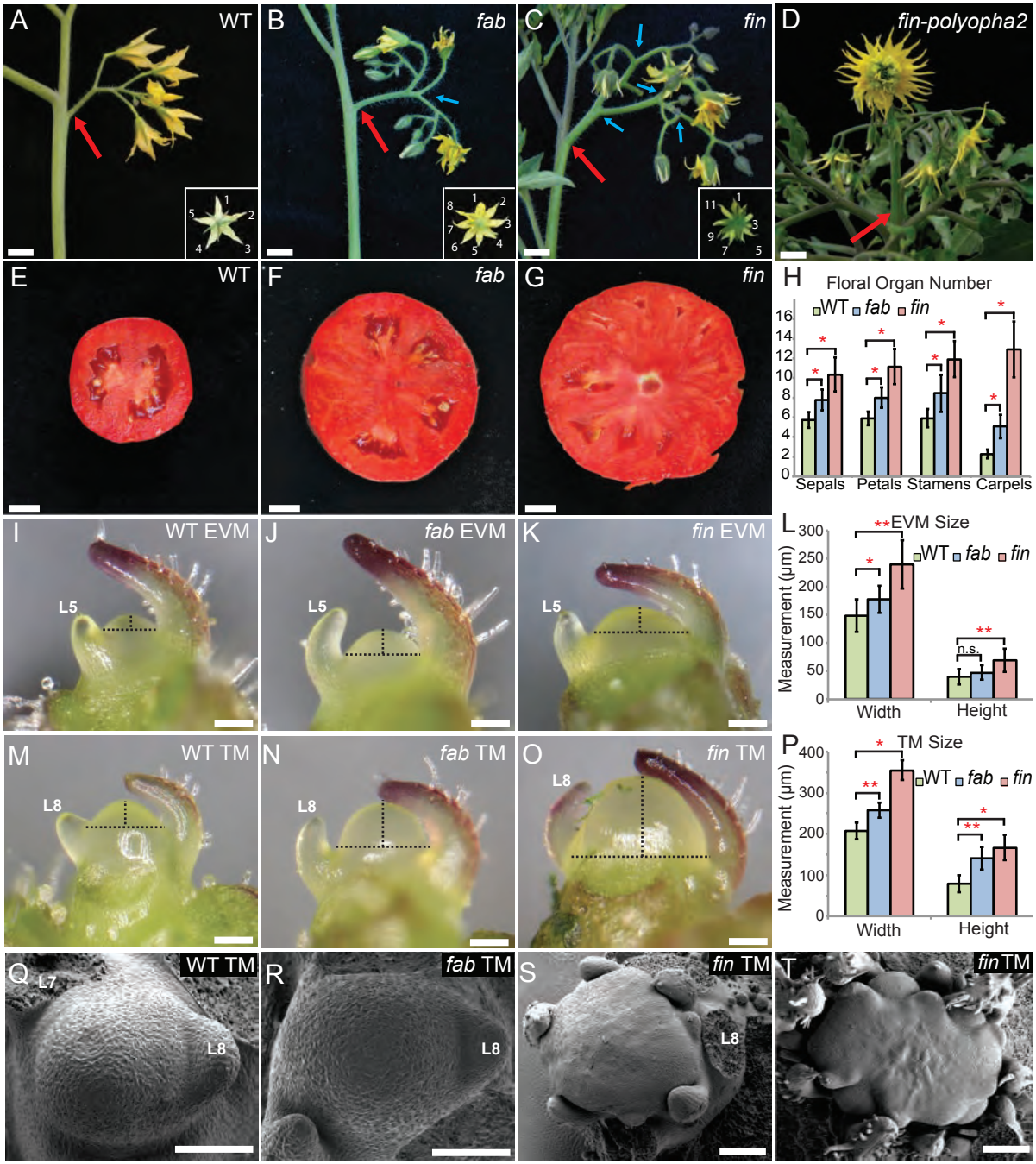
Tusnady, G.E., Simon, I., 2001. The HMMTOP transmembrane topology prediction server. Bioinformatics 17, 849-850.

Usadel, B., Nagel, A., Steinhauser, D., Gibon, Y., Blasing, O.E., Redestig, H., Sreenivasulu, N., Krall, L., Hannah, M.A., Poree, F., Fernie, A.R., Stitt, M., 2006. PageMan: an interactive ontology tool to generate, display, and annotate overview graphs for profiling experiments. BMC Bioinformatics 7, 535.

Wen, J., Lease, K.A., Walker, J.C., 2004. DVL, a novel class of small polypeptides: overexpression alters Arabidopsis development. Plant J 37, 668-677.

Wu, X., Chory, J., Weigel, D., 2007. Combinations of WOX activities regulate tissue proliferation during Arabidopsis embryonic development. Developmental biology 309, 306-316.

Figure 4.1. Similar mutant phenotypes are manifested early in development in both mutants, but are more extreme in *fin* than in *fab*. (A-H) Images and quantification of mature inflorescence (IF), flower and fruit phenotypes. Representative secondary inflorescences of wild type (A), *fab* (B) and *fin* (C) and the primary IF of *fin* (D). The *fin-polyopha2* mutant from *S. pimpinellifolium* is shown in panel D; all other *fin* images and are from the *S. lycopersicum fin-e4489* reference allele. Floral image insets within each panel show the increased number of sepals and petals in both *fab* and *fin* mutants. The base of each IF is marked with a red arrow and points of IF branching are marked with blue arrows; scale bars = 1cm. (E-G) Mature fruits illustrate the increase in locule number and fruit size in *fab* and *fin* mutants that converts the wild type M82 roma-sized tomatoes to beefsteak-sized; scale bars = 1cm. (H) Quantification of floral organ number (mean ±SD). Asterisks indicate statistical difference between genotypes; students t-test, $p<0.0001$. (I-K, M-O) A side view of meristems reveals an early emergence of the fasciated phenotype in both mutants by the early vegetative meristem (EVM/8dag) stage (I-K), which is more noticeable in the transition meristem TM/15dag) stage (M-O). Dashed lines mark the height and width dimensions used for meristem size quantification (L and P). Bars represent the mean values ±SD of meristem height and width at the EVM stage (L) and TM stage (P) for both mutants and wild type. Asterisks indicate statistical difference between genotypes; Wilcoxon rank sums test, $*p<0.01$, $**p<0.001$. Both mutants have significantly larger meristems as early as the EVM stage (L). Genotype and stage indicated in the top right of each panel; scale bars = 100μm. (Q-T) Scanning Electron Micrograph top view of the wild type (Q), *fab* (R) and *fin* (S and T) TM stages illustrates the great overproliferation and loss of regular phylotaxy in *fin* mutants by the onset of the TM stage (S). Many additional leaves are formed in *fin* and the meristem disorganization and enlargement becomes progressively worse prior to forming the first flower (T); scale bars=100μm.


A WT
B fab
C fin
D fin-polyopha2
E WT
F fab
G fin
H Floral Organ Number
WT fab fin
Sepals Petals Stamens Carpels
I WT EVM
J fab EVM
K fin EVM
L5
L EVM Size
Measurement (µm)
Width Height
n.s
M WT TM
N fab TM
O fin TM
L8
P TM Size
Q WT TM
R fab TM
S fin TM
T fin TM
L7

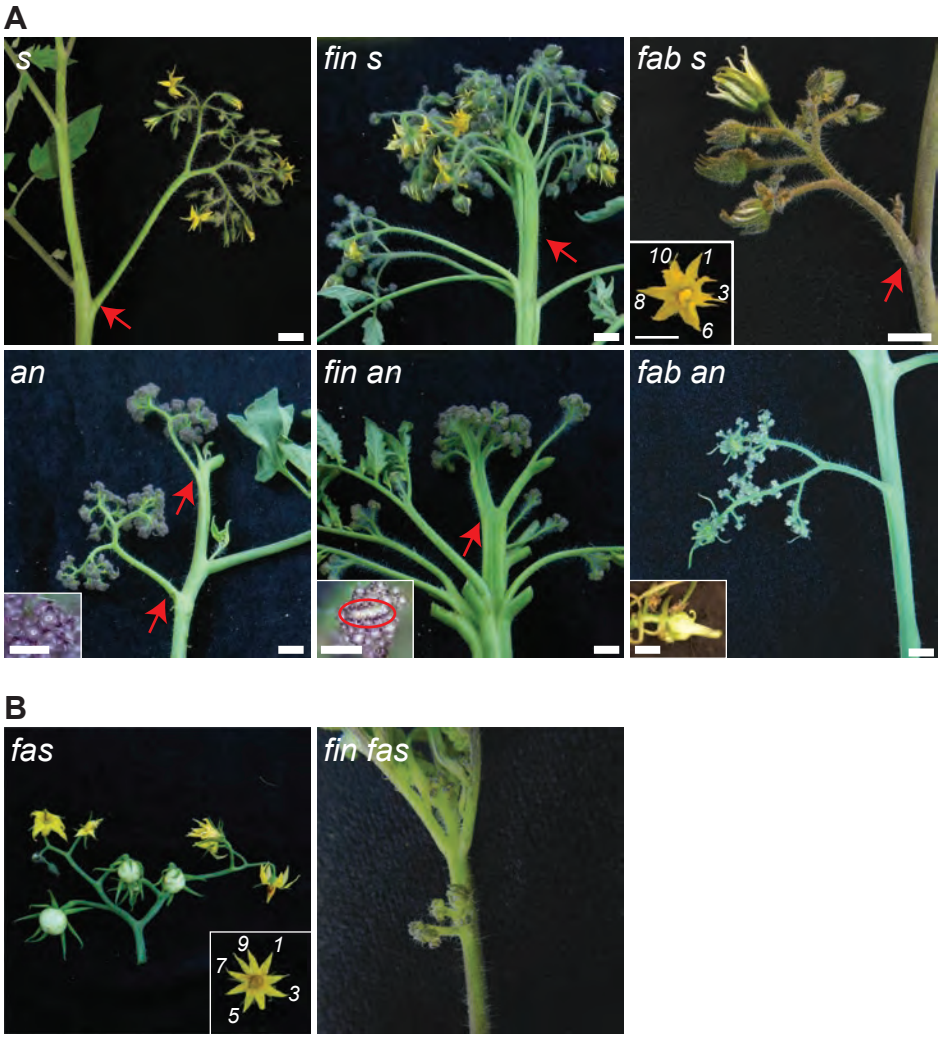


Figure 4.2. *FAB* and *FIN* act in separate pathways from *S* and *AN* and *FAS* to suppress inflorescence branching. (A) Images of inflorescences from double mutant combinations between both *fab* and *fin* when combined with *s* or *an* mutations. Red arrows point to the bottom of each inflorescence. The *s* and *an* single mutants are shown (left) for reference. The *fin s* (top middle) and *fab s* (top right) double mutants have both branching resembling or exceeding the *s* single mutants and fasciation of the flowers characteristic of the *fab* and *fin* single mutant phenotypes, respectfully (also see Figure 1A). *fin an* double mutants (bottom middle) show fasciation of the shoot characteristic of *fin* mutants and an inflorescence that is more highly branched than either *an* or *fin* alone and extremely fasciated floral meristems (inset). *fab an* double mutants (bottom right) are branched to a similar degree as either single mutant and form mostly small *an*-like meristem clusters along the inflorescence, but sometimes form terminal *fab*-like fasciated floral meristems that develop an enlarged pistil-like structure (inset), but do not form complete, viable flowers. (B) Double mutant analysis between *fin* and *fas*. *fas* single mutant inflorescence shows the characteristic inflorescence branching and flower and fruit fasciation (left). A flower is inset with petals labeled to demonstrate the increase in organ number compared to wild type plants (Figure 1). *fin fas* double mutants develop small inflorescences with inviable flowers (right). Scale bars = 1 cm.

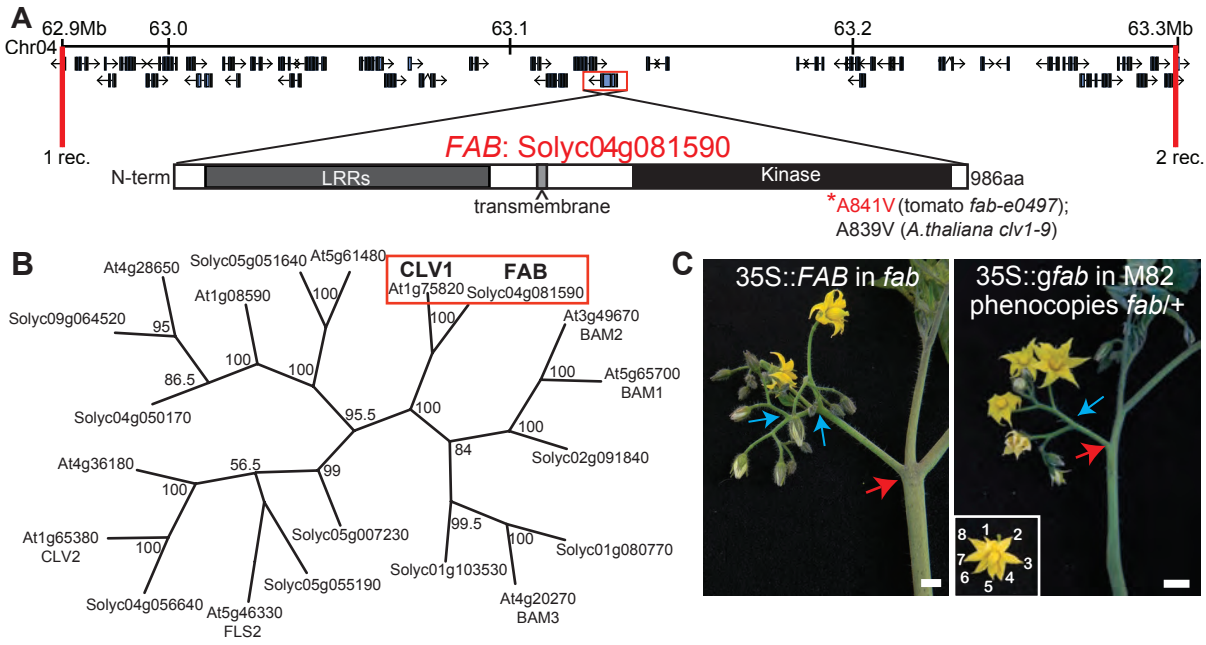


Figure 4.3. *FAB* encodes the tomato ortholog of *CLAVATA1*. (A) The mapping interval of *FAB*. Vertical red lines show the positions of the closest mapping markers used with the number of recombinants (rec) listed below. *FAB* is boxed in red and a gene model is shown below indicating characteristic domains and motifs. A red asterisk marks the site of the point mutation found in *fab-e0497*; the amino acid substitution at this position for both tomato and the identical substitution for *A. thaliana clv1-9* (Dievart et al. 2003) are also indicated. (B) Phylogenetic tree including tomato and *Arabidopsis* CLV1 and CLV2 and related proteins demonstrates that tomato FAB is the closest ortholog to *Arabidopsis* CLV1 (boxed in red). Maximum parsimony consensus tree with 100 replicate bootstrap values indicated at the nodes. (C) Images of inflorescences from representative transgenic plants demonstrates that as expected, 35S::g*FAB* fails to complement *fab* mutants (left), however in >20% of T1 transformants, expression of 35S::g*fab* in M82 phenocopies the *fab* heterozygous dominant negative phenotype (right; see Figure S2 for quantitation of *fab*/+ phenotype). Arrows mark the base of the IF (red) and site of branching (blue). Inset shows a flower with the number of petals labeled. Scale bar = 1cm.

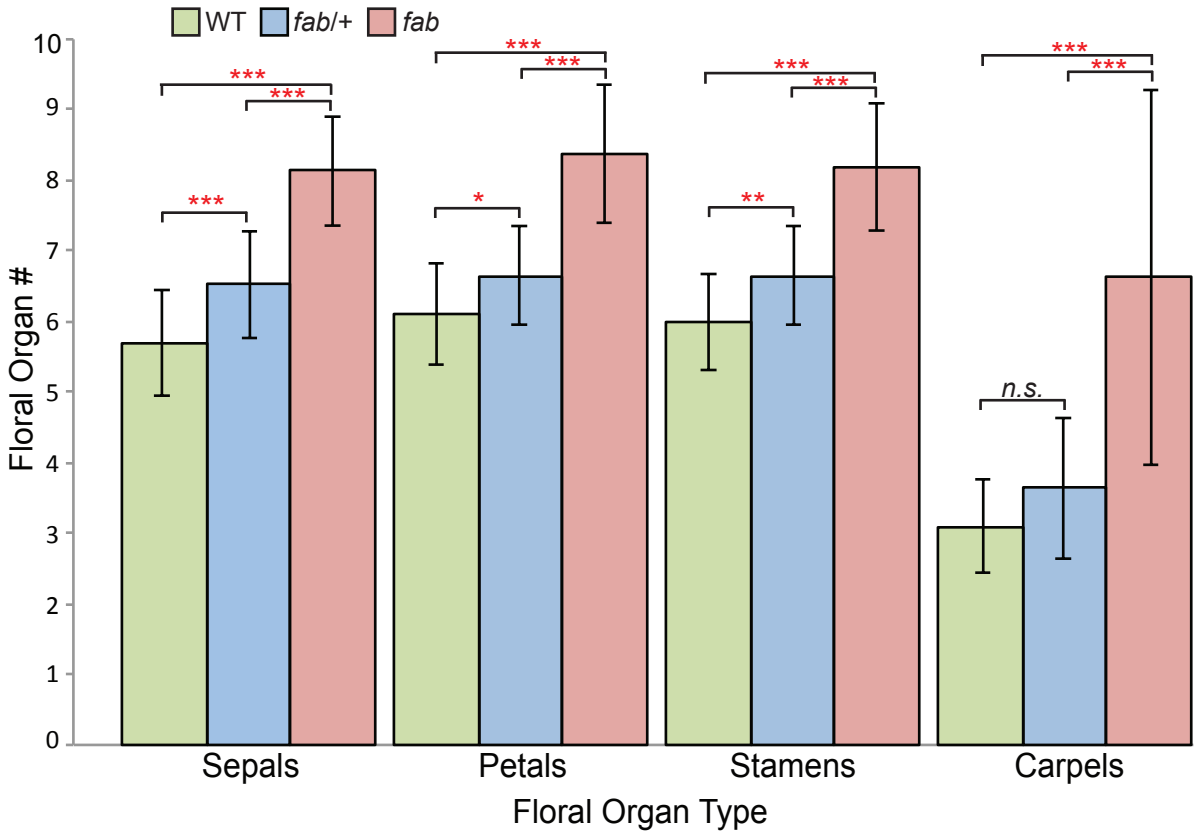


Figure 4.4. Quantitative phenotyping reveals weak semi-dominance for floral organ numbers in *fab/+* heterozygotes. Sepal, petal, stamen and carpel numbers for wild type, *fab/+* heterozygotes and *fab* homozygous mutants are represented in the bar graph (mean ±SD, n>29 flowers for each genotype). Significant differences between means were tested by a Student's t-test and significant values are indicated with asterisks (*p<0.005, **p<0.001, ***p<0.0001). n.s.=not significant.

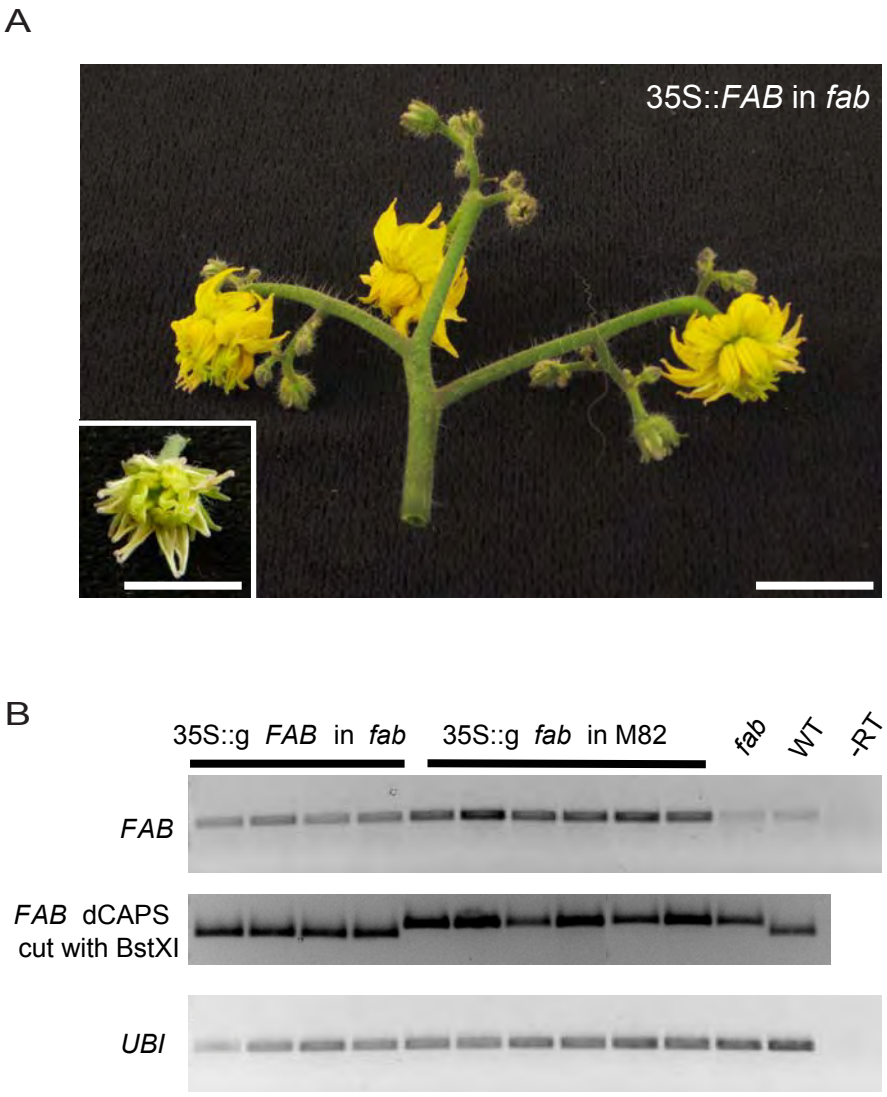


Figure 4.5. Exacerbated *fab* mutant phenotypes when overexpressing *FAB* in a *fab* mutant background is not a result of co-suppression. (A) Extreme fasciation phenotype observed in a few cases when overexpressing the wild type version of *FAB* in a *fab* mutant background. The fasciation of the flowers and inflorescence branching is more extreme and the flowers and overall inflorescence structure are smaller than in *fab* homozygous mutants (see Figure 1). (B) Semi-quantitative RT-PCR verifies that the transgenes, which are driven under the CaMV 35S promoter, are expressed and at higher levels than in non-transgenic plants (Top row). Tomato *UBIQUITIN* was used as a loading control (Bottom Row). Non-transgenic *fab*, WT, and a minus reverse transcription (-RT) negative control are shown at the far right for comparison. (C) A dCAPs marker verifies that the correct transgenic construct is overexpressed in each background; the enzyme *BstXI* cuts the wild type, but not mutant version of the PCR-amplified fragment. Because of transgene overexpression compared to the endogenous copy, only the transgenic construct is detected (middle row).

Figure 4.6. Extreme synergistic phenotype of *fab fin* double mutants. Images of a mature *fab fin* double mutant plant (left) with close-up images of the primary inflorescence (middle and right). Red arrows point to the base of the primary IF which fails to produce flowers. White arrows mark released axillary shoots. Stereoscopic imaging (right) illustrates extreme enlargement, invagination and proliferation of ectopic meristems at the meristem flanks.

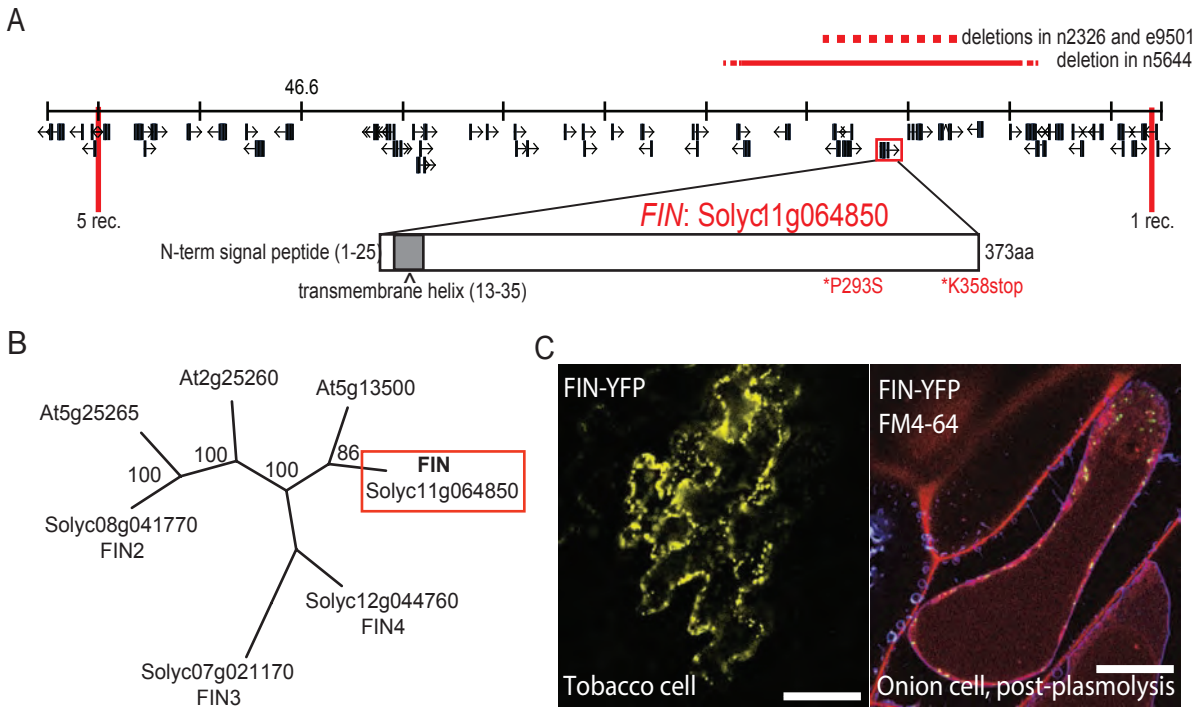


Figure 4.7. *FIN* encodes a small, membrane-localized protein of unknown function. (A) The mapping interval of *FIN*. Red lines show the positions of the mapping markers with the number of recombinants (rec) listed below. *FIN* is boxed in red and a gene model is displayed below with characteristic motifs indicated. Asterisks mark the two identified point mutations. Horizontal red lines above the map indicate the deletion mutants, with dashed lines marking the approximate boarders of the deletions. (B) Phylogenetic tree of the full tomato and *Arabidopsis* FIN protein family. The maximum parsimony consensus tree is shown with 100 replicate bootstrap values indicated at each node. Tomato FIN is boxed in red. (C) 35S::FIN-YFP protein localization to the membrane visualized by confocal microscopy in transiently transfected tobacco cells (left) and onion epidermal cells (right). Onion cells were counterstained with FM4-64 (red) and plasmolysis was induced to detach the membrane from the cell wall, indicating that puncta of FIN-YFP localize to the plasma membrane. Scale bars = 50μm.

Figure 4.8. A time-lapse movie showing FIN movement at the plasma membrane and throughout the intracellular space. (See accompanying video). Puncta of FIN-YFP expression are detected along the plasma membrane (Figure 4.7). These puncta are consistently observed to move along the membrane and through the intracellular space. Fusion of individual puncta and fusion of puncta with the plasma membrane are commonly observed, suggestive of FIN trafficking in vesicles.

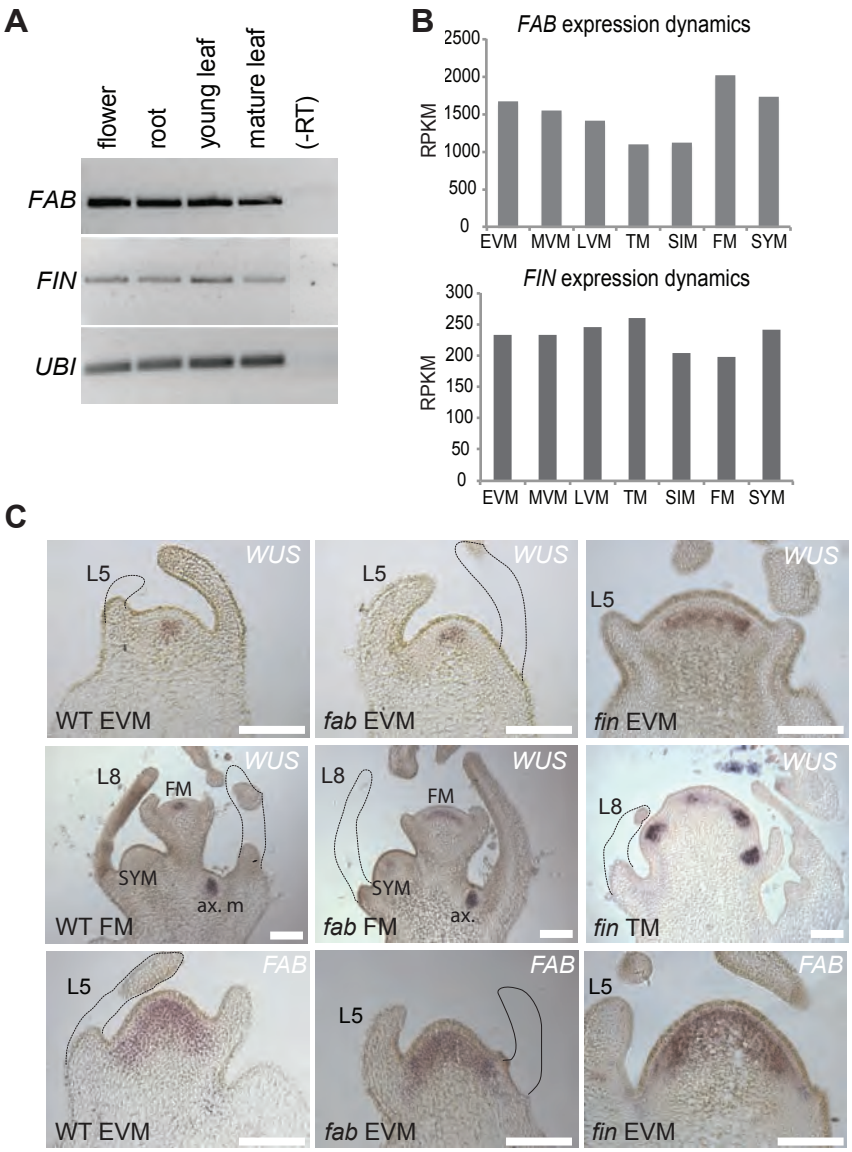


Figure 4.9. Spatial and temporal expression analysis of *fab* and *fin* mutants. (A) Semi-quantitative RT-PCR for *FAB* and *FIN* transcripts on a panel of tissues indicates broad expression of both genes. Tomato *UBIQUITIN* (*UBI*) was used as a loading control. (B) *FAB* and *FIN* are expressed in all shoot apical meristem types during the transition from vegetative to reproductive growth according to previous transcriptome profiling (Park et al. 2012). Meristem stages are marked below each bar (early, mid and late vegetative meristems; EVM/8dag, MVM/10dag and LVM/13dag, transition meristem (TM/15dag), FM and SIM/18dag). (C) Spatial regulation of *WUS* (top two rows) and *FAB* (bottom row) by *in situ* hybridization using antisense probes. *WUS* expression domain is grossly expanded in *fin* mutants compared to wild type by the EVM stage. While *WUS* expression is not significantly changed in the *fab* EVM, it expands laterally in the *fab* FM. *WUS* expression is also detected in the sympodial shoot meristem (SYM) and in axillary meristems (ax.) that will renew apical tomato growth. At 15dag, when WT and *fab* have transitioned to flowering, *fin* meristems are still in the transition state and form several ectopic branches, marked here by puncta of *WUS* expression, prior to flowering. Genotype and stage is indicated in the bottom left of each panel and the probes are indicated in the top right. Scale bar = 100μm.

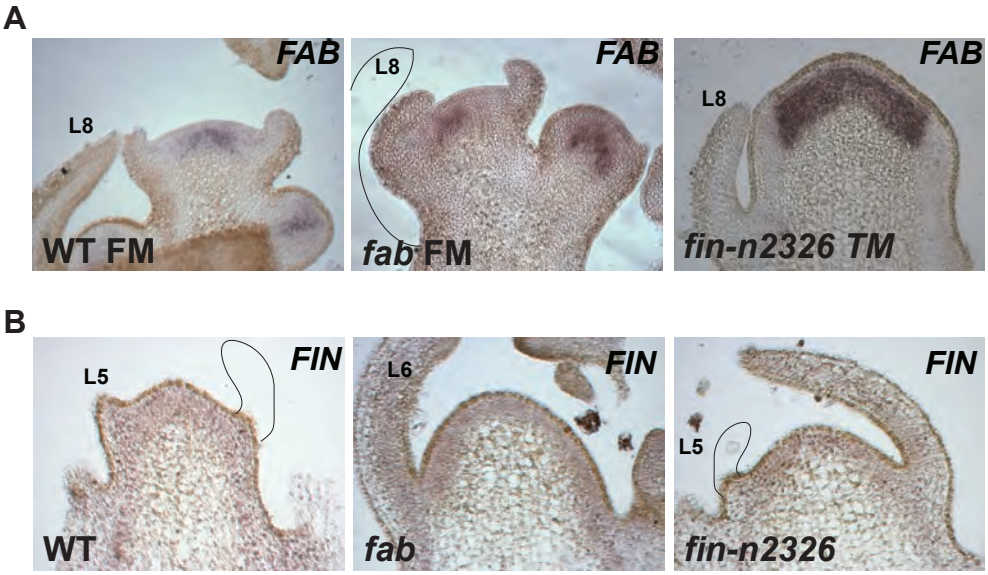


Figure 4.10. *FAB* transcripts are spatially localized below the L2 in all stages, whereas *FIN* transcripts are either too low to detect or diffuse throughout the meristem. (A) *FAB* expression is restricted below the L2 layer. Wild type and *fab* FM and SIMs as well as *fin* TM show the same spatial expression as in vegetative meristems (Figure 5C). (B) *FIN* expression levels are too low to detect or too diffuse throughout the meristem to see a distinct expression pattern. Images are representative of early to mid-vegetative meristems (EVM/8dag) probed with the full-length *FIN* anti-sense probe. The *fin*-n2326 deletion allele was probed as a negative control. A shorter N-terminal FIN probe and a mixture of the full-length and N-terminal probe were also tested and yielded the same result (not shown). Scale bars = 100μm.

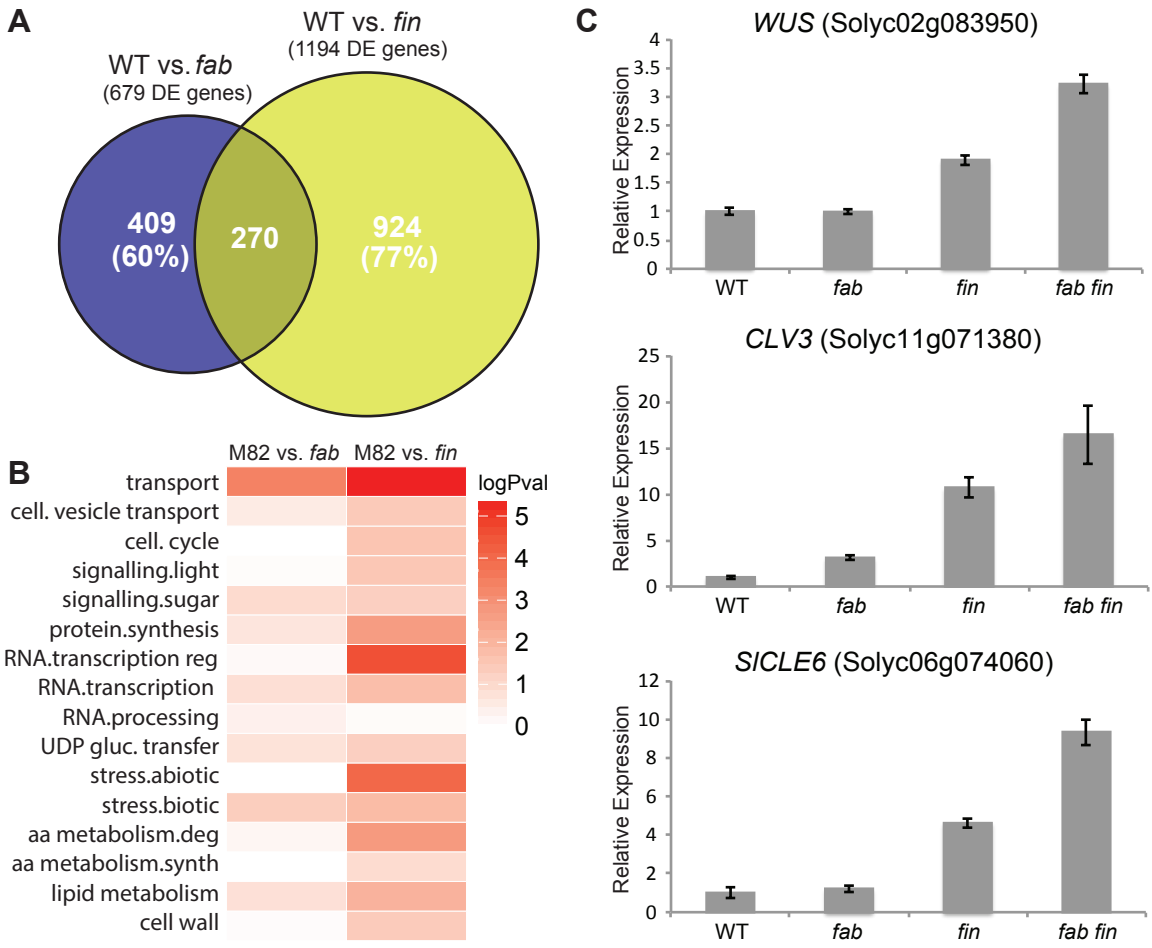


Figure 4.11. Transcriptome profiling reveals that *CLV3* and an additional *CLE* are overexpressed in *fin* mutants. (A) A venn diagram indicates the number of differentially expressed (DE) genes in *fab* versus M82, *fin* versus M82 and the overlap between the two sets (>30% overlap). (B) Different functional groups are enriched in *fab* versus *fin* mutants as determined by MapMan classifications. Major categories that were differentially expressed in either of the mutants and determined to be overrepresented as derived from Fisher's exact tests of the proportions of the selected categories in the mutants versus the proportions in the genome annotation. Scaled $-\log_{10}$P-values are represented in the heat-map; red indicates a significant overrepresentation of genes that are differentially expressed in a particular gene category for the given mutant. (C) Quantitative RT-PCR of *WUS*, *CLV3* and an additional putative *CLE* gene. Two technical replicates of EVM meristems from each genotype were run and expression levels were normalized to levels of *UBIQUITIN* transcript.

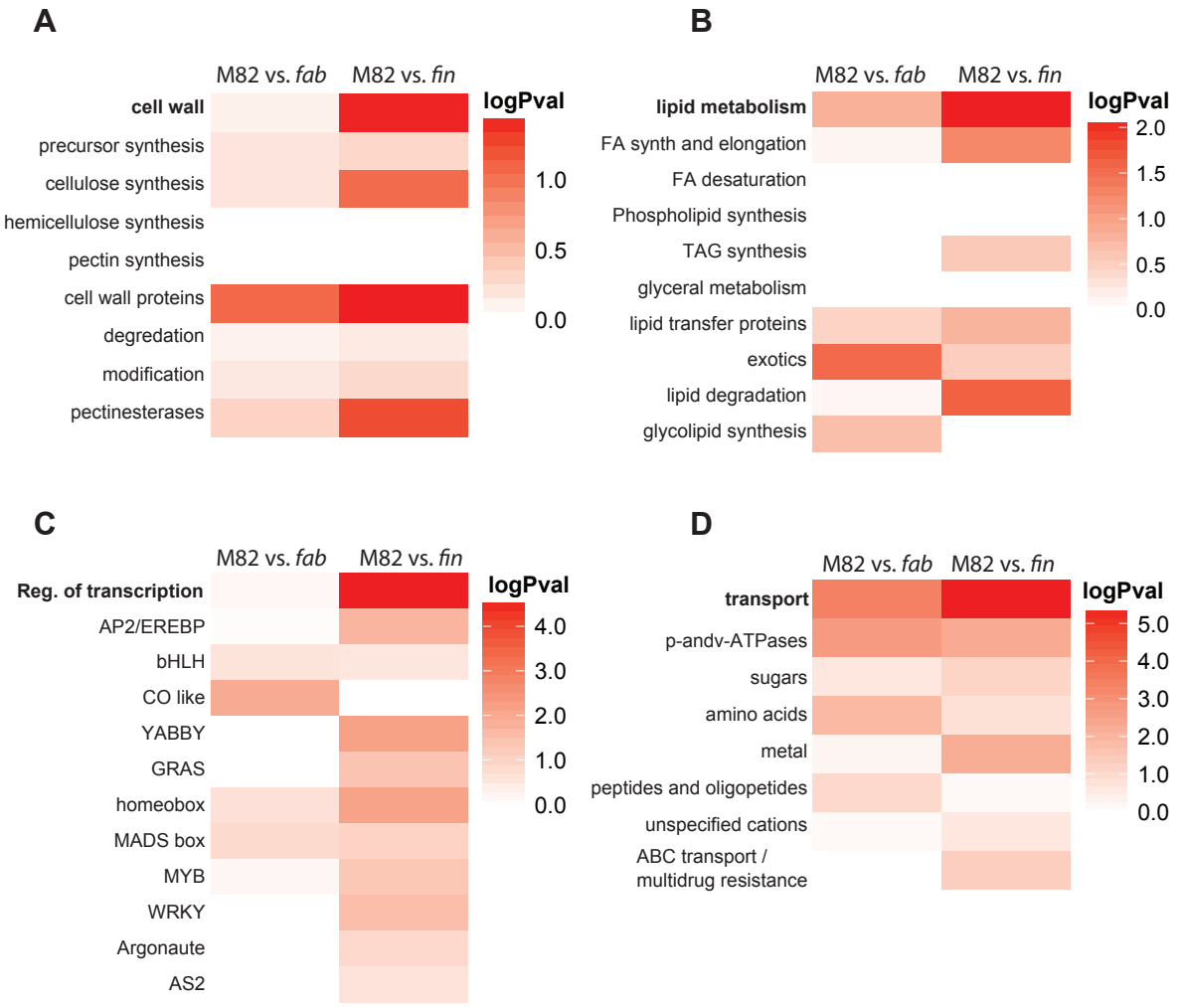


Figure 4.12. Sub-categories of differentially expressed genes from MapMan analysis. Sub-categories derived from main functional groups (Figure 6) that are enriched in *fab* and/or *fin* mutants. Selected sub-categories for Cell Wall (A), Lipid Metabolism (B), Regulation of Transcription/transcription factor families (C), Transport (D) that were determined differentially expressed and overrepresented as derived from Fisher's exact tests of the proportions of the selected categories in the mutants versus the proportions in the genome annotation. Scaled $-\log_{10}$P-values are represented in the heat-map; red indicates a significant overrepresentation of genes that are differentially expressed in a particular gene category for the given mutant.

A

CLE	dodecapeptide	Gene ID
AtCLV3	RTVPSGPDPLHH	AT2G27250
	: ... :	
SlCLE1	RRVPSEPDPIHN	Solyc02g067550
SlCLE2	RRVPNGPDPIHN	Solyc05g006610
SlCLE3	RKVRKGSDPIHN	Solyc05g007650
SlCLE4	RRVPTGPNAIHN	Solyc02g087470
SlCLE5	REVPSSPDPLHN	Solyc01g098890
SlCLE6	REAPMSPDPLHH	Solyc06g074060
SlCLE7/SlCLV3	RGVPAGPDPLHH	Solyc11g071380
SlCLE8	RQIPTGPDPLHH	Solyc01g014100
SlCLE9	RTAPTGPSPIHH	Solyc10g019090
SlCLE10	RVVPGGPDSQHH	Solyc07g053370
SlCLE11	RVAPQGPDAQHH	Solyc07g062670
SlCLE12	HDVPSGANPVQN	Solyc05g053630
SlCLE13	HEVPSGANPESN	Solyc09g091810
SlCLE14	HEVPSGPNPISN	Solyc09g061410
SlCLE15	HKPPSGPNPNGN	Solyc11g066120
SlCLE16	HLVPGGPNPLHN	Solyc03g025960

B

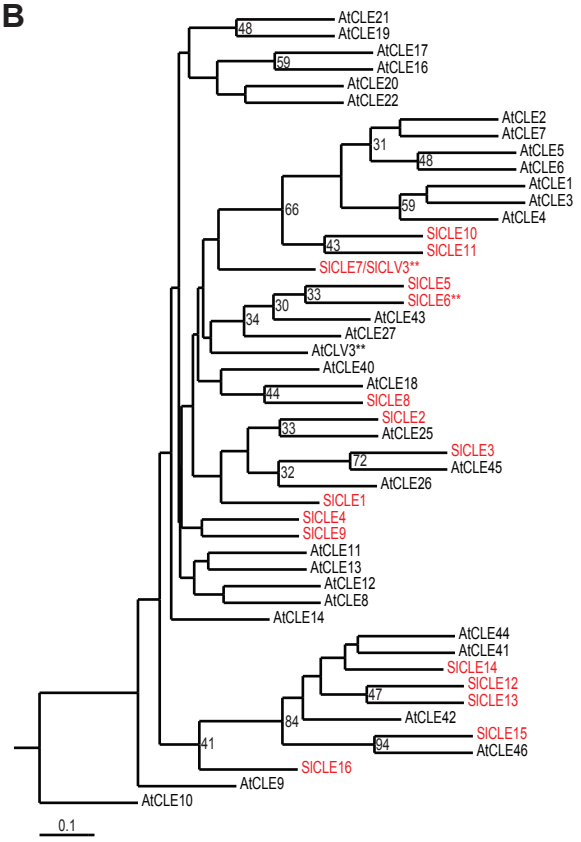
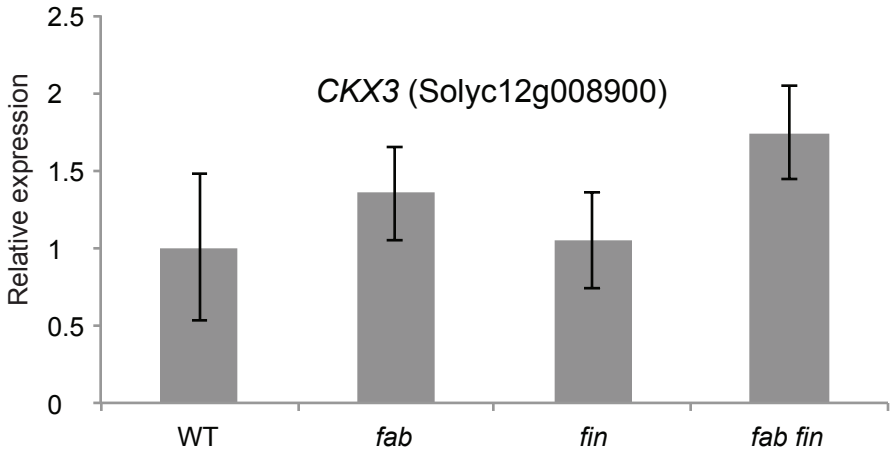


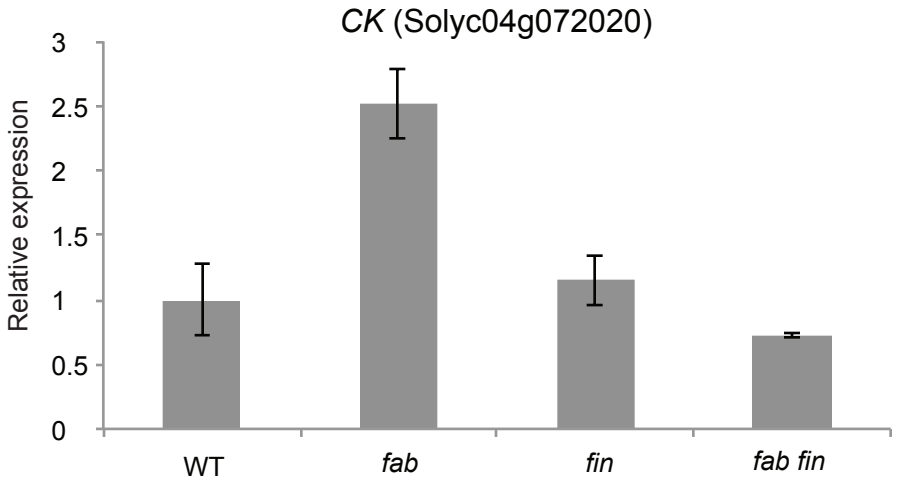
Figure 4.13. Alignment and phylogenetic analysis of tomato CLE peptides. (A) Alignment of the CLE dodecapeptide motif of annotated tomato CLEs. Sixteen tomato CLEs were identified by BLASTP analysis using the 31 *Arabidopsis* dodecapeptide sequences and a multiple sequence alignment was constructed using ClustalW2. Residues are colored based on their physiochemical properties and consensus symbols are shown at the top of the tomato sequences indicating positions with strongly similar properties (:) and weakly similar properties (.) with the *Arabidopsis thaliana* CLV3 dodecapeptide sequence shown for reference. Eleven CLEs (CLE1-11) start with an Arginine residue (R), characteristic of the CLV3 peptide class and five (SlCLE12-16) start with a Histidine residue (H), characteristic of the TDIF peptide class (Ito et al. 2006; Sawa et al. 2006; Jun et al. 2008). (B) An unrooted neighbor-joining phylogenetic tree of dodecapeptide CLE motif from 32 *Arabidopsis* CLEs and 16 tomato CLEs. Tomato CLEs are colored red. *Arabidopsis* CLV3, tomato CLV3 (SlCLE7) and the additional CLE highly overexpressed in *fin* mutants (SlCLE6) are highlighted with asterisks (**). Bootstrap values of 30% and above from 100 replicates are included at the nodes and the scale bar indicates the number of amino acid substitutions per site.

Figure 4.14. qRT-PCR of stem cell marker genes reveals the stem cell niches of *fab* and *fin* mutants are proportional to meristem size. (A-C) qPCR for tomato homologs of three stem cell markers identified in *Arabidopsis* (Aggarwal et al. 2010). Bars represent the fold change of gene expression over wild type (mean ± SE for two technical replicates), normalized to levels of *UBIQUITIN*. (A) qPCR of a *CKX3* homolog (Solyc12g008900), which is enriched in *WUS* and *CLV3* positive cells. (B) qPCR of a Choline Kinase (*CK)* homolog (Solyc04g072020), which is specific to *WUS* positive cells. (C) qPCR of an AtPUM-10 homolog (Solyc04g073980), which is specific to *CLV3* positive cells.

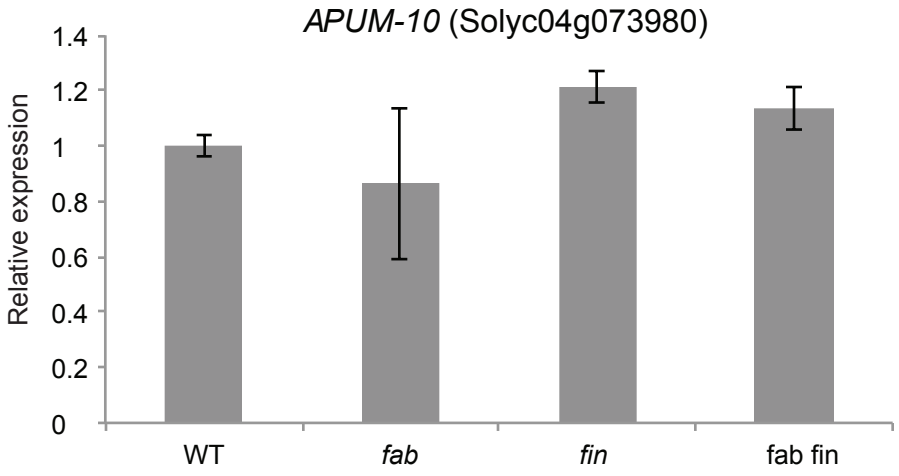
A


CKX3 (Solyc12g008900)
Relative expression
2.5
2
1.5
1
0.5
0
WT
fab
fin
fab fin

B


CK (Solyc04g072020)
Relative expression
3
2.5
2
1.5
1
0.5
0
WT
fab
fin
fab fin

C


APUM-10 (Solyc04g073980)
Relative expression
1.4
1.2
1
0.8
0.6
0.4
0.2
0
WT
fab
fin
fab fin

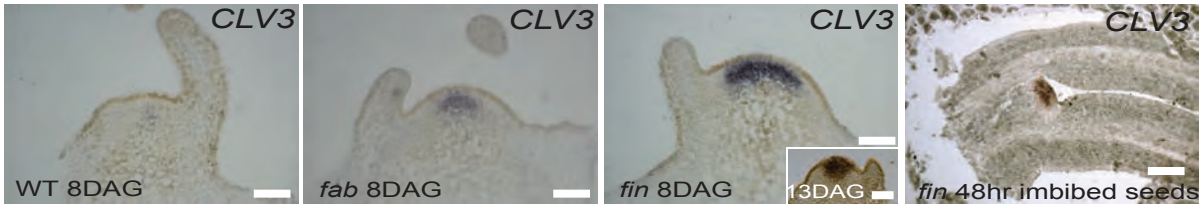


Figure 4.15. *CLV3* expression domain is expanded in *fin* mutants. *In situ* hybridization of meristems of the indicated genotype and age (lower left of each panel) probed with *CLV3*. *CLV3* is lowly expressed and the transcript is restricted to a few cells in WT early vegetative meristems (EVM/8dag; far left panel, images shown are from sections fixed after 48 hours of detection). Overexpression of *CLV3* and lateral expression of the expression domain is evident in both *fab* and *fin* EVMs (middle panels; images shown are from sections fixed after 12 hours of detection, however this same level and pattern of expression was evident within 4 hours of detection). High levels of *CLV3* transcript were detected throughout the entire shoot apical meristem of mature *fin* embryos before true leaves are evident (far right panel; embryo dissected from mature seeds imbibed in water for 48 hours). Scale bars = 100μm.

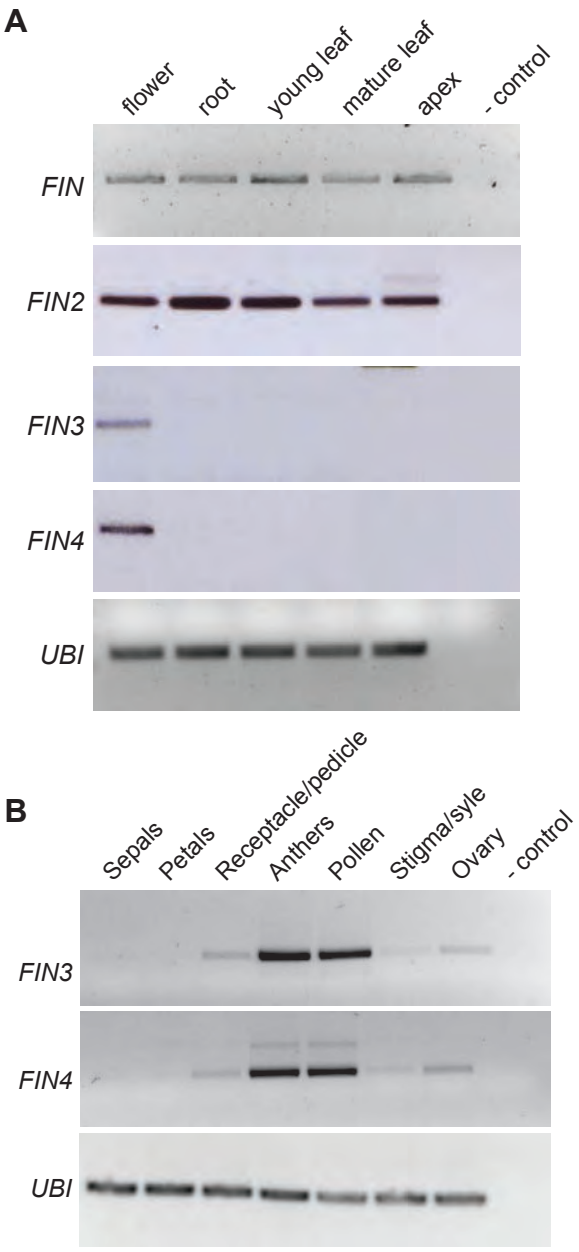


Figure 4.16. Expression analysis suggests specialized function within the flower for two tomato *FIN* genes. (A) Semi-quantitative RT-PCR of the four tomato *FIN* family members on a panel of tissues indicates broad expression of *FIN* (this study) and *FIN2* in all tissue types, but specific expression of *FIN3* and *FIN4* in the flower. Tomato *UBIQUITIN* (*UBI*) was used as a loading control. (B) Semi-quantitative RT-PCR of *FIN3* and *FIN4* on a panel of floral organ tissues. Expression of both genes is enriched within the pollen and anthers (which contain pollen). Tomato *UBIQUITIN* (*UBI*) was used as a loading control.

Table 4.1. DE gene list for *fab* vs. WT. (see electronic files) Transcriptome profiling results including gene ID, logFC, logCPM, P-value, raw counts, normalized counts, and ITAG descriptions are listed. Tab 1 lists genes in order of gene ID and also includes a list of the closest *Arabidopsis* orthologs. Tab 2 lists genes in order of expression level from most downregulated to most upregulated.

Table 4.2. DE gene list for *fin* vs WT. (see electronic files) Transcriptome profiling results including gene ID, logFC, raw counts, normalized counts, and ITAG descriptions are listed. Tab 1 lists genes in order of gene ID and also includes a list of the closest *Arabidopsis* orthologs. Tab 2 lists genes in order of expression level from most downregulated to most upregulated.

Table 4.3. Top 25 downregulated and upregulated genes in *fab*. Top 25 downregulated (highlighted blue) and upregulated genes (highlighted red). Gene ID, logFC, logCPM, P-value and ITAG descriptions are listed.

Gene ID	logFC	logCPM	PValue	ITAG desc
Solyc07g025170	-9.8603338	1.41934142	2.16E-17	Exocyst complex component 1 (AHRD V1 *-*- Q6P1Y9_MOUSE); contains Interpro domain(s) IPR019160 Exocyst complex, component Exoc1
Solyc10g054870	-7.6509639	-0.7558089	7.67E-09	Triosephosphate isomerase (AHRD V1 ***- A9PH17_POPTR); contains Interpro domain(s) IPR000652 Triosephosphate isomerase
Solyc07g025160	-6.756478	4.47693492	1.15E-24	Exocyst complex component 1 (AHRD V1 ***- Q5PPR2_MOUSE); contains Interpro domain(s) IPR019160 Exocyst complex, component Exoc1
Solyc03g078330	-4.6768525	-0.3567603	2.33E-08	Fatty acid elongase 3-ketoacyl-CoA synthase (AHRD V1 **** Q6DUV5_BRANA); contains Interpro domain(s) IPR012392 Very-long-chain 3-ketoacyl-CoA synthase
Solyc12g035550	-4.5086838	0.41399289	4.31E-10	Ycf1 (AHRD V1 ***- D2KLR1_OLEEU); contains Interpro domain(s) IPR008896 Ycf1
Solyc06g069870	-3.6337917	0.12849979	9.24E-08	Dehydration-responsive family protein (AHRD V1 ***- D7KPN7_ARALY); contains Interpro domain(s) IPR004159 Protein of unknown function DUF248, methyltransferase putative
Solyc07g019430	-3.5348014	0.95224144	4.00E-09	Unknown Protein (AHRD V1)
Solyc01g087340	-3.4284102	1.33721489	9.29E-08	Protein binding protein (AHRD V1 ***- D7L741_ARALY)
Solyc07g064870	-3.3419345	-0.0922063	2.71E-05	Endoglucanase 1 (AHRD V1 ***- B6U0J0_MAIZE); contains Interpro domain(s) IPR001701 Glycoside hydrolase, family 9
Solyc12g011430	-3.3419345	-0.1477534	1.08E-06	Hydroxyacylglutathione hydrolase (AHRD V1 **** D9J0E0_9CARY); contains Interpro domain(s) IPR001279 Beta-lactamase-like
Solyc09g007790	-3.2088084	0.4684401	1.06E-06	Arabidopsis thaliana genomic DNA chromosome 5 P1 clone MOK16 (AHRD V1 ***- Q9LYW2_ARATH); contains Interpro domain(s) IPR007608 Protein of unknown function DUF584
Solyc02g063040	-3.1521546	0.655311	1.84E-07	Unknown Protein (AHRD V1)
Solyc01g109390	-2.827643	-0.2295942	8.68E-06	PVR3-like protein (AHRD V1 ***- B6TVY8_MAIZE); contains Interpro domain(s) IPR013770 Plant lipid transfer protein and hydrophobic protein, helical
Solyc03g006830	-2.7835749	1.30721379	4.78E-06	MADS-box transcription factor 2 (AHRD V1 **** D1MFS7_HEVBR); contains Interpro domain(s) IPR002100 Transcription factor, MADS-box
Solyc02g076690	-2.7598586	1.87290377	1.42E-07	Cathepsin B-like cysteine proteinase (AHRD V1 **-* CYSP_SCHMA); contains Interpro domain(s) IPR013128 Peptidase C1A, papain
Solyc06g073560	-2.7012765	0.65159752	8.09E-07	Isovaleryl-CoA dehydrogenase (AHRD V1 **** Q0MX57_BETVU); contains Interpro domain(s) IPR009100 Acyl-CoA dehydrogenase/oxidase, middle and N-terminal
Solyc10g047110	-2.6632428	-0.3630863	2.47E-05	Peroxidase (AHRD V1 **** Q5K4K5_GOSHI); contains Interpro domain(s) IPR002016 Haem peroxidase, plant/fungal/bacterial
Solyc05g010680	-2.6610254	0.01528856	0.00032212	Abhydrolase domain-containing protein FAM108B1 (AHRD V1 **-- C0HAS5_SALSA); contains Interpro domain(s) IPR000073 Alpha/beta hydrolase fold-1
Solyc10g079860	-2.6255242	-0.0852916	1.58E-05	Beta-1 3-glucanase (AHRD V1 ***- Q68V46_OLEEU); contains Interpro domain(s) IPR000490 Glycoside hydrolase, family 17
Solyc11g073120	-2.6126583	0.56323036	2.08E-06	MYB transcription factor (AHRD V1 *-*- Q56UT4_ORYSJ); contains Interpro domain(s) IPR015495 Myb transcription factor
Solyc09g066000	-2.5891274	-0.1105421	1.84E-05	Prenylated rab acceptor PRA1 (AHRD V1 ***- A2Q469_MEDTR); contains Interpro domain(s) IPR004895 Prenylated rab acceptor PRA1
Solyc11g045520	-2.5806515	0.39895497	0.00022014	1-AMINOCYCLOPROPANE-1-CARBOXYLATE OXIDASE-like protein (AHRD V1 **-- Q94CL5_ARATH); contains Interpro domain(s) IPR005123 Oxoglutarate and iron-dependent oxygenase
Solyc01g106380	-2.5669479	1.82694069	2.21E-05	Genomic DNA chromosome 5 BAC clone T12B11 (AHRD V1 ***- Q9FGE8_ARATH)
Solyc02g089770	-2.5637628	1.16065225	2.71E-05	Dihydroflavonol-4-reductase (AHRD V1 ***- B6TK03_MAIZE); contains Interpro domain(s) IPR016040 NAD(P)-binding domain
Solyc02g092720	-2.5537548	0.29239808	6.39E-06	Unknown Protein (AHRD V1)
Solyc00g021630	2.79936425	-0.4049867	4.53E-05	NADH-ubiquinone oxidoreductase chain 6 (Fragment) (AHRD V1 ***- Q2PZB4_BRAOB); contains Interpro domain(s) IPR001457 NADH:ubiquinone/plastoquinone oxidoreductase, chain 6
Solyc06g074750	2.80536034	1.60452015	2.71E-07	Histone H2B (AHRD V1 *-*- B6SGC3_MAIZE); contains Interpro domain(s) IPR000558 Histone H2B
Solyc04g078060	2.86369297	-0.1304198	3.14E-05	Small hydrophobic protein (AHRD V1 *-*- Q9ZQY8_ARATH)
Solyc10g047530	2.89004938	-0.5597884	1.81E-05	Phototropic-responsive NPH3 family protein (AHRD V1 **-- D7KEU3_ARALY); contains Interpro domain(s) IPR004249 NPH3
Solyc10g054900	2.89004938	-0.6204044	3.21E-05	Proline-rich protein (AHRD V1 ***- Q9M6T7_NICGL); contains Interpro domain(s) IPR010616 Protein of unknown function DUF1210
Solyc03g117390	2.89852353	-0.324558	2.72E-05	Unknown Protein (AHRD V1)
Solyc11g020330	2.94318225	2.12004083	5.70E-07	class IV heat shock protein (AHRD V1 ***- B6TXB5_MAIZE); contains Interpro domain(s) IPR002068 Heat shock protein Hsp20
Solyc08g075760	2.98198061	0.09808744	1.54E-06	CYCLOPS (AHRD V1 ***- A7TUE1_MEDTR)
Solyc11g071430	2.98593346	0.82279685	3.70E-07	Cc-nbs-lrr, resistance protein
Solyc08g066790	3.01934855	0.01361881	9.77E-06	Phospholipase D (AHRD V1 **** Q9FR61_SOLLC); contains Interpro domain(s) IPR011402 Phospholipase D, plant
Solyc04g078040	3.05293336	0.77724847	6.42E-07	Mpv17 protein (AHRD V1 **-- Q178F4_AEDAE); contains Interpro domain(s) IPR007248 Mpv17/PMP22
Solyc10g086620	3.07847563	-0.1343432	4.63E-06	Tropinone reductase (AHRD V1 **-- B6TAE7_MAIZE); contains Interpro domain(s) IPR002347 Glucose/ribitol dehydrogenase
Solyc04g064620	3.19291681	0.25811322	3.61E-07	RAG1-activating protein 1 homolog (AHRD V1 **-- R1AP1_DROPS); contains Interpro domain(s) IPR018179 RAG1-activating protein 1 homologue
Solyc06g074760	3.25689073	0.18167832	1.03E-06	Neuralized (Fragment) (AHRD V1 *--- Q177Q7_AEDAE); contains Interpro domain(s) IPR001841 Zinc finger, RING-type
Solyc07g063110	3.25779624	-0.2199689	1.08E-06	Beta-1,3-galactosyl-O-glycosyl-glycoprotein beta-1,6-N-acetylglucosaminyltransferase (AHRD V1 *--* GCNT1_BOVIN); contains Interpro domain(s) IPR003406 Glycosyl transferase, family 14
Solyc03g111810	3.34768488	-0.1843633	9.20E-07	Sieve element-occluding protein 3 (AHRD V1 **-- B5THF7_MEDTR)
Solyc05g005130	3.63346878	1.7949169	3.08E-10	Cc-nbs-lrr, resistance protein with an R1 specific domain
Solyc03g025700	3.97402359	0.52184914	1.04E-09	Pentatricopeptide repeat-containing protein (AHRD V1 ***- D7KHY5_ARALY); contains Interpro domain(s) IPR002885 Pentatricopeptide repeat
Solyc01g098370	4.02324516	-0.0702339	3.14E-08	LRR receptor-like serine/threonine-protein kinase, RLP
Solyc07g041260	4.02324516	-0.0068861	1.67E-08	Histidine decarboxylase (AHRD V1 *-*- Q9MA74_ARATH); contains Interpro domain(s) IPR002129 Pyridoxal phosphate-dependent decarboxylase
Solyc06g072000	4.19273051	0.336566	4.31E-10	Katanin p60 ATPase-containing subunit A-like 1 (AHRD V1 *--- KATL1_DANRE); contains Interpro domain(s) IPR003959 ATPase, AAA-type, core
Solyc06g074430	4.71606208	8.3990531	5.15E-17	60s acidic ribosomal protein-like protein (AHRD V1 ***- Q3HVP0_SOLTU); contains Interpro domain(s) IPR001813 Ribosomal protein 60S
Solyc11g072790	6.04440934	0.56468089	1.70E-13	WUSCHEL-related homeobox 3B (AHRD V1 *-*- A8Y7Y2_MAIZE); contains Interpro domain(s) IPR001356 Homeobox
Solyc06g054090	7.51437747	-1.0823117	3.53E-08	Ribosomal-protein-alanine N-acetyltransferase (AHRD V1 **** D3FS71_BACPE); contains Interpro domain(s) IPR000182 GCN5-related N-acetyltransferase
Solyc10g076330	8.23373164	-0.3274605	5.68E-11	Kelch-like protein (AHRD V1 ***- Q84S70_ORYSJ); contains Interpro domain(s) IPR013989 Development and cell death domain

Table 4.4. Top 25 downregulated and upregulated genes in *fin*. Top 25 downregulated (highlighted blue) and upregulated genes (highlighted red). Gene ID, logFC, logCPM, P-value and ITAG descriptions are listed.

Gene ID	logFC	logCPM	PValue	ITAG desc	Notes
Solyc05g014280	-10.19873131	1.351454	4.28E-18	Heat shock protein (AHRD V1 ***- A9QVH3_9FABA); contains Interpro domain(s) IPR002068 Heat shock protein Hsp20	
Solyc10g086250	-10.00318519	1.195659	2.02E-17	MYB transcription factor (AHRD V1 **-- A9YY82_SOLTU); contains Interpro domain(s) IPR015495 Myb transcription factor	
Solyc04g074430	-9.199958366	0.353567	6.47E-14	Phi-1 protein (Fragment) (AHRD V1 *-*- Q4LAX0_CAPCH); contains Interpro domain(s) IPR006766 Phosphate-induced protein 1 conserved region	
Solyc04g074440	-8.889347112	0.06843	8.38E-13	Os06g0220000 protein (Fragment) (AHRD V1 ***- Q0DDJ2_ORYSJ); contains Interpro domain(s) IPR006766 Phosphate-induced protein 1 conserved region	
Solyc08g066270	-8.889347112	0.057332	9.37E-13	Receptor like kinase, RLK	
Solyc07g006730	-8.862437467	0.057809	9.37E-13	Multidrug resistance protein mdtK (AHRD V1 ***- MDTK_CITK8); contains Interpro domain(s) IPR015521 MATE family transporter related protein	
Solyc02g080220	-8.749479868	-0.05633	2.36E-12	Pectinesterase (AHRD V1 ***- B9RXQ4_RICCO); contains Interpro domain(s) IPR000070 Pectinesterase, catalytic	
Solyc03g081240	-8.626921055	-0.17519	7.17E-12	Pseudo response regulator (AHRD V1 ***- B9IDF0_POPTR); contains Interpro domain(s) IPR001789 Signal transduction response regulator, receiver region	
Solyc09g006010	-8.527646024	-0.30338	2.07E-11	Pathogenesis related protein PR-1 (AHRD V1 ***- Q9SC15_SOLTU); contains Interpro domain(s) IPR001283 Allergen V5/Tpx-1 related	
Solyc11g006300	-8.527646024	-0.29535	2.07E-11	3-oxo-5-alpha-steroid 4-dehydrogenase family protein (AHRD V1 ***- D7M7I5_ARALY); contains Interpro domain(s) IPR001104 3-oxo-5-alpha-steroid 4-dehydrogenase, C-terminal	
Solyc10g079860	-8.457452147	-0.36139	3.61E-11	Beta-1 3-glucanase (AHRD V1 ***- Q68V46_OLEEU); contains Interpro domain(s) IPR000490 Glycoside hydrolase, family 17	
Solyc02g090560	-8.38366756	-0.43074	6.44E-11	Calcium-transporting ATPase 1 (AHRD V1 ***- Q7XBH9_CERRI); contains Interpro domain(s) IPR001757 ATPase, P-type, K/Mg/Cd/Cu/Zn/Na/Ca/Na/H-transporter	
Solyc08g075630	-5.771638714	0.069437	8.77E-11	Cc-nbs-lrr, resistance protein	
Solyc09g011660	-5.516576354	0.748082	1.18E-12	Universal stress protein 1 (AHRD V1 ***- B0YQX0_GOSAR); contains Interpro domain(s) IPR006016 UspA	
Solyc08g068710	-5.202015639	-0.54079	2.63E-09	N-acetyltransferase (AHRD V1 ***- B6SUK9_MAIZE); contains Interpro domain(s) IPR000182 GCN5-related N-acetyltransferase	
Solyc07g043150	-5.136908819	0.362509	4.03E-11	UDP-glucosyltransferase (AHRD V1 **** A7M6U9_IPONI); contains Interpro domain(s) IPR002213 UDP-glucuronosyl/UDP-glucosyltransferase	
Solyc04g007050	-5.117410871	-0.60803	2.67E-08	Cc-nbs-lrr, resistance protein	
Solyc12g009270	-5.073172111	-0.6567	4.31E-08	Pectinesterase (AHRD V1 **** C0PST8_PICSI); contains Interpro domain(s) IPR006501 Pectinesterase inhibitor	
Solyc06g069000	-5.042448784	0.269047	3.80E-10	Uncharacterized GPI-anchored protein At4g28100 (AHRD V1 ***- UGPI7_ARATH)	
Solyc09g083290	-4.795056956	0.570072	1.03E-10	Auxin responsive protein (AHRD V1 ***- D9IQE6_CATRO); contains Interpro domain(s) IPR003311 AUX/IAA protein	
Solyc03g020040	-4.446635481	-0.33155	1.60E-08	Proteinase inhibitor II (AHRD V1 ***- B3F0C1_TOBAC); contains Interpro domain(s) IPR003465 Proteinase inhibitor I20, Pin2	
Solyc05g007710	-4.434667396	0.61641	6.81E-10	MYB transcription factor (AHRD V1 *--- Q6R098_ARATH); contains Interpro domain(s) IPR015495 Myb transcription factor	
Solyc10g086530	-4.130115438	-0.09539	3.72E-08	GRAS family transcription factor (AHRD V1 **-- B9HBM9_POPTR); contains Interpro domain(s) IPR005202 GRAS transcription factor	
Solyc08g068260	-4.116515388	-0.66559	9.11E-07	Unknown Protein (AHRD V1)	
Solyc11g012530	-4.098485914	-0.09805	3.72E-08	Unknown Protein (AHRD V1)	
Solyc07g053540	3.865773926	0.891653	1.06E-08	Fasciclin-like arabinogalactan protein 4 (AHRD V1 ***- A9XTK9_GOSHI); contains Interpro domain(s) IPR000782 FAS1 domain	
Solyc12g077420	3.882836146	-0.58061	7.79E-07	Zinc finger protein 207 (AHRD V1 ***- Q28FC3_XENTR)	
Solyc04g025510	3.941339711	1.207644	2.15E-09	Os02g0194300 protein (Fragment) (AHRD V1 *-*- Q0E347_ORYSJ); contains Interpro domain(s) IPR014854 Nse4	
Solyc12g098450	3.958071231	0.443927	2.55E-08	DUF803 domain membrane protein (AHRD V1 **-- Q4WRL6_ASPFU); contains Interpro domain(s) IPR008521 Protein of unknown function DUF803	
Solyc06g074060	4.185500999	0.657859	3.72E-09	Unknown Protein (AHRD V1)	CLE6
Solyc10g008120	4.241407089	0.288251	1.48E-08	O-methyltransferase 3 (AHRD V1 ***- B0ZB57_HUMLU); contains Interpro domain(s) IPR016461 O-methyltransferase, COMT, eukaryota	
Solyc02g077270	4.344358276	-0.15266	2.12E-08	Calcium ion binding protein (AHRD V1 *-*- B6U1M5_MAIZE); contains Interpro domain(s) IPR011992 EF-Hand type	
Solyc05g009480	4.359830902	0.422066	4.53E-09	Unknown Protein (AHRD V1)	
Solyc04g071400	4.453013809	-0.02845	7.23E-09	Nodulin-like protein (AHRD V1 ***- B6TIX8_MAIZE); contains Interpro domain(s) IPR000620 Protein of unknown function DUF6, transmembrane	
Solyc06g072000	4.521021669	0.563096	1.28E-09	Katanin p60 ATPase-containing subunit A-like 1 (AHRD V1 *--- KATL1_DANRE); contains Interpro domain(s) IPR003959 ATPase, AAA-type, core	
Solyc07g064380	4.737106333	0.224947	8.34E-10	Serine/threonine-protein phosphatase 7 long form homolog (AHRD V1 *--- PPP7L_ARATH); contains Interpro domain(s) IPR019557 Aminotransferase-like, plant mobile domain	
Solyc02g093740	4.993629678	1.015201	1.88E-11	MYB transcription factor (AHRD V1 *-** Q9LY95_ARATH); contains Interpro domain(s) IPR015495 Myb transcription factor	
Solyc11g071380	5.022086512	6.37405	5.87E-16	Unknown Protein (AHRD V1)	CLV3
Solyc05g005130	5.130856006	3.22151	2.37E-15	Cc-nbs-lrr, resistance protein with an R1 specific domain	
Solyc11g072500	5.367994814	1.381945	5.54E-13	Dof zinc finger protein 12 (AHRD V1 *-*- A5HWG1_HORVD); contains Interpro domain(s) IPR003851 Zinc finger, Dof-type	
Solyc11g069770	5.549603376	0.101702	9.78E-11	Transcription factor MADS-box (AHRD V1 *-*- A4Q7P2_MEDTR); contains Interpro domain(s) IPR002100 Transcription factor, MADS-box	
Solyc12g038510	5.742993607	2.487324	3.66E-16	MADS box transcription factor 11 (AHRD V1 ***- D9IFM5_ONCHC); contains Interpro domain(s) IPR002487 Transcription factor, K-box	
Solyc06g062960	6.20895747	0.736078	2.76E-13	Unknown Protein (AHRD V1); contains Interpro domain(s) IPR012876 Protein of unknown function DUF1677, plant	
Solyc04g008970	6.223584068	0.747365	2.56E-13	FIP1 (AHRD V1 **-- B6U778_MAIZE); contains Interpro domain(s) IPR004182 GRAM	
Solyc12g036470	7.885676628	-0.87053	3.56E-09	Transcription factor (AHRD V1 *-*- D6MKM4_9ASPA); contains Interpro domain(s) IPR011598 Helix-loop-helix DNA-binding	
Solyc06g076320	7.929885879	-0.82795	2.44E-09	DVL13 (AHRD V1 *-*- Q8L7D0_ARATH); contains Interpro domain(s) IPR012552 DVL	
Solyc11g006310	8.310380003	-0.45394	1.17E-10	GEX1 (Fragment) (AHRD V1 ***- Q49L63_ARATH)	
Solyc07g006570	8.46860427	-0.28715	2.73E-11	S8-RNase (Fragment) (AHRD V1 **-- Q5XPJ2_MALDO); contains Interpro domain(s) IPR001568 Ribonuclease T2	
Solyc10g061950	8.498268497	-0.26657	2.37E-11	Unknown Protein (AHRD V1)	
Solyc11g012650	8.690414776	-0.083	4.91E-12	TPD1 (AHRD V1 *-*- Q6TLJ2_ARATH)	

Chapter 4 Addendum

In Chapter 4, I described the characterization of two tomato mutants, *fasciated and branched* (*fab*) and *fasciated inflorescence* (*fin*), both of which develop fasciated flowers and fruits and branched inflorescences. While *FAB* encodes a receptor kinase protein that is the tomato ortholog of *Arabidopsis* CLAVATA1, *FIN* was described as encoding a small transmembrane protein of unknown function. Aside from an N-terminal signal peptide and a predicted transmembrane domain, no other conserved motifs were found that gave insight into its function.

However, an exciting development came just after the submission of my thesis. The *Arabidopsis* FIN homologs were identified as hydroxyproline O-arabinosyltransferase (HPAT) proteins (Ogawa-Ohnishi et al. 2013). HPATs are specialized glycosyltransferase enzymes that catalyze the addition of arabinose sugar moieties onto the hydroxyl group of hydroxyproline (Hyp) residues, and additional enzymes are thought to further lengthen the arabinose chain (Egelund et al. 2007; Gille et al. 2009; Velasquez et al. 2011). Arabinosylation is an important post-translational modification in plants that is commonly found on hydroxyproline-rich glycoproteins (HRGPs), such as extensins and arabinogalactan proteins, which are critical components of the plant cell wall and essential for cell wall integrity, cell-cell interactions, and defense (reviewed in; Ellis et al. 2010; Lamport et al. 2011). Moreover, a number of plant peptide signaling hormones with varying roles in development and stress response are also arabinosylated (reviewed in; Matsubayashi 2012). This finding was particularly intriguing because CLE signaling peptides, including CLV3p, are one class of peptides known to have this modification, most typically with a tri-arabinose chain on the conserved Hyp7 residue within the dodecapeptide CLE motif (Ohyama et al. 2009). Hyp7 modification of CLV3p is thought to induce a conformational change in the peptide that may increase its stability and/or signaling potency (e.g. modified peptide may be bound more efficiently by CLV1 and other receptor complexes) (Ohyama et al. 2009). However, the importance of this modification, particularly in meristem

maintenance, is still somewhat controversial. For instance, *Arabidopsis clv3* meristem defects can be rescued by expression of a transgenic form of CLV3 in which the critical proline-7 within the CLE motif is mutated to alanine (Song et al. 2012). Moreover, unmodified exogenous CLE peptides sufficiently rescue the *clv3* meristem defect at high levels in *Arabidopsis* (Shinohara and Matsubayashi 2013).

In contrast to tomato, single knockouts of the *Arabidopsis* HPATs do not have visible phenotypes, suggesting a redundancy among the three homologs (Ogawa-Ohnishi et al. 2013; CAM, personal communication). Furthermore, none of the *hpat* double mutant combinations present meristem defects either, however, pollen transmission defects were discovered when attempting to make the double mutant combinations (CAM, personal communication; see Chapter 5, general discussion, for more details). It is possible that the *Arabidopsis* triple *hpat* mutants will have meristem defects or it is also possible that the meristem defect is specific to tomato. Nonetheless, the tomato *fin* mutant phenotypes suggest that this modification is indeed critical for proper meristem maintenance in tomato.

Due to the extremely enhanced fasciation of *fab fin* double mutants, my original interpretation was that these two genes functioned in separate pathways. However, if FIN is responsible for arabinosylating CLV3p, and this modification is essential for CLV3p function, then FIN is likely acting upstream of CLV signaling. In hindsight, the molecular characterization of *fin* is consistent with this hypothesis. Transcriptome profiling (verified by qRT-PCR) revealed that *SlCLV3* and a second tomato *CLE*, *SlCLE6*, are highly overexpressed in *fin* mutant meristems, while *SlWUS* expression is only increased two-fold (Figure 4.11). Moreover, *in situ* hybridization demonstrated that both *SlWUS* and *SlCLV3* expression domains are greatly expanded in *fin* mutants (Figures 4.9 and 4.15). I hypothesize that in *fin* mutants, the SlCLE peptides are not modified properly with arabinose moieties, and this disrupts the normal balance of the CLV-WUS feedback loop. Specifically, less potent peptides are secreted to the apoplast reducing the perception of SlCLV3p levels in the underlying cells, perhaps due to decreased

binding affinity to SlCLV1 and parallel receptor complexes. Reduction in SlCLV3p perception leads to increased *SlWUS* expression, which feeds back to upregulate *SlCLV3* and potentially additional *SlCLEs* (Figure 4.17). Thus, extremely high levels of unmodified CLEs may accumulate before there is repression of *WUS*. Under normal circumstances, the extra accumulation of properly CLV3p would lead to smaller meristems due to the downregulation of *WUS*. However, unmodified CLV3p is likely not as stable or not recognized as efficiently by the receptor complexes responsible for signaling to repress *WUS*, and therefore, although CLV3p accumulates, *WUS* levels remain high, resulting in meristem fasciation.

I have been leading efforts on a series of follow-up experiments to test the functional link between *FIN* and arabinosylation in the control of meristem size in tomato. We are taking several approaches to test the arabinosyltransferase activity of FIN. In collaboration with the proteomics facility at CSHL and with a group at U.C. Berkley, we are investigating whether there are differences in arabinosylation levels between wild type and *fin* mutants *in planta*. First, we are looking both globally at the arabinosylation levels of cell wall components such as extensins, which are known to be Hyp-arabinosylated. Second, we are looking specifically for differences in arabinosylation levels in wild type versus *fin* tomato protoplasts overexpressing *SlCLV3*. We expect to see decreased levels of arabinosylation in *fin* for both experiments compared to wild type. Moreover, in collaboration with the Matsubayashi group, we will test whether tomato FIN, purified from stably expressing tobacco BY-2 cells, has the ability to catalyze the addition of arabinose to hydroxyprolinated *SlCLV3* peptide. Finally, in a complementary approach, I have begun to test whether exogenous CLV3 peptide can rescue the meristem defect in *fin* mutants. Currently, I have only been able to work with unmodified *Sl*CLV3p, however, to truly test the biological significance of glycosylation in meristem maintenance, properly modified SlCLV3p should be used. Again, we will work with the Matsubayashi group to synthesize and test these peptides on tomato *fin* meristem growth. If the *fin* meristem defect is a result of an inability to modify CLV3p, then modified peptides should fully rescue the meristem size. However, it is

possible that there are parallel pathways in which FIN acts to control meristem size, and therefore, we may only see a partial rescue. Indeed, the consequences of losing FIN function likely extends beyond CLV3p modification and the CLV-WUS signaling pathway. In addition to the CLEs and other signaling peptides, we expect that FIN also modifies HRGPs, and therefore, defects in cell wall structure and rigidity may also underlie *fin* fasciation (see Chapter 5, general discussion, for more details).

References

Egelund J, Obel N, Ulvskov P, Geshi N, Pauly M, Bacic A, Petersen BL. 2007. Molecular characterization of two Arabidopsis thaliana glycosyltransferase mutants, rra1 and rra2, which have a reduced residual arabinose content in a polymer tightly associated with the cellulosic wall residue. *Plant Mol Biol* **64**: 439-451.

Ellis M, Egelund J, Schultz CJ, Bacic A. 2010. Arabinogalactan-proteins: key regulators at the cell surface? *Plant Physiol* **153**: 403-419.

Gille S, Hansel U, Ziemann M, Pauly M. 2009. Identification of plant cell wall mutants by means of a forward chemical genetic approach using hydrolases. *Proceedings of the National Academy of Sciences of the United States of America* **106**: 14699-14704.

Lamport DT, Kieliszewski MJ, Chen Y, Cannon MC. 2011. Role of the extensin superfamily in primary cell wall architecture. *Plant Physiol* **156**: 11-19.

Matsubayashi Y. 2012. MBSJ MCC Young Scientist Award 2010. Recent progress in research on small post-translationally modified peptide signals in plants. *Genes Cells* **17**: 1-10.

Ogawa-Ohnishi M, Matsushita W, Matsubayashi Y. 2013. Identification of three hydroxyproline O-arabinosyltransferases in Arabidopsis thaliana. *Nat Chem Biol* **9**: 726-730.

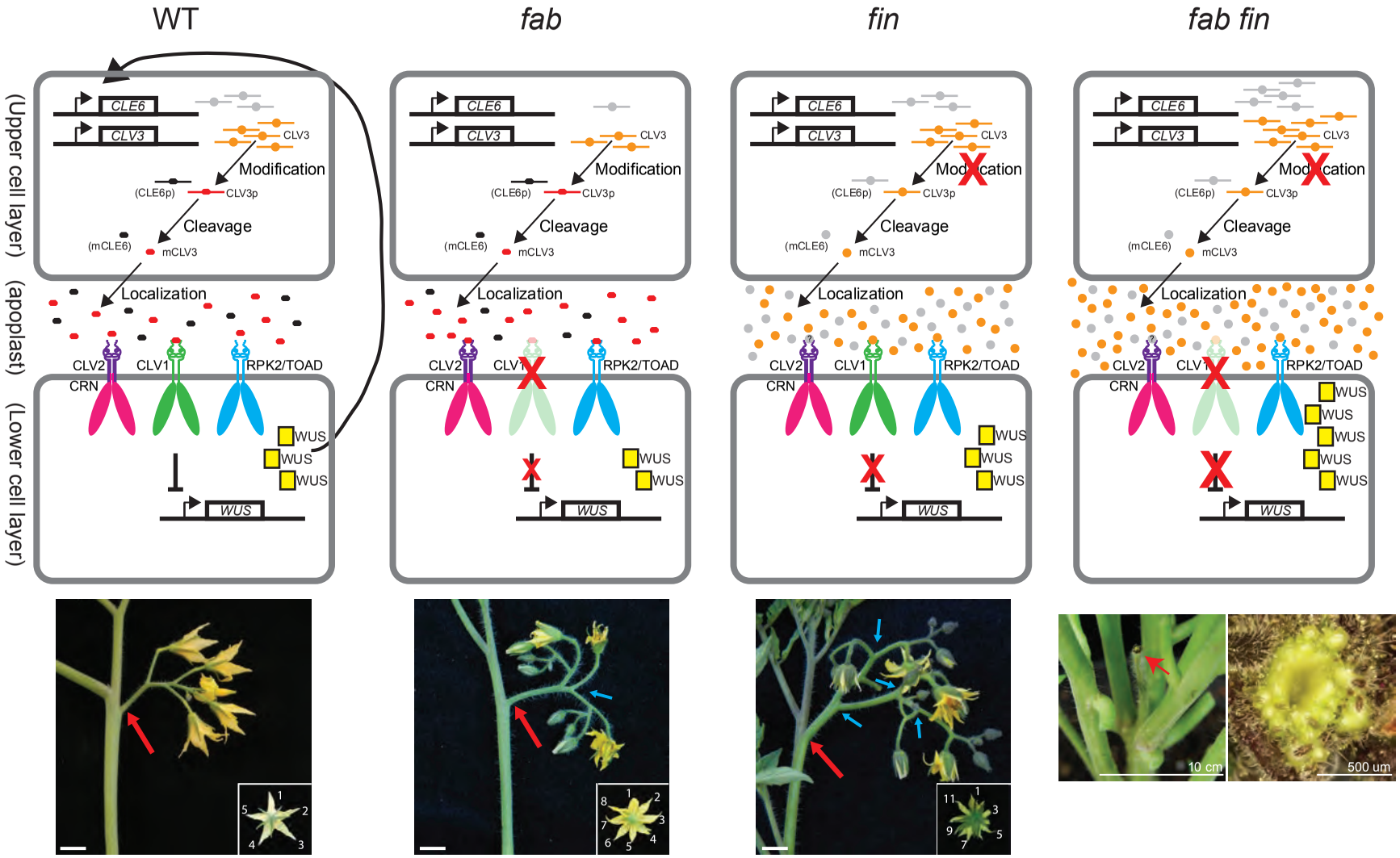
Ohyama K, Shinohara H, Ogawa-Ohnishi M, Matsubayashi Y. 2009. A glycopeptide regulating stem cell fate in Arabidopsis thaliana. *Nat Chem Biol* **5**: 578-580.

Shinohara H, Matsubayashi Y. 2013. Chemical synthesis of Arabidopsis CLV3 glycopeptide reveals the impact of hydroxyproline arabinosylation on peptide conformation and activity. *Plant Cell Physiol* **54**: 369-374.

Song XF, Yu DL, Xu TT, Ren SC, Guo P, Liu CM. 2012. Contributions of individual amino acid residues to the endogenous CLV3 function in shoot apical meristem maintenance in Arabidopsis. *Molecular plant* **5**: 515-523.

Velasquez SM, Ricardi MM, Dorosz JG, Fernandez PV, Nadra AD, Pol-Fachin L, Egelund J, Gille S, Harholt J, Ciancia M et al. 2011. O-glycosylated cell wall proteins are essential in root hair growth. *Science* **332**: 1401-1403.

Figure 4.17. The CLV-WUS feedback loop is progressively weakened in *fab*, *fin*, and *fab fin* mutants. From left to right, four panels are shown to depict a model of CLV-WUS signaling (top) and inflorescence phenotypes in WT, *fab*, *fin*, and *fab fin* situations (bottom). The model depicts two cell layers; the overlying cell layer expresses *SlCLV3* and potentially other CLEs like *SlCLE6*. Under WT conditions, *SlCLV3* is expressed (orange lines), modified with the addition of hydroxyprolines and Hyp O-arabinosylation, cleaved to a mature peptide (red polygons), and localized to the apoplast. Once in the apoplast, mature CLV3p is recognized by CLV1 (green receptor) and parallel receptor complexes, which are present in the underlying cell layers. Perception of CLV3p triggers a signaling cascade that represses *SlWUS* in these underlying cell layers. The WUS protein (yellow boxes) feeds back to upregulate the expression of *SlCLV3* in the overlying layers, thus canalizing CLV-WUS signaling. In the *fab* mutant situation, CLV3p is properly modified and localized, however, signaling through the CLV1 receptor is eliminated, which results in a slight derepression of *SlWUS*. A slight increase in WUS leads to a slight increase in *SlCLV3* as well, and hence, mild fasciation and branching. In *fin*, CLV3p is not properly modified and this impacts signaling through all of the parallel receptor complexes, leading to a higher level of *SlWUS* derepression and *SlCLV3* overexpression. *CLV3* overexpression would normally lead to smaller meristem size, however, without proper modification, this peptide is a less potent signaling molecule and the breaking of the CLV-WUS feedback loop ultimately leads to more extreme fasciation and branching. Finally, in the *fab fin* double mutants, the feedback loop is severely compromised. CLV3p is not properly modified and one of the major receptor complexes is lost, which results in the most extreme fasciation phenotype.


WT
fab
fin
fab fin
(Upper cell layer)
(apoplast)
(Lower cell layer)
CLE6
CLV3
CLV3
Modification
(CLE6p)
CLV3p
Cleavage
(mCLE6)
mCLV3
Localization
CLV2
CRN
CLV1
RPK2/TOAD
WUS
WUS
WUS
WUS
10 cm
500 um

5 Conclusions and Perspectives

Evolving models of meristem signaling

In characterizing the tomato inflorescence branching mutants *fasciated and branched* (*fab*) and *fasciated inflorescence* (*fin*) I found a connection between the control of meristem size and inflorescence branching (Chapter 4). I demonstrated that *FAB* is the ortholog of *Arabidopsis CLV1*, a component of a highly conserved plant meristem maintenance pathway, and importantly, this is the first component of the CLV-WUS pathway identified in tomato. Interestingly, while *FAB* encodes a known component of a meristem maintenance pathway, *FIN* encodes a glycosyltransferase, specifically a hydroxyproline O-arabinosyltransferase (HPAT). The identification of FIN, predicted to modify a number of cell wall proteins and signaling peptides including CLV3p, adds an interesting new layer of complexity to the evolving model of meristem maintenance. Additional work is necessary to further functionally characterize the individual roles of *FAB* and *FIN* in meristem maintenance, to determine the functional links between them, and to place them in a larger developmental context.

CLAVATA signaling: conserved and novel features

Emerging examples have revealed variations on the canonical CLV-WUS signaling pathway and stress that there is much to be learned about meristem maintenance, even regarding highly conserved and well-studied components. FAB/CLV1 is a LRR-receptor kinase that is functionally conserved in plants as demonstrated in several organisms including *Arabidopsis* (Clark et al. 1997), maize (Bommert et al. 2005), rice (Suzaki et al. 2004), and tomato (this study). Work in these systems has found that CLV1 acts to regulate stemness in different meristem types, including the vegetative SAM, the floral meristem (FM) and in the root apical meristem (RAM), but subtle differences have been observed depending on the system and

meristem type (Clark et al. 1997; Suzaki et al. 2004; Bommert et al. 2005; Stahl et al. 2013). For instance, in contrast to *Arabidopsis*, where *CLV1* expression is restricted to the L3 layer in the meristem, *FON1*, the rice ortholog, is expressed diffusely throughout the meristem and also in developing floral organs (Suzaki et al. 2004). Also, CLV1 can act as a homomer, but it also acts in heteromeric complexes with other proteins, such as the ARABIDOPSIS CRINKLY4 (ACR4) receptor kinase, which functions in the *Arabidopsis* root (Stahl et al. 2013). Moreover, work across systems has found that these receptor complexes can perceive different peptide ligands, and there is variation in the downstream targets that CLV signaling acts upon. For instance, in the *Arabidopsis* SAM, CLV1 perceives CLV3p and signaling acts to restrict *WUS* expression, and WUS protein then feeds back to activate expression of *CLV3* to complete the negative feedback loop (Clark et al. 1997; Brand et al. 2000; Schoof et al. 2000; Ogawa et al. 2008; Nimchuk et al. 2011). However, in the root, CLV1 and ACR4 perceive CLV40p to restrict expression of *WOX5,* a *WUSCHEL*-related homeobox (Stahl et al. 2009; Stahl et al. 2013). Confirmation that FAB perceives CLV3p in tomato SAMs is one experiment that is important to complete. Furthermore, molecular work thus far in tomato has focused on characterization of *FAB* in vegetative shoot apical meristems (SAMs). It will also be interesting to determine if *FAB* has a role in different tomato meristem types, for instance, as in *Arabidopsis*, *FAB* may act in the root. It is also possible that there are subtle differences in regulation in different meristem types that mature from the SAM; for instance, comparing how stemness is maintained in the primary shoot meristem (PSM) versus the sympodial shoot meristem (SYM) and within the inflorescence meristems (SIMs) or floral meristem (FM).

Another exciting area for future work will be to further investigate the connection between *FIN* family members and related glycosyltransferases in CLV-WUS signaling. *fab fin* double mutants showed a more extreme phenotype than either of the single mutants, suggesting that *FIN* may also function in a parallel pathway with overlapping function with the CLV pathway. Tellingly, recent work in *Arabidopsis* demonstrated that the FIN homologs are predicted

to attach arabinose sugar moieties to hydroxyprolinated residues on various hydroxyproline rich glycoproteins (HRGPs) and also signaling peptides, including members of the CLE peptide family (Ogawa-Ohnishi et al. 2013). CLE peptides have varied roles in plant development and continue to be an exciting area of investigation (Jun et al. 2008; Yamada and Sawa 2013). I found 16 predicted CLEs in tomato (Figure 4.13), the functions of which are unknown. My expression profiling found that tomato *SlCLE7/SlCLV3* and *SlCLE6* are highly overexpressed in *fin* (Figure 4.11), and I hypothesize that *fin* is unable to produce properly modified CLE peptides such as SlCLV3p, thus breaking the canonical CLV-WUS feedback loop (Figure 4.17). Therefore, it is important to investigate whether *FIN* is directly responsible for modifying SlCLV3p, and if lack of this modification explains the fasciation and inflorescence branching observed in *fin*. Several experiments are underway to test this hypothesis (Chapter 4 Addendum).

To further elucidate the role of *FIN,* work in other systems will also be informative. *FIN* is a member of a small, highly conserved family that has just begun to be explored. Interestingly, while the *FIN* gene family is highly conserved across plants, very different phenotypes have been observed in various models. For instance, mutants in the *Medicago FIN* homolog called *ROOT DETERMINED NODULATION 1* (*RDN1*), do not have SAM or inflorescence defects, rather, they present a hyper-nodulation phenotype (Schnabel et al. 2011). Moreover, work in our lab and by the Matsubayashi group in Japan, has found that *Arabidopsis fin-like (finl/hpat)* single mutants do not have obvious phenotypes, but when trying to produce *hpat1 hpat3* double mutant combinations, a male transmission defect was encountered, indicating *finl/hpat* members are necessary for pollen development (Ogawa-Ohnishi et al. 2013; CAM, personal communication). Interestingly, *hpat1 hpat3* pollen is viable and able to elongate down the transmitting tract, but fails to properly target ovules (CAM, personal communication). Finally, in moss, which does not have pollen or a multicellular meristem, knockouts of one of the two *FIN* homologs caused larger colonies to form with increased caulonema growth (elongated cells that grow outwards to colonize the surrounding substrate; CAM, personal communication). These findings raise the

question: what is the common theme and what is the ancestral function of *FIN*? One possibility, particularly considering the phenotypes in *Arabidopsis* pollen and moss caulonemal cells, is that *FIN* plays a role in directionality of growth by perception of an extracellular signal. Are there any other commonalities between the systems that give us any clues? One interesting observation, considering the results in *Arabidopsis*, is that CLV3 is expressed in both the meristem and in pollen in both tomato (not published) and *Arabidopsis* (Schmid et al. 2005). Two tomato *FIN* genes are also specifically enriched in the pollen (Figure 4.16). Perhaps there is a conserved role for FIN in CLE perception and/or the perception of additional signaling molecules across all systems and this will be very interesting to pursue.

Fasciation and branching: beyond CLV-WUS signaling

Determining factors that regulate inflorescence complexity remains a challenge in plant biology and has direct implications for improving crop yield. This work provided the first functional evidence that the CLV-WUS meristem maintenance pathway is conserved in tomato and demonstrated a link between increased meristem size and inflorescence branching. However, many other factors, such as maturation and determinacy of meristems, meristem organization, phytohormone signaling, and cellular mechanics contribute to inflorescence branching as discussed below (reviewed in Besnard et al. 2011; Tanaka et al. 2013; Kyozuka et al. 2014; Park et al. 2014; Zadnikova and Simon 2014).

As meristems mature, they must strike a balance between maintaining themselves and outgrowth of lateral organs. Aboveground growth first arises from a vegetative shoot apical meristem that produces the shoot and leaves. Upon receiving intrinsic and environmental maturation cues, the meristem undergoes a transition to a reproductive state in which flowers are formed. Recent work, particularly in the Solanaceae (reviewed in Park et al. 2014) and the grasses (reviewed in Kyozuka et al. 2014), has demonstrated that the timing and duration of phase

changes during the meristem maturation process, as well as the determinacy of the meristem at each maturation state, are important for controlling inflorescence architecture types.

As described in previous sections, tomato typically produces a multi-flowered inflorescence with ~6-8 flowers arranged on a single linear truss. However, delays in meristem maturation, either in the transition meristem (TM) or sympodial inflorescence meristem (SIM) stages affords an opportunity for overproliferation of lateral meristems, and hence, branching within the inflorescence (Park et al. 2012). Meristems in the TM and SIM maturation states are still indeterminate, meaning that they maintain a level of stemness that allows them to produce lateral meristems. This is in contrast to the determinate nature of the floral meristem (FM), the final meristem maturation state in the inflorescence, which can only produce floral organs before terminating. Similarly, grass inflorescences, such as the rice panicle described here, are also multi-flowered and typically branched structures. However, inflorescences development in the grasses is somewhat different than in the eudicots (reviewed in Kyozuka et al. 2014). Upon the transition to flowering, the SAM transitions into an inflorescence meristem (IM), which is still indeterminate in nature. Unlike in other inflorescence types, however, the IM does not directly give rise to determinate FMs, rather, it allows for the formation of lateral indeterminate meristems called branch meristems (BMs), which form primary branches. The indeterminacy of the BMs allows them to generate additional lateral BMs to give rise to secondary and higher-order branches. Ultimately the BMs undergo a phase change to determinate spikelet meristems (SMs), which then form one or more flowers. Importantly, similar to the delay in the "meristem maturation clock" that underlies much of the branching variation observed in tomato (Park et al. 2012), prolonged BM identity, and hence an extended indeterminate state, allows for more branching within the rice inflorescence.

Thus, determining the factors that regulate the maturation, and likewise, the determinacy of meristems throughout the transition to flowering, is essential for understanding the formation of diverse inflorescence architectures. A few specific factors have been identified that help

coordinate this maturation process, although the functions sometimes differ between eudicots and the grasses. In tomato, a majority of inflorescence branching variation traces back to mutations in a single gene, *COMPOUND INFLORESCENCE* (*S*), which encodes a WUSCHEL homeobox transcription factor (Lippman et al. 2008). Interrogation of transcriptome profiles at defined meristem stages during the transition to flowering demonstrated that the extreme branching found in the *s* mutants is due to a delay in meristem maturation (Park et al. 2012). A prolonged state of indeterminacy, in both the TM and SIM stages, allows for the formation of ectopic lateral meristem and thus, more branching. Interestingly, loss of *s* expression causes a delay in expression of *ANANTHA* (*AN*), the tomato ortholog of *UNUSUAL FLORAL ORGANS* (*UFO*). In the eudicots like tomato, UFO/AN and LFY/FA form a floral activation complex, and thus, the timing of expression of these genes is important for promoting the transition to a determinate floral meristem (Molinero-Rosales et al. 1999; Chae et al. 2008; Lippman et al. 2008). However this function of the UFO and LFY homologs are not explicitly conserved in the grasses (reviewed in Kyozuka et al. 2014). Instead, defects in the rice homologs of *UFO/AN* and *LFY/FA* lead to an opposite phenotype, or premature SM identity and smaller inflorescences with less branching, whereas gain-of-function of *UFO* leads to more branching. In rice, the function of these genes appears to be to promote cell proliferation as opposed to floral activation. Interestingly, in several rice inflorescence mutants, increased cell proliferation correlates with increased branching (reviewed in Kyozuka et al. 2014). One possibility is that the cell proliferation allows for the meristem to persist in a state of indeterminacy longer leading to the formation of extra BMs, and therefore, more highly branched inflorescences. This is in contrast to rice mutants in the CLV-WUS meristem maintenance pathway, which have increased floral organ numbers, but do not present major defects in inflorescence architecture (Suzaki et al. 2004; Suzaki et al. 2009). Similarly, mutations in the maize CLV homologs lead to fasciated meristems, and hence, increased spikelet pair meristem (SPM) proliferation, resulting in ears with increased kernel row number and tassels with increased spikelet density (Bommert et al. 2005; Bommert et al. 2013a;

Bommert et al. 2013b). However, higher order branching due to extra BM proliferation is not seen in these mutants. Thus, while both meristem maintenance and meristem maturation contribute to inflorescence development, meristem determinacy appears to be the major factor regulating inflorescence branching in the grasses.

Other factors controlling meristem phase change are conserved across the eudicot-monocot divide. For instance, a member of the small family of *Arabidospsis* LSH and *Oryza* G1 (ALOG) family was found to modulate meristem phase change in both tomato and rice (reviewed in Kyozuka et al. 2014; Park et al. 2014). In tomato, loss-of-function mutations of the ALOG family member *TMF* lead to precocious activation of *AN*, resulting in the formation of a single flowered primary inflorescence (MacAlister et al. 2012). Similarly, *TAWAWA1* (*TAW1*), the rice homolog to *TMF*, is functionally conserved in rice (Yoshida et al. 2013). Loss of *TAW1* results in small inflorescences with fewer branches, likely due to precocious transition to SM identity. On the other hand, a gain-of-function mutant of *TAW1* delays SM specification, and therefore, the meristem stays in a prolonged state of indeterminacy, allowing for more lateral BMs to form, and resulting in increased branching.

In addition to meristem maturation and determinacy, proper organization within the meristem and developing organs, and signaling between them, is critical for maintaining meristem size and organ development. Each region – the meristem, boundary, and organ primordia – has distinct gene expression patterns and transcriptional regulation, mechanical properties, and hormone signaling, all of which differentiate the cell identity of each region and allows for communication between them (reviewed in Besnard et al. 2011; Zadnikova and Simon 2014). A breakdown in these factors can lead to meristem and/or organ defects (see below). Although boundaries are necessary to keep these distinct cell populations separate, signaling between the organ primordia and the meristem, via the boundary region, is essential for both meristem maintenance and organ development.

Meristem specific genes have been discussed extensively in previous sections. Briefly, members of the KNOX family such as SHOOT MERISTEMLESS (STM) along with CUP-SHAPED COTYLEDON (CUC) transcription factors promote meristem identity. In parallel, the CLV-WUS feedback loop acts specifically within the meristem to regulate shoot apical meristem size; loss of *CLV* genes result in fasciated meristems, whereas loss of *WUS* leads to a reduction in meristem size (Clark et al. 1997; Brand et al. 2000; Schoof et al. 2000). In the boundary and organ primordia, as cells commit to becoming an organ, different factors are at play to promote the differentiation of cells into the developing organ and promote organ polarity. One key change in the boundary and organ primordia is that the KNOX genes, which normally promote meristem identity, must be repressed. These genes are maintained in a repressive state in the boundary region and developing organ primordia by several factors including the organ-specific transcription factors ASYMETRIC LEAVES 1 and 2 (AS1/AS2) and polycomb repressive complexes (Ikezaki et al. 2010; Lodha et al. 2013). Finally, signaling from the developing organ primordia back to the meristem is also integral to restricting meristem organization and size. For instance, the petunia gene *HAIRY MERISTEM* (*HAM*), which encodes a GRAS transcription factor, signals non-cell autonomously from the organ primordia back on the meristem to promote meristem maintenance (Stuurman et al. 2002). Moreover, in *Arabidopsis*, abaxial-specific expression of YABBY genes in the organ primordia promotes the expression of a second GRAS transcription factor in the boundary called *LATERAL SUPPRESSOR* (*LAS*) (Goldshmidt et al. 2008). While the mechanism is not clear, LAS relays this signal to the meristem to promote proper meristem organization. Loss of *YABBY* expression leads to aberrant phyllotaxis and disorganization of the SAM (Goldshmidt et al. 2008; Sarojam et al. 2010). Interestingly, defects in a YABBY transcription factor, *fasciated* (*fas*), was the only fasciated tomato mutant characterized before the work presented in this thesis. While *FAS* is not connected to CLV-WUS signaling in the regulation of meristem maintenance *per se*, it does have an effect on meristem

organization, and as the name suggests, loss of *FAS* leads to fasciation of the flowers and fruits, and mild inflorescence branching (Cong et al. 2008).

In addition to the aforementioned transcription factors and signaling pathways, interplay between cell wall mechanics and phytohormone distribution within the meristem are also crucial for controlling meristem size via restriction of organ initiation and outgrowth (reviewed in Besnard et al. 2011; Zadnikova and Simon 2014). Notably, regulation of cellular mechanics such as cell wall stiffness, is essential for maintaining proper control of organ positioning, initiation, and outgrowth. For instance, microtubules are positioned parallel to axes of greatest stress around the meristem and direct the insertion of cell wall modifying enzymes such as the cellulose synthase (CESA) complex (Gutierrez et al. 2009). Cellulose deposition, and the direction of cellulose microfibril formation, is critical to plant cell wall structure, integrity, and anisotropy. Tellingly, mutations in *Arabidopsis* CESA INTERACTIVE PROTEIN 1, which helps anchor CESA to the microtubules, results in enlarged and twisted stems with altered phyllotaxy (Landrein et al. 2013). Concomitantly, PINFORMED (PIN) auxin efflux carriers orient themselves anticlinal to points of organ initiation, correlating with microtubule, and thus microfibril positioning (Heisler et al. 2010). This positioning is especially pronounced at the boundary between the meristem and developing organ primordia, where PINs are responsible for pumping auxin into the meristem and to maxima at the points of organ outgrowth, while boundary regions become devoid of auxin (reviewed in Zadnikova and Simon 2014). Importantly, organ initiation and phyllotaxy is likely regulated, in part, by organ primordia sensing the auxin levels of its neighbors (Jonsson et al. 2006; Smith et al. 2006). Meanwhile, auxin regulates the expression of cell wall modifying enzymes, thus creating a feedback loop that interconnects changes in cell wall composition and auxin positioning, both of which impact organ outgrowth. Recently, additional examples have emerged connecting cell wall properties with organ outgrowth. For instance, work in *Arabidopsis* has shown that pectin methyl-esterification is important for organ outgrowth and phyllotaxy, and arabinosylation of cell wall proteins is

important for suppressing lateral root growth (Peaucelle et al. 2008; Peaucelle et al. 2011; Roycewicz and Malamy 2014).

I hypothesized that a breakdown in CLV-WUS signaling may drive the fasciation and branching in *fin* due to a loss of modification of CLV3p. However, consistent with the aforementioned evidence for a connection between cell wall structure and the control of organ positioning and outgrowth, *fin* fasciation and branching may also be, at least partly, a consequence of a weakened cell wall. Specifically, a reduction in arabinosylation of cell wall strengthening components such as extensins in *fin,* could lead to reduced rigidity of the cell wall, and hence, formation of ectopic leaves. These ectopic organs would then disrupt signaling between the organ primordia and the meristem leading to further disorganization and enlargement of the meristem. In turn, meristem enlargement would allow for formation of ectopic meristems in the leaf axils, and hence, branching. Of course, these two mechanisms are not necessarily separate and may also be intertwined – losing arabinosylation of CLV3p and loss of cell wall integrity may have compounding effects on meristem size. An alternative hypothesis is that weakening of the cell wall leads to increased cell proliferation in the *fin* meristems and allows the meristem to persist in an indeterminate state longer during the transition to flowering. More time spent in the transition to flowering would allow for overproliferation of IMs and thus, increased branching. While loss of cell wall rigidity may feedback to upregulate the CLV-WUS loop, transcriptome profiling uncovered a dramatic increase in *SlCLV3* expression in *fin* that does not correlate with the increase in meristem size (Figures 4.11, 4.13, Table 4.2). Therefore, currently my results support the hypothesis for a direct role of CLV3p glycosylation in the regulation of meristem size. An exciting area for further work will be to investigate the role of glycosylation in diverse developmental contexts including meristem maintenance, from both the perspective of CLV-WUS signaling and from the perspective of cell wall mechanics.

Further impetus to investigate the role of glycosylation in meristem maintenance has come from characterization of another fasciated tomato mutant in the lab called *fasciated and*

necrotic (*fan*), which has similar, albeit milder, fasciation and branching than *fin*. Interestingly, *fan* mutants also senescence prematurely (ST and ZBL, unpublished). *fan* was recently found to be defective in the homolog of *Arabidopsis* XEG113, another glycosyltransferase that is predicted to catalyze the addition of the third arabinose moiety to arabinose chains (Gille et al. 2009). Similar to the *fin*/*hpat* situation, *Arabidopsis xeg113* mutants do not have shoot apical meristem defects. Rather, the most evident defects are hypocotyl elongation and shortened lateral roots, again pointing towards species-specific differences for *FAN/XEG113* homologs (Gille et al. 2009; Roycewicz and Malamy 2014). Yet, importantly, the tomato *fan* phenotype provides additional evidence that glycosylation is essential for tomato meristem maintenance. Even more, the milder fasciation phenotype in *fan* compared to *fin* suggests sensitivity to the length of the arabinose chain. This has been proposed previously in *Arabidopsis*, although the importance of glycosylation of CLV3p remains controversial (Song et al. 2012; Shinohara and Matsubayashi 2013). As FIN is likely responsible for adding the first arabinose group to hydroxyprolinated proteins/peptides and FAN is predicted to catalyze the addition of the third sugar to the elongating arabinose chain, we might expect that *fin* may be epistatic over *fan*. Puzzlingly, however, the *fin fan* double mutant is more extreme than either single mutant (ZBL and ST, personal communication). In fact, young *fin fan* double mutants are indistinguishable from *fab fin* double mutants – the primary SAM becomes extremely enlarged and fails to produce mature flowers (Figure 4.6, *fab fin* double; *fin fan*, not shown). As in the case of the *fab fin* double mutants, it is possible that FIN and FAN have both overlapping and separate targets. For instance, FAN may act on some proteins/signaling peptides that are not modified by FIN *per se*, but perhaps by one of the FIN homologs. Further work is necessary to tease apart the roles of FIN and FAN glycosylation in meristem maintenance, both of which may act in concert with FAB/CLV1 in CLV-WUS signaling, and may also have separate parallel roles in the regulation of meristem size and inflorescence structure.

The discovery that two glycosyltransferases act in the regulation of meristem size highlights that there is still much to learn about meristem maintenance in tomato. One approach to uncover additional factors that participate in the regulation of meristem size and branching is to perform an enhancer-suppressor screen, which I initiated in the spring of 2013 for *fab*. I treated *fab* seeds with EMS, M1 seed from >1000 individuals was collected in the summer field, and screening in the M2 generation can begin next season. We hope that this will uncover novel factors to further our understanding of how inflorescence architecture is determined.

Gene dosage and fine-tuning of dosage to manipulate plant architecture

Work presented in this dissertation also provides interesting evidence that fine-tuning gene dosage, particularly for conserved factors at key developmental time points such as the transition to flowering, can have a profound effect on plant architecture and reproductive success. As described previously, recent work demonstrated that single-gene overdominance involving heterozygosity for the flowering hormone florigen resulted in increased yield in tomato (Krieger et al. 2010). Importantly, through transcriptional profiling and detailed phenotyping we demonstrate that increased yield is a direct result of a sensitivity to *sft* dosage (Jiang et al. 2013; see Appendix B). Different alleles of sft were not tested explicitly in this study. However, in the recapitulation of this dosage effect using micro-RNAs, we were able to uncover transgenic plants with varying strengths of *SFT* knockdown and this translated to different degrees of *sp* suppression phenotypes. This suggested that fine-tuning of SFT dosage is possible. I extended the characterization of florigen heterozygosity to *Arabidopsis* to determine if this effect was conserved across diverse species. Interestingly, I found that while florigen dosage is sensed, as reflected in plant stature intermediate to the parental lines, the overall plant yield is not increased as in tomato (Figures 2.5 and 2.6). This difference traces to differences in the growth habit of the plants (Figure 1.3); while *Arabidopsis* only undergoes one transition to flowering, tomato

undergoes multiple transitions and therefore, florigen dosage can be sensed at each transition allowing for a few additional inflorescences to form on each shoot (Jiang et al. 2013; see Appendix B). This highlights that genetic context is key. However, this does not mean that the results are special to tomato, but it does stress that species-specific tuning will be necessary depending on many factors including growth habit. Many questions remain such as: Is the yield heterosis a special case between *SFT* and *SP*? Can we achieve the same effects by manipulating other components of flowering pathways? How much can yield be boosted by manipulating these pathways?

It may seem logical that manipulation of flowering time would be a target for increasing flowering and crop production. Indeed, flowering time has been selected for during the domestication of many crops, and *SFT* and *SP* orthologs have been identified as some of the key contributing factors (Kwak et al. 2008; Pin et al. 2010; Tian et al. 2010; Navarro et al. 2011; Repinski et al. 2012). However, most domestication in tomato has selected for fruit size and not growth habit or flowering time (Tanksley 2004; Peralta and Spooner 2005). This means that there may be a lack of allelic variation in genes controlling flowering time in tomato, but, argues that this pathway could be a key target for further yield improvement.

Further evidence that fine-tuning of dosage in flowering pathways could alter plant architecture came from characterizing an interesting moderately branched inflorescence variant called *bifurcating inflorescence* (*bif*). QTL mapping revealed a single major QTL that overlaps an interval containing *COMPOUND INFLORESCENCE* (*S*). *s* mutants have highly branched inflorescences due to a delay of the "meristem maturation clock" (Park et al. 2012); *s* mutants stall at two phases of meristem development allowing for overproliferation of meristems and hence branching. I described genetic and molecular evidence (Figure 3.3), indicating that transcriptional changes in *bif-S* mimics the expression profile of the weakly branched *S. peruvianum* (Park et al. 2012), which may explain the mild branching in *bif* variants. In parallel, work characterizing *s*/+ heterozygotes also found mild branching indicating that similar to the

case of *sft*/+ heterozygosity dosage sensitivity during the transition to flowering might be key to this effect. Combined, the above results argue that there is, in fact, much room for "fine-tuning" key pathways such as flowering for further manipulating plant architecture and improving crop yield. However, it also highlights that the available alleles in the natural germplasm are limited and that introduction of artificial alleles may be necessary to further fine-tune these networks for agricultural gain.

Advancing crop improvement with bioengineering

In breeding for improved agricultural lines there is a fine balance between improving desired traits such as increased yield without the expense of another such as fruit sugar content. Recent work, including the work presented in this dissertation, suggest that fine-tuning of gene expression could push these limits as described above (Jiang et al. 2013; see Appendix B). Natural variation and mutagenesis screens have allowed for the identification of variants that can be useful for breeding, however, variants are limited and time-consuming to find and introgress into elite germplasm. Therefore, to rapidly advance breeding programs, bioengineering is necessary.

Recent advancements in genome editing technology will advance both basic plant biology research and applied bioengineering in crops (reviewed in Belhaj et al. 2013). As most plants do not have natural mechanisms for homologous recombination, classically, the introduction of transgenes has been by random insertion often introducing multiple insertions in random locations. This leads to variation in expression of the transgene from both the position of insertion and copy number. However, several methods have recently emerged that allow for targeting specific loci. These include zinc finger nucleases (ZFNs), TAL effector nucleases (TALENs) and more recently, CRISPR (clustered regularly interspaced short palindromic repeats)/Cas (CRISPR-associated) type II prokaryotic adaptive immune system (CRISPR/Cas

system). Of these systems, the CRISPR/Cas system is gaining the most popularity, because unlike the other systems, it does not require the engineering of specific DNA binding proteins, rather, it is guided to the target site by a small engineered RNA. This technology represents a potential game-changer for plant genome editing. Indeed, recent reports demonstrate that the CRISPR/Cas system is highly effective at targeting specific loci via protoplast and agroinfiltration systems in several plant species, including Arabidopsis, rice, and sorghum (Jiang et al. 2013).

Importantly, CRISPRs have the potential to enable a range of genome manipulations including single gene knockouts, deletions of chromosomal segments, gene insertions and gene replacements (reviewed in Belhaj et al. 2013). From a basic biology perspective, this will speed up research in a variety of ways, including reverse genetics applications for targeted knockouts of genes of interests. From an applied perspective, this technology has the potential to advance the creation of new breeding lines in a variety of ways. For instance, targeted gene insertions may allow rapid stacking of transgenes. Moreover, as discussed above, fine-tuning of gene expression might be achieved with specific alleles and this technology would allow for replacement of a gene at its endogenous locus with an edited version. There has been a recent explosion of proof-of-principle experiments demonstrating that this system works in plants. It will be interesting to see how consistent the results are and how powerful the system will prove to be for crop improvement.

Concluding Remarks

Multiple pathways converge to regulate the development of diverse plant architecture types. My work has highlighted important mechanisms including, 1) a role for meristem size/stem cell maintenance in the regulation of inflorescence branching, and, 2) a role for dosage-sensitivity during the transition to flowering, which impacts both plant growth habit and inflorescence architecture. In both cases, genetic context is key; highly conserved factors can behave very

differently in varying contexts. Yet, these mechanisms and conserved factors provide rich opportunity for furthering our knowledge of plant development and both may be harnessed and manipulated for improving crop yield.

References

Belhaj K, Chaparro-Garcia A, Kamoun S, Nekrasov V. 2013. Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. *Plant Methods* 9: 39.

Besnard F, Vernoux T, Hamant O. 2011. Organogenesis from stem cells in planta: multiple feedback loops integrating molecular and mechanical signals. *Cell Mol Life Sci* 68: 2885-2906.

Bommert P, Je BI, Goldshmidt A, Jackson D. 2013a. The maize Galpha gene COMPACT PLANT2 functions in CLAVATA signalling to control shoot meristem size. *Nature*.

Bommert P, Lunde C, Nardmann J, Vollbrecht E, Running M, Jackson D, Hake S, Werr W. 2005. thick tassel dwarf1 encodes a putative maize ortholog of the Arabidopsis CLAVATA1 leucine-rich repeat receptor-like kinase. *Development* 132: 1235-1245.

Bommert P, Nagasawa NS, Jackson D. 2013b. Quantitative variation in maize kernel row number is controlled by the FASCIATED EAR2 locus. *Nature genetics* 45: 334-337.

Brand U, Fletcher JC, Hobe M, Meyerowitz EM, Simon R. 2000. Dependence of Stem Cell Fate in Arabidopsis on a Feedback Loop Regulated by CLV3 Activity. *Science* 289: 617-619.

Chae E, Tan QK, Hill TA, Irish VF. 2008. An Arabidopsis F-box protein acts as a transcriptional co-factor to regulate floral development. *Development* 135: 1235-1245.

Clark SE, Williams RW, Meyerowitz EM. 1997. The CLAVATA1 gene encodes a putative receptor kinase that controls shoot and floral meristem size in Arabidopsis. *Cell* 89: 575-585.

Cong B, Barrero LS, Tanksley SD. 2008. Regulatory change in YABBY-like transcription factor led to evolution of extreme fruit size during tomato domestication. *Nature genetics* 40: 800-804.

Gille S, Hansel U, Ziemann M, Pauly M. 2009. Identification of plant cell wall mutants by means of a forward chemical genetic approach using hydrolases. *Proceedings of the National Academy of Sciences of the United States of America* 106: 14699-14704.

Goldshmidt A, Alvarez JP, Bowman JL, Eshed Y. 2008. Signals derived from YABBY gene activities in organ primordia regulate growth and partitioning of Arabidopsis shoot apical meristems. *The Plant cell* 20: 1217-1230.

Gutierrez R, Lindeboom JJ, Paredez AR, Emons AM, Ehrhardt DW. 2009. Arabidopsis cortical microtubules position cellulose synthase delivery to the plasma membrane and interact with cellulose synthase trafficking compartments. *Nat Cell Biol* 11: 797-806.

Heisler MG, Hamant O, Krupinski P, Uyttewaal M, Ohno C, Jonsson H, Traas J, Meyerowitz EM. 2010. Alignment between PIN1 polarity and microtubule orientation in the shoot apical meristem reveals a tight coupling between morphogenesis and auxin transport. *PLoS biology* 8: e1000516.

Ikezaki M, Kojima M, Sakakibara H, Kojima S, Ueno Y, Machida C, Machida Y. 2010. Genetic networks regulated by ASYMMETRIC LEAVES1 (AS1) and AS2 in leaf development in Arabidopsis thaliana: KNOX genes control five morphological events. *Plant J* 61: 70-82.

Jiang K, Liberatore KL, Park SJ, Alvarez JP, Lippman ZB. 2013; see Appendix B. Tomato Yield Heterosis is Triggered by a Dosage Sensitivity of the Florigen Pathway that Fine-Tunes Shoot Architecture. *PLoS Genet* 9: e1004043.

Jiang W, Zhou H, Bi H, Fromm M, Yang B, Weeks DP. 2013. Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. *Nucleic acids research* 41: e188.

Jonsson H, Heisler MG, Shapiro BE, Meyerowitz EM, Mjolsness E. 2006. An auxin-driven polarized transport model for phyllotaxis. *Proceedings of the National Academy of Sciences of the United States of America* 103: 1633-1638.

Jun JH, Fiume E, Fletcher JC. 2008. The CLE family of plant polypeptide signaling molecules. *Cell Mol Life Sci* 65: 743-755.

Krieger U, Lippman ZB, Zamir D. 2010. The flowering gene SINGLE FLOWER TRUSS drives heterosis for yield in tomato. *Nat Genet* 42: 459-463.

Kwak M, Velasco D, Gepts P. 2008. Mapping homologous sequences for determinacy and photoperiod sensitivity in common bean (Phaseolus vulgaris). *J Hered* 99: 283-291.

Kyozuka J, Tokunaga H, Yoshida A. 2014. Control of grass inflorescence form by the fine-tuning of meristem phase change. *Curr Opin Plant Biol* 17C: 110-115.

Landrein B, Lathe R, Bringmann M, Vouillot C, Ivakov A, Boudaoud A, Persson S, Hamant O. 2013. Impaired cellulose synthase guidance leads to stem torsion and twists phyllotactic patterns in Arabidopsis. *Curr Biol* 23: 895-900.

Lippman ZB, Cohen O, Alvarez JP, Abu-Abied M, Pekker I, Paran I, Eshed Y, Zamir D. 2008. The making of a compound inflorescence in tomato and related nightshades. *PLoS biology* 6: e288.

Lodha M, Marco CF, Timmermans MC. 2013. The ASYMMETRIC LEAVES complex maintains repression of KNOX homeobox genes via direct recruitment of Polycomb-repressive complex2. *Genes Dev* 27: 596-601.

MacAlister CA, Park SJ, Jiang K, Marcel F, Bendahmane A, Izkovich Y, Eshed Y, Lippman ZB. 2012. Synchronization of the flowering transition by the tomato TERMINATING FLOWER gene. *Nature genetics* 44: 1393-1398.

Molinero-Rosales N, Jamilena M, Zurita S, Gomez P, Capel J, Lozano R. 1999. FALSIFLORA, the tomato orthologue of FLORICAULA and LEAFY, controls flowering time and floral meristem identity. *Plant J* 20: 685-693.

Navarro C, Abelenda JA, Cruz-Oro E, Cuellar CA, Tamaki S, Silva J, Shimamoto K, Prat S. 2011. Control of flowering and storage organ formation in potato by FLOWERING LOCUS T. *Nature* 478: 119-122.

Nimchuk ZL, Tarr PT, Ohno C, Qu X, Meyerowitz EM. 2011. Plant stem cell signaling involves ligand-dependent trafficking of the CLAVATA1 receptor kinase. *Curr Biol* 21: 345-352.

Ogawa M, Shinohara H, Sakagami Y, Matsubayashi Y. 2008. Arabidopsis CLV3 peptide directly binds CLV1 ectodomain. *Science* 319: 294.

Ogawa-Ohnishi M, Matsushita W, Matsubayashi Y. 2013. Identification of three hydroxyproline O-arabinosyltransferases in Arabidopsis thaliana. *Nat Chem Biol* 9: 726-730.

Park SJ, Eshed Y, Lippman ZB. 2014. Meristem maturation and inflorescence architecture-lessons from the Solanaceae. *Curr Opin Plant Biol* 17C: 70-77.

Park SJ, Jiang K, Schatz MC, Lippman ZB. 2012. Rate of meristem maturation determines inflorescence architecture in tomato. *Proceedings of the National Academy of Sciences of the United States of America* 109: 639-644.

Peaucelle A, Louvet R, Johansen JN, Hofte H, Laufs P, Pelloux J, Mouille G. 2008. Arabidopsis phyllotaxis is controlled by the methyl-esterification status of cell-wall pectins. *Curr Biol* 18: 1943-1948.

Peaucelle A, Louvet R, Johansen JN, Salsac F, Morin H, Fournet F, Belcram K, Gillet F, Hofte H, Laufs P et al. 2011. The transcription factor BELLRINGER modulates phyllotaxis by regulating the expression of a pectin methylesterase in Arabidopsis. *Development* 138: 4733-4741.

Peralta IE, Spooner DM. 2005. Morphological characterization and relationships of wild tomatoes (Solanum L. Sect. Lycopersicon). *Monogr Syst Bot, Missouri Bot Gard* 104: 227-257.

Pin PA, Benlloch R, Bonnet D, Wremerth-Weich E, Kraft T, Gielen JJ, Nilsson O. 2010. An antagonistic pair of FT homologs mediates the control of flowering time in sugar beet. *Science* 330: 1397-1400.

Repinski SL, Kwak M, Gepts P. 2012. The common bean growth habit gene PvTFL1y is a functional homolog of Arabidopsis TFL1. *Theor Appl Genet* 124: 1539-1547.

Roycewicz PS, Malamy JE. 2014. Cell wall properties play an important role in the emergence of lateral root primordia from the parent root. *J Exp Bot*.

Sarojam R, Sappl PG, Goldshmidt A, Efroni I, Floyd SK, Eshed Y, Bowman JL. 2010. Differentiating Arabidopsis shoots from leaves by combined YABBY activities. *The Plant Cell* 22: 2113-2130.

Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, Scholkopf B, Weigel D, Lohmann JU. 2005. A gene expression map of Arabidopsis thaliana development. *Nature genetics* 37: 501-506.

Schnabel EL, Kassaw TK, Smith LS, Marsh JF, Oldroyd GE, Long SR, Frugoli JA. 2011. The ROOT DETERMINED NODULATION1 gene regulates nodule number in roots of Medicago truncatula and defines a highly conserved, uncharacterized plant gene family. *Plant Physiol* 157: 328-340.

Schoof H, Lenhard M, Haecker A, Mayer KF, Jurgens G, Laux T. 2000. The stem cell population of Arabidopsis shoot meristems in maintained by a regulatory loop between the CLAVATA and WUSCHEL genes. *Cell* 100: 635-644.

Shinohara H, Matsubayashi Y. 2013. Chemical synthesis of Arabidopsis CLV3 glycopeptide reveals the impact of hydroxyproline arabinosylation on peptide conformation and activity. *Plant Cell Physiol* 54: 369-374.

Smith RS, Guyomarc'h S, Mandel T, Reinhardt D, Kuhlemeier C, Prusinkiewicz P. 2006. A plausible model of phyllotaxis. *Proceedings of the National Academy of Sciences of the United States of America* 103: 1301-1306.

Song XF, Yu DL, Xu TT, Ren SC, Guo P, Liu CM. 2012. Contributions of individual amino acid residues to the endogenous CLV3 function in shoot apical meristem maintenance in Arabidopsis. *Molecular plant* 5: 515-523.

Stahl Y, Grabowski S, Bleckmann A, Kuhnemuth R, Weidtkamp-Peters S, Pinto KG, Kirschner GK, Schmid JB, Wink RH, Hulsewede A et al. 2013. Moderation of Arabidopsis root stemness by CLAVATA1 and ARABIDOPSIS CRINKLY4 receptor kinase complexes. *Curr Biol* 23: 362-371.

Stahl Y, Wink RH, Ingram GC, Simon R. 2009. A signaling module controlling the stem cell niche in Arabidopsis root meristems. *Curr Biol* 19: 909-914.

Stuurman J, Jaggi F, Kuhlemeier C. 2002. Shoot meristem maintenance is controlled by a GRAS-gene mediated signal from differentiating cells. *Genes Dev* 16: 2213-2218.

Suzaki T, Ohneda M, Toriba T, Yoshida A, Hirano HY. 2009. FON2 SPARE1 redundantly regulates floral meristem maintenance with FLORAL ORGAN NUMBER2 in rice. *PLoS genetics* 5: e1000693.

Suzaki T, Sato M, Ashikari M, Miyoshi M, Nagato Y, Hirano HY. 2004. The gene FLORAL ORGAN NUMBER1 regulates floral meristem size in rice and encodes a leucine-rich repeat receptor kinase orthologous to Arabidopsis CLAVATA1. *Development* 131: 5649-5657.

Tanaka W, Pautler M, Jackson D, Hirano HY. 2013. Grass meristems II: inflorescence architecture, flower development and meristem fate. *Plant Cell Physiol* 54: 313-324.

Tanksley SD. 2004. The genetic, developmental, and molecular bases of fruit size and shape variation in tomato. *The Plant cell* 16 Suppl: S181-189.

Tian Z, Wang X, Lee R, Li Y, Specht JE, Nelson RL, McClean PE, Qiu L, Ma J. 2010. Artificial selection for determinate growth habit in soybean. *Proceedings of the National Academy of Sciences of the United States of America* 107: 8563-8568.

Yamada M, Sawa S. 2013. The roles of peptide hormones during plant root development. *Curr Opin Plant Biol* 16: 56-61.

Yoshida A, Sasao M, Yasuno N, Takagi K, Daimon Y, Chen R, Yamazaki R, Tokunaga H, Kitaguchi Y, Sato Y et al. 2013. TAWAWA1, a regulator of rice inflorescence architecture, functions through the suppression of meristem phase transition. *Proceedings of the National Academy of Sciences of the United States of America* 110: 767-772.

Zadnikova P, Simon R. 2014. How boundaries control plant development. *Curr Opin Plant Biol* 17C: 116-125.

Appendix A

8 Heterosis: The Case for Single-Gene Overdominance

Katie L. Liberatore[1,2*], Ke Jiang[1,2*], Dani Zamir[3], and Zachary B. Lippman[1,2]

[1] *Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, USA*
[2] *Watson School of Biological Sciences, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, USA*
[3] *The Robert H. Smith Institute of Plant Sciences and Genetics in Agriculture, The Hebrew University of Jerusalem, Rehovot, Israel*
[*] *Katie L. Liberatore and Ke Jiang contributed equally to this work.*

Introduction

For centuries, naturalists as well as plant and animal breeders have noted that prolonged inbreeding in normally outcrossing populations leads to progressive accumulation of inferior traits such as smaller, less vigorous, sickly, and often malformed offspring (Darwin, 1868; Charlesworth & Willis, 2009). The genetic and molecular basis of this "inbreeding depression" is still not completely understood (Crow, 2008). However, a widely supported hypothesis is that the accumulation of spontaneously formed deleterious recessive mutations is unmasked upon inbreeding, often culminating in maladaptive phenotypes (Charlesworth & Willis, 2009). The evolutionary implications of inbreeding depression are widespread, and this topic remains a focal point of population genetics research. Yet, it is the surprising and mysterious antithesis to inbreeding depression known as "hybrid vigor" that has captured the imaginations of breeders and scientists alike for more than a century. Hybrid vigor, the phenotypic superiority and improved fitness among progeny resulting from crossing genetically distinct parents, was first described by Charles Darwin and later refined by the maize geneticists George Shull and Edward East. Perhaps the most renowned demonstration of hybrid vigor is the mule—a stronger and more fit, albeit sterile, animal resulting from mating a male donkey with a female horse. Countless additional examples of both plant and animal hybrid vigor have been noted over the last century leading to the suggestion that increased heterozygosity in hybrid organisms provides a "magical" genetic and physiological advantage in growth and fecundity extending beyond the simple masking (i.e., complementation) of deleterious alleles. Thus, over the last century, and especially within the last decade, great efforts have been devoted to deciphering the genetic and molecular underpinnings of hybrid vigor in diverse organisms ranging from yeast to humans. In this chapter, we provide a general overview of hybrid vigor/heterosis from the perspective of both natural populations (hybrid vigor) and crops (heterosis), and we discuss how this field of study has evolved from the first classical genetic models put forth for maize. In particular, we explore current knowledge on examples of hybrid vigor implicating the most intriguing and controversial model

Polyploid and Hybrid Genomics, First Edition. Edited by Z. Jeffrey Chen and James A. Birchler.

known as "single-gene overdominance," which is based on heterozygosity and allelic interactions at a single gene.

Understanding Hybridization: Natural Phenomenon to Genetic Mystery

Hybrids are found throughout nature and in agriculture, and examples of hybridization span the gamut from whole-genome heterozygosity between distinct species to hybridization between nearly isogenic breeding lines that differ only in a small chromosomal segment (i.e., an introgression) harboring just a few dozen or even a single gene. Although hybridization between genetically distinct parents creates new allelic combinations genome-wide, drastic phenotypic changes in offspring are not always observed. Studies throughout the twentieth century involving crosses between different *Drosophila* species, however, revealed several examples where hybrid phenotypes extended beyond the bounds of parents for fitness traits such as viability (Dobzhansky, 1950), fertility (Fry et al., 1998), growth rate (Houle, 1989), and even cold tolerance (Jefferson et al., 1974). Such "transgressive variation" can be negative, in which progeny are less fit than their parents, or positive, in which progeny exceed the fitness of their progenitors (hybrid vigor/heterosis).

The mechanisms underlying these hybridization-induced transgressive phenotypes are being actively investigated at both the genetic and molecular levels, and studies have shown that several cases of reduced vigor and fitness trace back to simple negative epistatic interactions in hybrids. The most well-documented cases have been investigated for nearly a century in the context of speciation and hybrid incompatibility in *Drosophila* (Bateson, 1909; Muller, 1940, 1942; Wallace & Dobzhansky, 1962; Dobzhansky & Spassky, 1968). First Bateson and then Dobzhansky and Muller noted that crossing different *Drosophila* species often produced sterile or lethal progeny, and surprisingly, genetic analysis indicated that as few as two interacting loci were involved. This led to the formulation of the Bateson–Dobzhansky–Muller incompatibility (BDMI) model, which states that at least two loci, having evolved independently into new allelic forms either through sympatric (overlapping geographic distributions) or allopatric (nonoverlapping geographic distributions) speciation, show negative epistasis and reduced fitness upon "meeting" each other again in hybrids (Wallace & Dobzhansky, 1962; Dobzhansky & Spassky, 1968; Orr, 1996). Recently, the molecular identities of several BDMI genes have been discovered, and they represent a range of molecular functions from chromatin-binding factors (Brideau et al., 2006) to components of nuclear pore complexes (Tang & Presgraves, 2009). Suggestive that BDMI is widespread in nature and a driver of speciation is the rediscovery and molecular dissection of hybrid necrosis in plants (Bomblies et al., 2007). By intracrossing hundreds of *Arabidopsis thaliana* accessions in a large hybridization matrix scheme, it has been discovered that 2% of hybrids exhibit an autoimmune-like response at natural growth temperatures, which is triggered by epistatic interactions between two disease-resistance genes (Bomblies & Weigel, 2007; Bomblies et al., 2007). Additional examples of BDMI in *Arabidopsis* have been found, including embryonic lethality and stunted root growth, and interestingly these examples involve different classes of genes (Bikard et al., 2009). Thus, similar to *Drosophila*, the molecular dissection of BDMI in plants suggests that several pathways and networks involved in growth and fitness can drive negative epistasis. Importantly, BDMI interactions have also been found through environment-driven laboratory evolution in yeast, highlighting the spontaneous, rapid, and dramatic impacts that negative epistasis can have on populations and the process of speciation (Anderson et al., 2010). Strikingly, a recent example in *Arabidopsis* has shown that heterozygosity at just a single locus harboring tandemly repeated receptor-like kinase genes can

cause negative epistasis (Smith et al., 2011). It should be emphasized that for all these examples, genetic context and growth condition is key (Wallace & Dobzhansky, 1962; Dobzhansky & Spassky, 1968; Bomblies et al., 2007; Bikard et al., 2009; Anderson et al., 2010). For instance, hybrid necrosis in *Arabidopsis* is rescued at high temperatures, suggesting that BDMI is driven through interactions between genes and environmental selection pressures. With such dramatic phenotypic consequences resulting from negative epistatic interactions in hybrids involving just one or two genes, the obvious question has been whether the opposite side of the hybridization coin, hybrid vigor (also referred to as heterosis), might be based on similarly simple genetic mechanisms.

Hybrid Vigor versus Heterosis

Before moving further, we propose to make a distinction between the terms "hybrid vigor" and "heterosis." Hybrid vigor hereinafter is defined as natural hybridization and heterozygosity that contributes to the selection and adaptation of the most "fit" individuals, although not necessarily the largest or most prolific. Heterosis, on the other hand, is defined as representing specific cases of hybrid vigor following artificial selection and domestication (as in selective breeding programs for crops and livestock), in which the offspring are larger and more prolific (i.e., higher yielding) than elite inbred lines. Thus, to achieve heterosis, hybrid traits have been selected from an anthropocentric point of view and they exhibit "positive" transgressive phenotypes. Conversely, hybrid vigor in natural populations is based on fitness, and thus hybrids may not always show transgressive variation. Along the same lines, heterosis does not necessarily confer adaptation; for example, increased vegetative growth in hybrids can be advantageous and considered heterotic in vegetable cultivation, but may not translate to better fitness in the wild. This distinction is key since superior vigor over parents may increase productivities in hybrid crops or domesticated animals, but may be selected against in nature.

The role of hybrids in driving evolution and speciation in the absence of transgressive variation is well described. For example, hybrids between two closely related species may give rise to a population of fitter individuals that outperform their parents in certain environments to the extent that natural selection maintains this hybrid population by selecting against backcrosses to parental lines. Alternatively, geographical or ecological barriers to the parents can preserve the hybrid population. Such cases of hybrid vigor have been studied extensively in sunflowers, Louisiana Irises, and birds (Welch & Rieseberg, 2002; Hermansen et al., 2011; Taylor et al., 2011). For example, the sunflower homoploid hybrid *Helianthus paradoxus* shows higher salt tolerance, enabling this hybrid species to occupy a special niche in salt marshes that neither parental species can inhabit (Welch & Rieseberg, 2002). Importantly, although this unique physiological adaptation affords the hybrid with a selective advantage in this harsh environment, it is otherwise intermediate in phenotype relative to its parental species in nonselective conditions, and thus not necessarily more vigorous as in cases of crop heterosis (see below). This example brings about a critical point—one cannot ignore that in the wild, natural selection puts constant pressure on genomes to maintain fitness within a population, which to the human perspective may not represent the "best" phenotypes (Mather, 1955). The distinction between agriculturally "heterotic" traits and better adaptations in nature due to "hybrid vigor" must be noted because studies have involved a range of phenotypes spanning both natural populations and agricultural organisms. We highlight these distinctions also to foreshadow that both hybrid vigor and heterosis are based on multiple mechanisms, depending on organisms, trait, and growth condition, examples of which are discussed in detail below.

Inbreeding Depression and Heterosis in Breeding

For over 200 years, breeders and experimental biologists alike have documented the phenomenon of increased vigor in hybrid offspring. The earliest examples of hybrid vigor, as elegantly presented, although not precisely defined, by Charles Darwin in *The Variation of Animals and Plants Under Domestication*, were primarily from accounts of animal breeders (Darwin, 1868). Here, Darwin describes that a certain level of inbreeding is necessary to foster domestication and maintain "pure" breeding lines. However, he provides compelling evidence from diverse species of wild and domesticated animals such as cattle that "the infusion of fresh blood" is needed for maintaining healthy populations over time (Darwin, 1868). Moreover, intermating progressively more divergent breeding stocks often resulted in increased size, vigor, and fertility, and thus breeders recognized that a certain level of outcrossing was beneficial, although the biological basis was a mystery. Indeed, inbreeding depression has been classically detected in many organisms since Darwin's time by directly observing each generation of progeny following successive rounds of selfing. More recently, inbreeding depression has been studied by estimating the genetic load in populations of various organisms using molecular markers (Charlesworth & Willis, 2009). Given the evidence that deleterious recessive mutations of small or large effects are widespread in natural populations, it has been suggested that these mutations are the major causes of inbreeding depression (Charlesworth & Willis, 2009). The corollary of this hypothesis is that complementation of these same accumulated deleterious recessive mutations in hybrids manifests hybrid vigor (Charlesworth & Willis, 2009).

The first controlled experiments to understand the positive attributes of hybridization were performed in plants. Although Gregor Mendel and plant breeders of the early nineteenth century anecdotally noted the increased vigor of hybrid plants, Darwin completed the first set of extensive experimentation on the subject (Mendel, 1865; Darwin, 1876). Darwin observed increased vigor in hybrid progeny resulting from crosses of different parental lines across many genera of plants, although the specific species and breeding lines, as well as the environment, impacted the magnitude of vigor. Although the benefits of hybridization were evident to plant and animal breeders of the time, and a remarkable allusion to transgressive variation and hybrid vigor extends as far back as biblical times (Gen. 30:31–43), this phenomenon received little attention until its rediscovery by maize geneticists in the early 1900s, which set the foundation for deciphering the genetic basis of what is now formally known as "heterosis." While George Shull and Edward East both reported severely diminished vigor in maize inbred lines over several generations of inbreeding, Shull first reported recovery and dramatically improved vigor in hybrids (East, 1908; Shull, 1908). Over the following decade, the utilization of this transgressive variation became widespread in maize breeding, and the debate over two prominent theories for the genetic basis of heterosis, "dominance" and "overdominance," had begun (East, 1908; Shull, 1908; Bruce, 1910; Jones, 1917; Singleton, 1941).

Since the 1930s, utilization of hybridization has led to a significant increase in yield by up to 50%, depending on the crop (Singleton, 1941; Duvick, 1999, 2001, 2005; Davies, 2003). Although newer techniques such as transgenics are being implemented for crop improvement (e.g., to introduce resistance to diseases and pests, and herbicides for weed control), plant breeding through hybridization and heterosis still plays a major role in generating new elite high-yielding varieties (Davies, 2003). Importantly, it is believed that the upper limit of heterosis in plants has not been achieved and that an improved understanding of how heterotic effects are modified by genetic background and environmental conditions can enable the breaking of yield barriers (Castle, 1926; Crow, 1948; Birchler et al., 2001, 2005, 2007; Burke & Arnold, 2001; Gur & Zamir, 2004; Hochholdinger & Hoecker, 2007; Lippman & Zamir, 2007; Springer & Stupar, 2007b; Chen, 2010; Goff, 2010). Thus, understanding and harnessing the genetic and molecular basis of heterosis remains a focus in crop breeding.

Hypotheses on the Genetic Basis of Heterosis

The genetic basis of heterosis has eluded researchers for over a century despite extensive genetic and molecular experimentation. For instance, with new gene expression detection technologies, widespread studies in the last 5 years have focused on transcriptome analysis in hybrids compared to progenitor lines (Hochholdinger & Hoecker, 2007). What has become clear is that drawing meaningful conclusions from these molecular profiling experiments is challenging. Determining the proper stage for sampling, ensuring matched stages are taken for all genotypes, and differentiating statistically significant versus biologically relevant molecular changes from technical artifacts arising from the use of multiple profiling platforms (i.e., microarrays and sequencing) are just a few of the difficulties encountered (Guo et al., 2004, 2006; Sun et al., 2004; Auger et al., 2005; Vuylsteke et al., 2005; Huang et al., 2006; Stupar & Springer, 2006; Swanson-Wagner et al., 2006, 2009; Meyer et al., 2007; Song et al., 2007; Springer & Stupar, 2007a; Uzarowska et al., 2007; Hoecker et al., 2008; Li et al., 2009a, 2009b; Stokes et al., 2010). Although these genomic tools have been useful for providing insight into the molecular profiles of hybrids and the response of the transcriptome following hybridization, they have not yet been successful at resolving the basis for hybrid vigor or heterosis. The challenge remains going beyond simple correlations between gene expression changes in hybrids and heterotic phenotypes (Lippman & Zamir, 2007).

Of all approaches, classical and quantitative genetics have provided some of the greatest progress in deciphering heterosis and addressing two conventional models: dominance and overdominance (East, 1908, 1909; Shull, 1908; Birchler et al., 2003; Hochholdinger & Hoecker, 2007; Lippman & Zamir, 2007; Springer & Stupar, 2007b; Chen, 2010). Dominance complementation, which is the most widely accepted model (Lippman & Zamir, 2007; Gore et al., 2009), presumes that superior dominant alleles complement nonoverlapping deleterious recessive alleles at potentially hundreds or even thousands of loci across the genome in the F1 hybrid. Overdominance, on the other hand, posits that intralocus allelic interactions at just a single heterozygous gene can cause heterosis, or that several overdominant loci of small effect contribute cumulatively to heterosis. Many examples supporting the dominance hypothesis have been reported (Jones, 1917; Xiao et al., 1995; Peters et al., 2003). Most recently, with the introduction of cost-effective high-throughput sequencing, genotyping within and between populations has afforded the opportunity to evaluate the scope of allelic variation, and the extent of heterozygosity in a genome-wide context. For example, genomic diversity analysis of maize inbreds revealed an extraordinary level of variation within domesticated germplasm. For example, compared to the benchmark reference genome B73, the genome of any given maize line is missing approximately 5% gene space, in addition to variation in noncoding regions and repetitive elements (Gore et al., 2009; McMullen et al., 2009). This provides a tremendous amount of genetic material for complementation when the inbred genomes are united upon hybridization. A high number of gene presence/absence polymorphisms among maize inbred lines have also been discovered, providing even more indirect support for the dominance hypothesis (Springer et al., 2009; Lai et al., 2010).

But in the context of agriculture, the relative role and relevance of dominance complementation versus overdominance in explaining heterosis is far from settled. There are key distinctions between crop plants and wild populations with respect to the genetics of hybridization. In crop plants, in which most quantitative genetic analyses have been performed, inbred genetic backgrounds with improved performance can often be generated with ease by purging of deleterious mutations due to intensive artificial selection. In this respect, compared with wild populations, greater homozygosity and less deleterious mutations in crops reduce the likelihood that dominance complementation is the primary explanation for improved vigor in a hybrid. Consequently, evidence for overdominance, and

also a third model, epistasis, involving interactions among multiple genes and alleles (Hochholdinger & Hoecker, 2007; Lippman & Zamir, 2007; Springer & Stupar, 2007b; Birchler et al., 2010; Chen, 2010), is still being pursued in the context of domesticated organisms. Below, we explore current support for overdominance, drawing examples from across the plant and animal kingdoms with particular emphasis on single-locus heterosis in crops.

Overdominance and Quantitative Genetics

Unlike the dominance model, overdominance does not rely on elite alleles *per se*; rather, intralocus allelic interactions, regardless of allelic relationship, drive transgressive phenotypes. Overdominance has proven difficult to study because it requires an otherwise isogenic background, and epistatic interactions between loci can mask overdominant effects (Hochholdinger & Hoecker, 2007; Lippman & Zamir, 2007; Springer & Stupar, 2007b; Birchler et al., 2010; Chen, 2010). In addition, other mechanistic twists such as pseudo-overdominance, which describes cases of dominance that mimic overdominance due to linkage in repulsion of two or more deleterious recessive alleles, and epistasis, have been observed in a number of organisms and have further complicated the efforts to demonstrate convincing examples of overdominance (Stuber et al., 1992; Xiao et al., 1995; Yu et al., 1997; Li et al., 2001; Luo et al., 2001; Stupar & Springer, 2006; Ishikawa, 2009).

Although controversial, evidence for overdominance continues to be sought after in domesticated organisms, particularly crops, in which isogenic backgrounds and linkage information are readily available. As quantitative genetic approaches such as quantitative trait loci (QTL) mapping have become accessible in diverse systems due to high-throughput marker discovery and now genotyping-by-sequencing tools, greater efforts have been placed on breaking down the genetic basis of heterosis. Early on, interval mapping techniques were used in the model system *Arabidopsis*, and it revealed a major potential overdominant QTL associated with viability heterosis (Mitchell-Olds, 1995). Around the same time, QTL studies in crop species such as rice and maize provided support for all four major genetic models for heterosis: dominance, overdominance, pseudo-overdominance, and epistasis. Interestingly, one of these early QTL studies in maize found heterotic traits mainly associated with heterozygous genotypes, suggesting a predominant role of overdominance in maize heterosis (Stuber et al., 1992). However, this experiment failed to narrow down the QTL to smaller chromosome segments due to the lack of high-density genetic markers at the time, and one QTL was later shown to be based on two closely linked QTL that act in a dominance complementation manner (Graham et al., 1997)—a classic scenario of pseudo-overdominance. As demonstrated by these early studies, in many cases the lack of high-density markers makes overdominance and pseudo-overdominance indistinguishable. In addition, QTL studies in rice unveiled another mechanistic twist, epistasis, in which dozens or even hundreds of linked and unlinked loci in genomic space interact to cause synergistic heterotic effects (Li et al., 2001; Luo et al., 2001). Unfortunately, the precise mapping of heterotic loci putatively originating from epistatic interactions is notoriously difficult to resolve. Indeed, epistasis in segregating mapping populations can modify heterotic phenotypes and therefore confound QTL mapping and prevent the isolation of overdominant loci.

Nonetheless, the early QTL analyses in both model and crop systems yielded promising results in the dissection of heterosis. At the same time, they also highlighted the difficulty in pinpointing the exact loci responsible for heterosis, as successful QTL studies are often impeded by many technical challenges including a lack of high-density markers, inconsistent or imprecise phenotyping, and low recombination frequency in the mapping populations. Even a simple yeast QTL study, with none of the aforementioned technical problems, revealed that a case of overdominance traced back to a

complex locus of tightly linked genes having *cis*- and *trans*-acting epistatic interactions (Steinmetz et al., 2002). Moreover, critics of a widespread role for overdominant QTL in heterosis often point to the loss of heterotic phenotypes in F2 and subsequent generations (i.e., hybrid breakdown), which suggests that heterozygosity is needed at multiple loci and these "magical" combinations are reshuffled and lost during recombination.

A popular approach to circumvent issues with epistasis and hybrid breakdown has been through the creation of introgression line (IL) populations, where a series of inbred lines are generated using classical genetics to each carry a short chromosomal segment from a divergent parent. Ideally, the epistatic interactions that are not closely linked will be eliminated and recombination will not affect the small introgression, leaving only overdominance and pseudo-overdominance originating from one or more genes in the introgression as possible explanations for heterosis when the introgression is in a heterozygous state. The use of ILs was pioneered in tomato (*Solanum lycopersicum*), and a large phenomic study involving ILs and IL hybrids from a wild tomato species *Solanum pennellii* identified several chromosomal segments that confer heterosis (Semel et al., 2006). By evaluating a suite of 35 reproductive and nonreproductive traits in both homozygous and heterozygous ILs, overdominant QTL were found predominantly for reproductive traits, such as flower, fruit, and seed production. Although ILs allow homing in on potentially true overdominant loci, pseudo-overdominance cannot be ruled out because multiple genes reside in each introgression. However, pseudo-overdominance involves closely linked recessive QTL, and such QTL would be expected to be randomly distributed in the genome, and therefore one would not expect a bias for overdominant QTL toward reproductive traits as was found in this study. This led to the suggestion, albeit tenuous, that the heterotic effects originating from the wild species introgressions are likely to be truly overdominant (Semel et al., 2006). Although no additional evidence has been provided to support overdominant QTL in tomato, an important advance in this study is that putative cases of overdominance can originate from a very small chromosomal segment—a finding that has since been found in IL studies in other systems, including *Arabidopsis* (Lisec et al., 2008, 2009) and mice (Ishikawa, 2009).

Cases for Single-Gene Overdominance

Although few and far between, it is important to note that examples of overdominance tracing back to a single gene have been observed in several organisms, including yeast, plants, and animals (Schuler, 1954; Mukai & Burdick, 1959; Rédei, 1962; Efron, 1974; Hall & Wills, 1987; Grobet et al., 1997; Schuelke et al., 2004; Mosher et al., 2007; Delneri et al., 2008). These scattered reports, linking heterozygosity at a single gene to transgressive phenotypes, remain the most tantalizing from both fundamental and applied perspectives because single-gene overdominance could easily be leveraged for crop improvement. One hypothesis for the molecular nature of single-gene overdominance is that a combination of alleles encoding proteins adapted to different conditions, for example, isozymes, generates better performance in a wider condition spectrum, and thus an overdominant effect on the hybrid phenotype. For instance, nearly 40 years ago, Efron (1974) described a case of single-gene heterosis in maize for two inbred lines homozygous for different alcohol dehydrogenase (ADH) alleles that produced isozymes seemingly adapted to produce maximum enzyme activities in two different tissues: scutellum and pollen. Hybrids between the two lines now heterozygous for the two ADH alleles showed optimized enzyme activities in both tissues, thus expanding the enzyme activity spectrum and producing a more balanced overall metabolic efficiency (Efron, 1974). Intriguingly, heterozygosity for two distinct temperature-sensitive alleles of an alcohol dehydrogenase isozyme (ADH1) in yeast confers transgressive alcohol tolerance over either parent (Hall & Wills, 1987).

Interestingly, as in maize, the overdominance in yeast also seems to result from intermediate enzyme activity in the heterozygous individuals. Together, these studies imply that an expanded activity spectrum due to the unification of alleles adapted to different conditions might be a simple molecular mechanism to explain at least some cases of single-gene overdominance.

However, there are examples in the literature of single-gene overdominance involving mutant alleles in the heterozygous condition that cannot be explained by the aforementioned molecular mechanism, as only the dose of a gene and gene product is altered in mutant heterozygotes, and there is therefore no manipulation of activity spectrums as for isozymes. For instance, a classical example of mutant single-gene overdominance is that of heterozygosity for a mutant allele of hemoglobin. Individuals homozygous for mutated hemoglobin suffer from sickle cell anemia because the mutated protein causes the formation of abnormal crescent-shaped blood cells that have aggregation problems and are not efficient oxygen carriers; however, individuals who are heterozygous for one mutant copy of this gene do not suffer from this disease because they are also able to make normal blood cells, and remarkably they also have a higher resistance to malaria than individuals who are homozygous wild-type for the hemoglobin gene. Such early examples of mutant single-gene heterosis were criticized as rare and special conditional cases, and thus overdominance was still widely contested.

Classical work with *Drosophila* mutants, however, further alluded to the existence of overdominant genes, although the validity of these data was also greatly debated because of questions over environmental conditions, variation in genetic background, and standardization of growth conditions (Muller, 1928; Mukai & Burdick, 1959; Falk, 1960; Muller & Falk, 1960; Wallace & Dobzhansky, 1962; Wallace, 1963; Dobzhansky & Spassky, 1968). This early evidence for "mutant overdominance" in *Drosophila* (Mukai & Burdick, 1959), combined with an intriguing case involving the classical *erecta* mutant in *Arabidopsis* (Rédei, 1962), provided the impetus to carry out a large-scale study using tomato isogenic mutants to assess if, and to what extent, single-gene heterosis for yield can occur in heterozygous mutations (Semel et al., 2006; Krieger et al., 2010). A total of 35 homozygous isogenic fertile mutants in a field tomato variety (M82 *self-pruning* (sp^-) "determinate" background) were crossed with nonmutant parental control plants (M82 sp^- "determinate"), and hybrids were evaluated for increased fruit production (Semel et al., 2006). Remarkably, one mutant, *single flower truss* (*sft*), produced 60% more fruits in the heterozygous condition. Whereas homozygous *sft* mutants are severely delayed in flowering and they produce very few flowers compared to the nonmutant parent (M82 sp^-), *sft*/+ mutant heterozygotes produce more inflorescences and fruits compared to already high-yielding controls. It is important to note that epistasis was found to be an indispensable component of *SFT* heterosis: overdominance occurs only in the *sp* mutant background, revealing that interactions between *SP* and *SFT* must play a role (Krieger et al., 2010). Having only one functional *SFT* allele in *sft*/+ heterozygotes in the *sp* mutant background causes a weak dosage-dependent suppression of the *sp* determinate phenotype (Lifschitz et al., 2006; Krieger et al., 2010). Through a detailed genetic and phenomic analysis, the *sft*/+ overdominance was traced back to a developmental change in tomato shoot architecture, which is based on the "sympodial" compound shoot growth habit. Specifically, whereas *sp* mutants produce only three to four sympodial units (SYMs) and therefore inflorescences on each compound shoot due to the precocious termination of sympodial growth caused by the *sp* mutation (Pnueli et al., 1998; Krieger et al., 2010), heterozygosity for *sft* mutations enables one to two additional SYMs to develop, providing one to two additional inflorescences, each with seven to nine flowers, before termination (K. Jiang & Z.B. Lippman, unpublished data). As this suppression occurs on all primary and axillary shoots, the *sft*/+ dosage effect is quickly amplified to produce a whole plant with approximately 35% more inflorescences. This finding suggested an attractive new hypothesis for explaining heterosis, albeit in specific circumstances: weak semi-dominant effects on particular developmental processes

due to single-gene mutant heterozygosity can drive cumulative overdominance if the dosage effect is recurring and is amplified by the organism's specific development program. In this case, the sympodial growth habit of tomato, which is also found in other *Solanaceae* and is a hallmark of perennial plants such as vines and trees, amplifies the dosage effect caused by losing one functional copy of *SFT*.

Well before the tomato study was initiated, the *sft* and *sp* mutant phenotypes were found to be caused by mutations in the tomato orthologs of the *Arabidopsis* flowering hormone, florigen, encoded by *FT*, and its related antagonist, *TFL1*, respectively (Pnueli et al., 1998; Lifschitz et al., 2006). As a result, heterosis caused by these genes in tomato is due to manipulation of flowering, which could be argued in a special case. However, both old and new literature in diverse plants, including monocots and eudicots, suggest that the role of flowering and florigen in heterosis is a more general phenomenon and that dosage effects originating from allelic variation in flowering-time genes can drive transgressive variation for yield by subtle quantitative modulation of the plant reproductive transition. For example, a *Sorghum bicolor* mutant exhibiting strongly delayed flowering time in a day-length dependent manner also shows overdominance in the heterozygous state as a result of intermediate floweringtime (Quinby & Karper, 1946). Although the underlying gene has not been identified, the late-flowering phenotype of the mutant implies that *FT* could be involved, or at least a component of the florigen network (Quinby & Karper, 1946). In addition, a domestication QTL tracing back to a loss-of-function mutation in an *FT* paralog in sunflower also causes single-gene overdominance for flower size (Blackman et al., 2010). It should be noted that the dosage effect here may be more complex than in tomato, potentially involving interactions between a paralogous *FT* gene and its ancestor. However, importantly, all the aforementioned examples point toward a single gene, specifically the gene that encodes the vital mobile flowering signal, florigen. The dosage effect based on florigen and the manipulation of growth, transgressive or not, seems to be universal and likely to occur in all flowering plants. Indeed, in the diverse growth habits and plant systems from which FT-related heterosis has been observed, it may only be transgressive in certain developmental and environmental contexts, similar to many previously reported cases of full genome, IL, and single-gene heterosis (Efron, 1974; Li et al., 2001; Luo et al., 2001; Welch & Rieseberg, 2002; Semel et al., 2006; Mosher et al., 2007; Krieger et al., 2010). Intriguingly, a similar example of heterosis tracing back to two flowering-time genes functioning upstream of *FT* called *FRIGIDA* (*FRI*) and *FLOWERING LOCUS C* (*FLC*) was recently documented in *Arabidopsis* (Moore & Lukens, 2011). Beyond the flowering transition itself, heterosis has also been confirmed for another tomato mutant, *compound inflorescence* (*s*), which causes extensive inflorescence branching (Lippman et al., 2008). While homozygous *s* mutants have lower fruit production than wild-type due to low fruit set, half of the inflorescences in *s*/+ heterozygotes branch just once or twice, thereby effectively increasing inflorescence number without compromising the ability of the plants to set fruit (K. Jiang & Z.B. Lippman, unpublished data). Like for *sft*/+ heterozygotes, these findings suggest that deleterious "recessive" mutants affecting growth and development may have weak semidominant dosage effects in the heterozygous state that are only revealed upon comprehensive quantitative phenotyping. Such dosage effects due to one-component traits like flowering may be a major cause of yield heterosis in crops beyond tomato.

Intriguingly, while the above examples revolve around flowering pathways, cases of single-gene overdominance have been identified for other traits. For instance, in maize, fascinating, albeit inconclusive, support for single-gene heterosis involving heterozygosity for deleterious mutations impacting multiple traits beyond flowering has been revealed. Dollinger identified a series of recessive mutants in maize that negatively affect diverse aspects of development and growth, and therefore yield, and he crossed these mutants back to their isogenic inbred parents and observed

widespread heterotic phenotypes in the F1s (Dollinger, 1985). The heterotic effects from creating these mutant heterozygotes affected multiple aspects of growth, which suggested that just a single heterozygous gene could have dramatic pleiotropic impacts, resembling in many ways the heterosis caused by *sft*/+ mutant heterozygosity in tomato. Almost all aspects of yield were affected, including flowering time, plant height, ear size, kernel characteristics, and total yield. These findings suggest that maize mutations, classically defined as recessive, may in fact show dosage effects in the heterozygous condition, lending more support to the hypothesis that a single heterozygous mutation can drive heterosis through pleiotropic dosage-dependent changes on growth. From these several examples involving simpler genetic contexts of inbred lines, it may be reasonable to assume that dosage effects due to mutations are more ubiquitous in wild populations than previously expected due to the widespread masking of deleterious recessive alleles in nature (Charlesworth & Willis, 2009). Indeed, perhaps the genetic and molecular basis of both hybrid vigor and heterosis traces back not only to dominance complementation but also to dosage effects and pleiotropy.

Remarkably, single-gene overdominance extends beyond the plant kingdom to examples in animals as well. One intriguing example involves overdominance for muscle mass in various mammals including cattle, dogs, and humans. Originally described in cattle as the "double-muscling" phenotype (Grobet et al., 1997), this increased muscle mass has been directly associated with mutations in myostatin genes, and in some conditions these mutations are beneficial. For instance, whippet dogs that are heterozygous for a particular myostatin mutant allele are more muscular and have increased racing performance, while those homozygous for this mutation have excessive muscle that is detrimental to their athleticism (Mosher et al., 2007). A similar muscular disorder is found in humans—myostatin-related muscle hypertrophy (Schuelke et al., 2004). Similar to whippet dogs, humans who are homozygous for a particular mutation in myostatin (MSTN) have double muscle mass, whereas those heterozygous for the mutant allele have muscle mass intermediate to those individuals with two mutant alleles and those lacking the mutation (Schuelke et al., 2004), again suggesting that genetic dosage underlies the molecular mechanism for mutant single-gene overdominance.

Dosage: An Evolving Heterosis Model

When examined collectively, the cases of single-gene heterosis begin to evoke a picture of a tightly controlled balance between genes, protein products, and their phenotypic consequences in a complex multilayered network controlling development and reproduction, echoing an emerging theory of heterosis focusing on gene dosage and networks (Birchler et al., 2010), as well as protein metabolism (Goff, 2010). In this respect, the dramatic effects of single heterozygous mutations, although at this stage likely explaining only a subset of heterosis cases, manifest their outputs from quantitative modifications of development and growth networks whose underlying molecular components have been finely balanced by evolution. Robustness of such networks depends on the sensitivities to dosage changes of the individual components, especially for those central genes/components whose dosage changes likely lead to both unexpected and potentially dramatic pleiotropic and transgressive phenotypic consequences in both homozygous and heterozygous conditions. Dollinger, in elegantly summarizing his observations of single-gene heterosis with maize mutants, presciently stated that "wide ranging pleiotropic effects, either positive or negative and often quite large, represent the usual or normal situation for such visible recessive alleles when present as heterozygotes," and went further to foretell that "pleiotropic effects, whatever their genetic nature turns out be, provide a genetic basis for inbreeding depression and heterosis" (Dollinger, 1985). The implications of

these ideas are clear—any dosage perturbation of strongly pleiotropic genes, whether due to loss-of-function allelic changes as found in tomato *SFT* or gene duplications as in sunflower, is capable of causing dramatic and unexpected phenotypic consequences. Single-gene heterosis that has been observed independently for diverse genes and phenotypes across distantly related flowering plant lineages suggests that dosage-dependent heterosis arising from one or more heterozygous mutations may be more widespread than previously appreciated. It must also be remembered that single-gene heterosis is not confined to flowering pathways in plants (Efron, 1974; Dollinger, 1985). Moreover, it is significant that recent studies in organisms beyond plants have provided independent support for the role of single-gene dosage effects on heterosis-like phenomena. For example, a genome-wide screen of single-gene hemizygous mutations in diploid baker's yeast (*Saccharomyces cerevisiae*) revealed hundreds of genes showing "haploproficiency" for growth rate in multiple growth conditions (Delneri et al., 2008). An interesting observation from this study is that genes showing haploproficiency seem to be randomly distributed across the genome, consistent with the hypothesis that dosage effects from multiple components comprising diverse and complex networks of genes and gene products can drive transgressive phenotypes. The yeast study further highlights that transgression is often associated with specific growth conditions, consistent with the environment-dependent single-gene heterosis found in multicellular organisms. Indeed, semidominant dosage effects originating from single-gene heterozygosity frequently require a genetic or environmental "amplifier" to be translated into overall transgressive overdominant phenotypes, as is the case with tomato *SFT* heterosis depending on the *sp* mutation.

The dosage and network-centric view of heterosis is not new (Birchler et al., 2010). Emerging ideas on heterosis derived from the scattered cases of single-gene heterosis have been gaining momentum from a greater understanding and appreciation for the role of dosage-dependent regulatory mechanisms in growth and development. In several studies of aneuploidy and polyploidy in maize, it has been hypothesized that heteroallelic combinations of interacting proteins may improve the efficiency of interaction networks through combinatorial dosage effects, thus providing improved performance compared to homoallelic combinations (Riddle et al., 2010; Yao et al., 2011). Alternatively, subtle changes of mRNA or protein dosage due to heterozygosity may achieve a better balance or optimized state of molecular buffering in a complicated regulatory network, yielding improved efficiency and output from a signaling cascade (Veitia, 2010). This model is consistent with single-gene heterosis, as discussed above, because this single gene can be a "hub" in a network with a high number of connections, thus potentiating strong pleiotropic effects when dosage is altered. Of all examples, the "hub" hypothesis integrates well with *FT*-driven single-gene heterosis found in multiple plant lineages because *FT* is the central and universal integrator of multiple signal transduction cascades controlling environmental response and the decision to transition from vegetative growth to flowering (Jack, 2004). Another recent spin on the dosage-centric view of heterosis is centered around increased energy efficiency in hybrids achieved by allele-specific expression of dominant alleles and differences in protein metabolism in heterozygotes compared with parental lines (Goff, 2010). It should be noted that these hypotheses of dosage-dependent heterosis clearly distinguish themselves from the dominance/complementation hypothesis because dosage effects originate from semidominant functions of the genes involved, as opposed to full complementation.

In the past decade, both technological and theoretical advances have allowed acquisition and evaluation of much more data so that dosage effects and regulatory network structure and their response to perturbations can now be investigated in a highly quantitative manner. This prompted Birchler et al. (2010) to question the future value of restricting experimental design and interpretation to the classical heterosis models of dominance, overdominance, and epistasis for a more unified, dosage interaction-based framework. Indeed, dosage-centered thinking may prove more valuable

in the formulation and testing of hypotheses originating from the future system level integration of transcriptomic, proteomic, metabolomic, and phenomic data sets.

Conclusion

The spontaneous accumulation of deleterious recessive mutations genome-wide is a hallmark of population genetics, and this genetic load can be preserved through heterozygosity imposed through the selection of dominant complementing alleles. There is no doubt that the classic dominance model provides a key means for achieving population vigor, especially under environmental perturbations. Therefore, in natural settings, in light of the available genetic diversity, it is difficult to envision cases of single-gene overdominance. However, there are clearly loci and mutant alleles that are selected for in nature that may act in a semidominant dosage-dependent manner that translate to pleiotropic dosage effects in the heterozygous state that can act cumulatively or perhaps epistatically on several component traits to improve fitness of individuals and populations. Indeed, as some of these fitness traits are also attractive from the agricultural perspective, it is likely that breeders have inadvertently harnessed such mechanisms of overdominance to drive heterosis in agriculture. After all, the process of domestication is founded on the fixation of multiple mutations (Doebley et al., 2006). We therefore propose that nature is full of as yet undiscovered examples of allelic heterozygosity that are causing both weak and strong dosage effects on growth and development leading to subtle, yet significant, transgressive phenotypic effects. Future population genetics studies may reveal such loci and their dosage effects as driving forces of both heterosis in crops and hybrid vigor in nature.

Acknowledgments

We regret that due to space limitations and the expansive body of literature on the subject, we were unable to cite all relevant genetic and molecular studies on heterosis, especially in systems not mentioned in this chapter. We owe special thanks to Uri Krieger for our collaborations on heterosis and to Yuval Eshed and Eliezer Lifschitz for many valuable discussions on florigen and flowering. Research in the Zamir laboratory has been funded by the European Research Council YIELD project. KLL is supported by a National Science Foundation Graduate Research Fellowship (DGE-0914548). Research in the Lippman laboratory is supported through a Heterosis Challenge Grant (HCG) grant from the NSF Plant Genome Research Program (DBI-0922442).

References

Anderson, J.B., Funt, J., Thompson, D.A., et al. (2010) Determinants of divergent adaptation and Dobzhansky-Muller interaction in experimental yeast populations. *Curr Biol* **20** (15), 1383–1388.

Auger, D.L., Gray, A.D., Ream, T.S., Kato, A., Coe, E.H., Jr, & Birchler, J.A. (2005) Nonadditive gene expression in diploid and triploid hybrids of maize. *Genetics* **169** (1), 389–397.

Bateson, W. (1909) *Heredity and Variation in Modern Lights*. Cambridge University Press, Cambridge.

Bikard, D., Patel, D., Le Mette, C., et al. (2009) Divergent evolution of duplicate genes leads to genetic incompatibilities within *A. thaliana*. *Science* **323** (5914), 623–626.

Birchler, J.A., Auger, D.L., & Riddle, N.C. (2003) In search of the molecular basis of heterosis. *Plant Cell* **15** (10), 2236–2239.

Birchler, J.A., Bhadra, U., Bhadra, M.P., & Auger, D.L. (2001) Dosage-dependent gene regulation in multicellular eukaryotes: implications for dosage compensation, aneuploid syndromes, and quantitative traits. *Dev Biol* **234** (2), 275–288.

Birchler, J.A., Riddle, N.C., Auger, D.L., & Veitia, R.A. (2005) Dosage balance in gene regulation: biological implications. *Trends Genet* **21** (4), 219–226.
Birchler, J.A., Yao, H., & Chudalayandi, S. (2007) Biological consequences of dosage dependent gene regulatory systems. *Biochim Biophys Acta* **1769** (5–6), 422–428.
Birchler, J.A., Yao, H., Chudalayandi, S., Vaiman, D., & Veitia, R.A. (2010) Heterosis. *Plant Cell* **22** (7), 2105–2112.
Blackman, B.K., Strasburg, J.L., Raduski, A.R., Michaels, S.D., & Rieseberg, L.H. (2010) The role of recently derived FT paralogs in sunflower domestication. *Curr Biol* **20** (7), 629–35.
Bomblies, K., Lempe, J., Epple, P., et al. (2007) Autoimmune response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in plants. *PLoS Biol* **5** (9), e236.
Bomblies, K., & Weigel, D. (2007) Hybrid necrosis: autoimmunity as a potential gene-flow barrier in plant species. *Nat Rev Genet* **8** (5), 382–393.
Brideau, N.J., Flores, H.A., Wang, J., Maheshwari, S., Wang, X., & Barbash, D.A. (2006) Two Dobzhansky-Muller genes interact to cause hybrid lethality in *Drosophila*. *Science* **314** (5803), 1292–1295.
Bruce, A.B. (1910) The Mendelian theory of heredity and the augmentation of vigor. *Science* **32** (827), 627–628.
Burke, J.M., & Arnold, M.L. (2001) Genetics and the fitness of hybrids. *Annu Rev Genet* **35**, 31–52.
Castle, W.E. (1926) The explanation of hybrid vigor. *Proc Natl Acad Sci USA* **12** (1), 16–19.
Charlesworth, D., & Willis, J.H. (2009) The genetics of inbreeding depression. *Nat Rev Genet* **10** (11), 783–796.
Chen, Z.J. (2010) Molecular mechanisms of polyploidy and hybrid vigor. *Trends Plant Sci* **15** (2), 57–71.
Crow, J.F. (1948) Alternative hypotheses of hybrid vigor. *Genetics* **33** (5), 477–487.
Crow, J.F. (2008) Mid-century controversies in population genetics. *Annu Rev Genet* **42**, 1–16.
Darwin, C. (1868) *The Variation of Animals and Plants under Domestication*, vol. 2. John Murray, London.
Darwin, C. (1876) *The Effects of Cross and Self-fertilization in the Vegetable Kingdom*. John Murray, London.
Davies, W.P. (2003) An historical perspective from the Green Revolution to the gene revolution. *Nutr Rev* **61** (6 Pt 2), S124–S134.
Delneri, D., Hoyle, D.C., Gkargkas, K., et al. (2008) Identification and characterization of high-flux-control genes of yeast through competition analyses in continuous cultures. *Nat Genet* **40** (1), 113–117.
Dobzhansky, T. (1950) Genetics of natural populations. XIX. Origin of heterosis through natural selection in populations of *Drosophila pseudoobscura*. *Genetics* **35** (3), 288–302.
Dobzhansky, T., & Spassky, B. (1968) Genetics of natural populations. XL. Heterotic and deleterious effects of recessive lethals in populations of *Drosophila pseudoobscura*. *Genetics* **59** (3), 411–425.
Doebley, J.F., Gaut, B.S., & Smith, B.D. (2006) The molecular genetics of crop domestication. *Cell* **127**, 1309–1321.
Dollinger, E.J. (1985) Effects of visible recessive alleles on vigor characteristics in maize hybrid. *Crop Sci* **25**, 819–821.
Duvick, D.N. (1999) Heterosis: feeding people and protecting natural resources. In: *The Genetics and Exploration of Heterosis in Crops* (eds J.G. Coors & S. Pandey), American Society of Agronomy, Inc., Crop Science Society of America, Inc., Soil Science Society of America, Inc, Madison, WI.
Duvick, D.N. (2001) Biotechnology in the 1930s: the development of hybrid maize. *Nat Rev Genet* **2** (1), 69–74.
Duvick, D.N. (2005) The contribution of breeding to yield advances in maize (*Zea mays* L.). *Adv Agron* **86**, 83–145.
East, E.M. (1908) Inbreeding in corn. *Reports of the Connecticut Agricultural Experiment Station for Years 1907–1908*. Connecticut Agricultural Experiment Station.
East, E.M. (1909) The distinction between development and heredity in inbreeding. *Am Nat* **43** (507), 173–181.
Efron, Y. (1974) Specific differences in maize alcohol dehydrogenase: possible explanation of heterosis at the molecular level. *Nat New Biol* **241**, 41–42.
Falk, R. (1960) Are induced mutations in *Drosophila* overdominant? II. Experimental results. *Genetics* **46**, 737–757.
Fry, J.D., Nuzhdin, S.V., Pasyukova, E.G., & Mackay, T.F. (1998) QTL mapping of genotype-environment interaction for fitness in *Drosophila melanogaster*. *Genet Res* **71** (2), 133–141.
Goff, S.A. (2010) A unifying theory for general multigenic heterosis: energy efficiency, protein metabolism, and implications for molecular breeding. *New Phytol* **189** (4), 923–937.
Gore, M.A., Chia, J.M., Elshire, R.J., et al. (2009) A first-generation haplotype map of maize. *Science* **326** (5956), 1115–1117.
Graham, G.I., Wolff, D.W., & Stuber, C.W. (1997) Characterization of a yield quantitative trait locus on chromosome five of maize by fine mapping. *Crop Sci* **37** (5), 1601–1610.
Grobet, L., Martin, L.J.R., Poncelet, D., et al. (1997) A deletion in the bovine myostatin gene causes the double-muscled phenotype in cattle. *Nat Genet* **17** (1), 71–74.
Guo, M., Rupe, M.A., Yang, X., et al. (2006) Genome-wide transcript analysis of maize hybrids: allelic additive gene expression and yield heterosis. *Theor Appl Genet* **113** (5), 831–845.
Guo, M., Rupe, M.A., Zinselmeier, C., Habben, J., Bowen, B.A., & Smith, O.S. (2004) Allelic variation of gene expression in maize hybrids. *Plant Cell* **16** (7), 1707–1716.
Gur, A., & Zamir, D. (2004) Unused natural variation can lift yield barriers in plant breeding. *PLoS Biol* **2** (10), e245.

Hall, J.G., & Wills, C. (1987) Conditional overdominance at an alcohol dehydrogenase locus in yeast. *Genetics* **117** (3), 421–427.

Hermansen, J.S., Saether, S.A., Elgvin, T.O., Borge, T., Hjelle, E., & Saetre, G.P. (2011) Hybrid speciation in sparrows I: phenotypic intermediacy, genetic admixture and barriers to gene flow. *Mol Ecol* **20** (18), 3812–3822.

Hochholdinger, F., & Hoecker, N. (2007) Towards the molecular basis of heterosis. *Trends Plant Sci* **12** (9), 427–432.

Hoecker, N., Keller, B., Muthreich, N., et al. (2008) Comparison of maize (*Zea mays* L.) F1-hybrid and parental inbred line primary root transcriptomes suggests organ-specific patterns of nonadditive gene expression and conserved expression trends. *Genetics* **179** (3), 1275–1283.

Houle, D. (1989) Allozyme-associated heterosis in *Drosophila melanogaster*. *Genetics* **123** (4), 789–801.

Huang, Y., Zhang, L., Zhang, J., et al. (2006) Heterosis and polymorphisms of gene expression in an elite rice hybrid as revealed by a microarray analysis of 9198 unique ESTs. *Plant Mol Biol* **62** (4–5), 579–591.

Ishikawa, A. (2009) Mapping an overdominant quantitative trait locus for heterosis of body weight in mice. *J Hered* **100** (4), 501–504.

Jack, T. (2004) Molecular and genetic mechanisms of floral control. *Plant Cell* **16** (Suppl.), S1–17.

Jefferson, M.C., Crumpacker, D.W., & Williams, J.S. (1974) Cold temperature resistance, chromosomal polymorphism and interpopulation heterosis in *Drosophila pseudoobscura*. *Genetics* **76** (4), 807–822.

Jones, D.F. (1917) Dominance of linked factors as a means of accounting for heterosis. *Genetics* **2** (5), 466–479.

Krieger, U., Lippman, Z.B., & Zamir, D. (2010) The flowering gene SINGLE FLOWER TRUSS drives heterosis for yield in tomato. *Nat Genet* **42** (5), 459–463.

Lai, J., Li, R., Xu, X., et al. (2010) Genome-wide patterns of genetic variation among elite maize inbred lines. *Nat Genet* **42** (11), 1027–1030.

Li, B., Zhang, D.-F., Jia, G.-Q., Dai, J.-R., & Wang, S.-C. (2009a) Genome-wide comparisons of gene expression for yield heterosis in maize. *Plant Mol Biol Rep* **27**, 162–176.

Li, X., Wei, Y., Nettleton, D., & Brummer, C.E. (2009b) Comparative gene expression profiles between heterotic and non-heterotic hybrids of tetraploid *Medicago sativa*. *BMC Plant Biol* **9**, 107–119.

Li, Z.K., Luo, L.J., Mei, H.W., et al. (2001) Overdominant epistatic loci are the primary genetic basis of inbreeding depression and heterosis in rice. I. Biomass and grain yield. *Genetics* **158** (4), 1737–1753.

Lifschitz, E., Eviatar, T., Rozman, A., et al. (2006) The tomato FT ortholog triggers systemic signals that regulate growth and flowering and substitute for diverse environmental stimuli. *Proc Natl Acad Sci USA* **103** (16), 6398–6403.

Lippman, Z.B., Cohen, O., Alvarez, J.P., et al. (2008) The making of a compound inflorescence in tomato and related nightshades. *PLoS Biol* **6** (11), e288.

Lippman, Z.B., & Zamir, D. (2007) Heterosis: revisiting the magic. *Trends Genet* **23** (2), 60–66.

Lisec, J., Meyer, R.C., Steinfath, M., et al. (2008) Identification of metabolic and biomass QTL in *Arabidopsis thaliana* in a parallel analysis of RIL and IL populations. *Plant J Cell Mol Biol* **53** (6), 960–972.

Lisec, J., Steinfath, M., Meyer, R.C., et al. (2009) Identification of heterotic metabolite QTL in *Arabidopsis thaliana* RIL and IL populations. *Plant J Cell Mol Biol* **59** (5), 777–788.

Luo, L.J., Li, Z.K., Mei, H.W., et al. (2001) Overdominant epistatic loci are the primary genetic basis of inbreeding depression and heterosis in rice. II. Grain yield components. *Genetics* **158** (4), 1755–1771.

Mather, K. (1955) The genetical basis of heterosis. *Proc R Soc Lond B Biol Sci* **144** (915), 143–150.

McMullen, M.D., Kresovich, S., Villeda, H.S., et al. (2009) Genetic properties of the maize nested association mapping population. *Science* **325** (5941), 737–740.

Mendel, G. (1865) Versuche über Plflanzenhybriden. *Abhandlungen* 3–47.

Meyer, R.C., Steinfath, M., Lisec, J., et al. (2007) The metabolic signature related to high plant growth rate in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* **104** (11), 4759–4764.

Mitchell-Olds, T. (1995) Interval mapping of viability loci causing heterosis in *Arabidopsis*. *Genetics* **140** (3), 1105–1109.

Moore, S., & Lukens, L. (2011) An evaluation of *Arabidopsis thaliana* hybrid traits and their genetic control. *G3* **1** (7), 571–579.

Mosher, D.S., Quignon, P., Bustamante, C.D., et al. (2007) A mutation in the myostatin gene increases muscle mass and enhances racing performance in heterozygote dogs. *PLoS Genet* **3** (5), e79.

Mukai, T., & Burdick, A.B. (1959) Single gene heterosis associated with a second chromosome recessive lethal in *Drosophila melanogaster*. *Genetics* **44** (2), 211–232.

Muller, H.J. (1928) The measurement of gene mutation rate in *Drosophila*, its high variability, and its dependence on temperature. *Genetics* **13**, 279–357.

Muller, H.J. (1940) *Bearing of the Drosophila Work on Systematics*. Clarendon Press, Oxford.

Muller, H.J. (1942) Isolating mechanisms, evolution and temperature. *Biol Symp* **6**, 71–125.

Muller, H.J., & Falk, R. (1960) Are induced mutations in *Drosophila* overdominant? I. Experimental design. *Genetics* **46**, 727–735.
Orr, H.A. (1996) Dobzhansky, Bateson, and the genetics of speciation. *Genetics* **144** (4), 1331–1335.
Peters, A.D., Halligan, D.L., Whitlock, M.C., & Keightley, P.D. (2003) Dominance and overdominance of mildly deleterious induced mutations for fitness traits in *Caenorhabditis elegans*. *Genetics* **165** (2), 589–599.
Pnueli, L., Carmel-Goren, L., Hareven, D., et al. (1998) The SELF-PRUNING gene of tomato regulates vegetative to reproductive switching of sympodial meristems and is the ortholog of CEN and TFL1. *Development* **125** (11), 1979–1989.
Quinby, J.R., & Karper, R.E. (1946) Heterosis in sorghum resulting from the heterozygous condition of a single gene that affects duration of growth. *Am J Bot* **33** (9), 716–721.
Rédei, G.P. (1962) Single locus heterosis. *Z Ver* **93**, 164–170.
Riddle, N.C., Jiang, H., An, L., Doerge, R.W., & Birchler, J.A. (2010) Gene expression analysis at the intersection of ploidy and hybridity in maize. *Theor Appl Genet* **120** (2), 341–353.
Schuelke, M., Wagner, K.R., Stolz, L.E., et al. (2004) Myostatin mutation associated with gross muscle hypertrophy in a child. *N Engl J Med* **350** (26), 2682–2688.
Schuler, J.F. (1954) Natural mutations in inbred lines of maize and their heterotic effect. I. Comparison of parent, mutant and their F(1) hybrid in a highly inbred background. *Genetics* **39** (6), 908–922.
Semel, Y., Nissenbaum, J., Menda, N., et al. (2006) Overdominant quantitative trait loci for yield and fitness in tomato. *Proc Natl Acad Sci USA* **103** (35), 12981–12986.
Shull, G.H. (1908) The composition of a field of maize. *Am Breed Assn Rep* **4**, 269–301.
Singleton, W.R. (1941) Hybrid vigor and its utilization in sweet corn breeding. *Am Nat* **75**, 48–60.
Smith, L.M., Bomblies, K., & Weigel, D. (2011) Complex evolutionary events at a tandem cluster of *Arabidopsis thaliana* genes resulting in a single-locus genetic incompatibility. *PLoS Genet* **7** (7), e1002164.
Song, S., Qu, H., Chen, C., Hu, S., & Yu, J. (2007) Differential gene expression in an elite hybrid rice cultivar (*Oryza sativa* L.) and its parental lines based on SAGE data. *BMC Plant Biol* **7**, 49.
Springer, N.M., & Stupar, R.M. (2007a) Allele-specific expression patterns reveal biases and embryo-specific parent-of-origin effects in hybrid maize. *Plant Cell* **19** (8), 2391–2402.
Springer, N.M., & Stupar, R.M. (2007b) Allelic variation and heterosis in maize: how do two halves make more than a whole?. *Genome Res* **17** (3), 264–275.
Springer, N.M., Ying, K., Fu, Y., et al. (2009) Maize inbreds exhibit high levels of copy number variation (CNV) and presence/absence variation (PAV) in genome content. *PLoS Genet* **5** (11), e1000734.
Steinmetz, L.M., Sinha, H., Richards, D.R., et al. (2002) Dissecting the architecture of a quantitative trait locus in yeast. *Nature* **416** (6878), 326–330.
Stokes, D., Fraser, F., Morgan, C., et al. (2010) An association transcriptomics approach to the prediction of hybrid performance. *Mol Breed* **26**, 91–106.
Stuber, C.W., Lincoln, S.E., Wolff, D.W., Helentjaris, T., & Lander, E.S. (1992) Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. *Genetics* **132** (3), 823–839.
Stupar, R.M., & Springer, N.M. (2006) Cis-transcriptional variation in maize inbred lines B73 and Mo17 leads to additive expression patterns in the F1 hybrid. *Genetics* **173** (4), 2199–2210.
Sun, Q.X., Wu, L.M., Ni, Z.F., Meng, F.R., Wang, Z.K., & Lin, Z. (2004) Differential gene expression patterns in leaves between hybrids and their parental inbreds are correlated with heterosis in a wheat diallel cross. *Plant Sci* **166**, 651–657.
Swanson-Wagner, R.A., DeCook, R., Jia, Y., et al. (2009) Paternal dominance of trans-eQTL influences gene expression patterns in maize hybrids. *Science* **326** (5956), 1118–1120.
Swanson-Wagner, R.A., Jia, Y., DeCook, R., Borsuk, L.A., Nettleton, D., & Schnable, P.S. (2006) All possible modes of gene action are observed in a global comparison of gene expression in a maize F1 hybrid and its inbred parents. *Proc Natl Acad Sci USA* **103** (18), 6805–6810.
Tang, S., & Presgraves, D.C. (2009) Evolution of the *Drosophila* nuclear pore complex results in multiple hybrid incompatibilities. *Science* **323** (5915), 779–782.
Taylor, S.J., Willard, R.W., Shaw, J.P., Dobson, M.C., & Martin, N.H. (2011) Differential response of the homoploid hybrid species *Iris nelsonii* (Iridaceae) and its progenitors to abiotic habitat conditions. *Am J Bot* **98** (8), 1309–1316.
Uzarowska, A., Keller, B., Piepho, H.P., et al. (2007) Comparative expression profiling in meristems of inbred-hybrid triplets of maize based on morphological investigations of heterosis for plant height. *Plant Mol Biol* **63** (1), 21–34.
Veitia, R.A. (2010) A generalized model of gene dosage and dominant negative effects in macromolecular complexes. *FASEB J* **24** (4), 994–1002.
Vuylsteke, M., van Eeuwijk, F., Van Hummelen, P., Kuiper, M., & Zabeau, M. (2005) Genetic analysis of variation in gene expression in *Arabidopsis thaliana*. *Genetics* **171** (3), 1267–1275.
Wallace, B. (1963) Further data on the overdominance of induced mutations. *Genetics* **48** (5), 633–651.

Wallace, B., & Dobzhansky, T. (1962) Experimental proof of balanced genetic loads in *Drosophila*. *Genetics* **47**, 1027–1042.
Welch, M.E., & Rieseberg, L.H. (2002) Habitat divergence between a homoploid hybrid sunflower species, *Helianthus paradoxus* (Asteraceae), and its progenitors. *Am J Bot* **89** (3), 472–478.
Xiao, J., Li, J., Yuan, L., & Tanksley, S.D. (1995) Dominance is the major genetic basis of heterosis in rice as revealed by QTL analysis using molecular markers. *Genetics* **140** (2), 745–754.
Yao, H., Kato, A., Mooney, B., & Birchler, J.A. (2011) Phenotypic and gene expression analyses of a ploidy series of maize inbred Oh43. *Plant Mol Biol* **75** (3), 237–251.
Yu, S.B., Li, J.X., Xu, C.G., et al. (1997) Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. *Proc Natl Acad Sci USA* **94** (17), 9226–9231.

OPEN ACCESS Freely available online

PLOS | GENETICS

Tomato Yield Heterosis Is Triggered by a Dosage Sensitivity of the Florigen Pathway That Fine-Tunes Shoot Architecture

Ke Jiang[1], Katie L. Liberatore[1], Soon Ju Park[1], John P. Alvarez[2], Zachary B. Lippman[1]*

1 Watson School of Biological Sciences, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, United States of America, **2** Monash University, School of Biological Sciences, Clayton Campus, Melbourne, Victoria, Australia

Abstract

The superiority of hybrids has long been exploited in agriculture, and although many models explaining "heterosis" have been put forth, direct empirical support is limited. Particularly elusive have been cases of heterozygosity for single gene mutations causing heterosis under a genetic model known as overdominance. In tomato (*Solanum lycopersicum*), plants carrying mutations in *SINGLE FLOWER TRUSS* (*SFT*) encoding the flowering hormone florigen are severely delayed in flowering, become extremely large, and produce few flowers and fruits, but when heterozygous, yields are dramatically increased. Curiously, this overdominance is evident only in the background of "determinate" plants, in which the continuous production of side shoots and inflorescences gradually halts due to a defect in the flowering repressor *SELF PRUNING* (*SP*). How *sp* facilitates *sft* overdominance is unclear, but is thought to relate to the opposing functions these genes have on flowering time and shoot architecture. We show that *sft* mutant heterozygosity (*sft/+*) causes weak semi-dominant delays in flowering of both primary and side shoots. Using transcriptome sequencing of shoot meristems, we demonstrate that this delay begins before seedling meristems become reproductive, followed by delays in subsequent side shoot meristems that, in turn, postpone the arrest of shoot and inflorescence production. Reducing *SFT* levels in *sp* plants by artificial microRNAs recapitulates the dose-dependent modification of shoot and inflorescence production of *sft/+* heterozygotes, confirming that fine-tuning levels of functional *SFT* transcripts provides a foundation for higher yields. Finally, we show that although flowering delays by florigen mutant heterozygosity are conserved in *Arabidopsis*, increased yield is not, likely because cyclical flowering is absent. We suggest *sft* heterozygosity triggers a yield improvement by optimizing plant architecture via its dosage response in the florigen pathway. Exploiting dosage sensitivity of florigen and its family members therefore provides a path to enhance productivity in other crops, but species-specific tuning will be required.

Citation: Jiang K, Liberatore KL, Park SJ, Alvarez JP, Lippman ZB (2013) Tomato Yield Heterosis Is Triggered by a Dosage Sensitivity of the Florigen Pathway That Fine-Tunes Shoot Architecture. PLoS Genet 9(12): e1004043. doi:10.1371/journal.pgen.1004043

Editor: Nathan M. Springer, University of Minnesota, United States of America

Received July 8, 2013; **Accepted** November 6, 2013; **Published** December 26, 2013

Funding: This research was supported by an NSF Graduate Research Fellowship (DGE-0914548) to KLL, and a grant from the NSF Plant Genome Research Program (DBI-0922442) to ZBL. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

* E-mail: lippman@cshl.edu

Introduction

More than a century ago, simple garden studies by Darwin revealed a remarkable phenomenon in which crossing related varieties of plants produced hybrid progeny with superior growth and fecundity compared to their parents [1]. Understanding this hybrid vigor began with population genetics theories postulating that outcrossing facilitates adaptation and improves fitness by shuffling allelic diversity to thwart inbreeding depression [2]. However, it was the agricultural exploitation of hybrid vigor, or "heterosis," in both crop and animal breeding that propelled efforts to dissect its genetic and molecular bases [3–10]. Maize geneticists noted early on that inbreeding prior to hybridization drives yield heterosis, and heterotic effects generally improve with greater genetic distance between parental lines [3]. These observations led to the notion that heterosis derives from genome-wide masking of independently accrued deleterious recessive mutations. Extensive quantitative genetic, transcriptomic, and genomic sequencing studies in crop and model plants have provided widespread indirect support for a "dominance complementation" model [2,6,11]; however, there is lingering evidence that a model known as overdominance might also contribute to heterosis [5–8]. Overdominance has long been an appealing explanation, because theoretically heterozygosity at only a single gene is needed to cause heterotic effects, presumably from intra-locus allelic interactions functionally superseding any one allelic form. However, the relevance of overdominance for yield and whether allelic interactions are the underlying cause remains controversial, primarily because quantitative trait locus (QTL) mapping studies reporting overdominant QTL have failed to pinpoint responsible genes [12–16]. Importantly, though, there have been scattered reports of single gene overdominance over the years, and among these have been several unexplained examples from yeast, plants, and animals involving heterozygosity for single gene loss-of-function mutations [17–24].

We previously reported a dramatic case of overdominance for tomato yield in multiple environments and planting densities resulting from loss-of-function mutations in the gene *SINGLE*

Author Summary

For over a century, it has been known that inbreeding harms plant and animal fitness, whereas interbreeding between genetically distinct individuals can lead to more robust offspring in a phenomenon known as hybrid vigor, or heterosis. While heterosis has been harnessed to boost agricultural productivity, its causes are not understood. Especially controversial is a model called "overdominance," which states in its simplest form that a single gene can drive heterosis, although multiple overdominant genes can also contribute. In tomato, a mutation in just one of two copies of a gene encoding the flowering hormone called florigen causes remarkable increases in yield, but it is not known why. We show that yield increases are triggered by a fine-tuning of florigen levels that cause subtle delays in the time it takes all shoots to produce flowers. The resulting plant architecture maximizes yield in varieties that dominate the processing tomato industry. We show that while similar changes in flowering occur when one copy of florigen is mutated in the model crucifer plant *Arabidopsis*, yield is not increased, suggesting that, while manipulating florigen holds potential to improve crop productivity, the tuning of florigen and related genes will have to be tailored according to species.

FLOWER TRUSS (*SFT*) encoding the generic flowering hormone florigen [25]. Tomato yield, on both a per plant basis and in the context of tons per acre, depends partly on fruit size, but is mainly driven by the production of dozens of multi-flowered inflorescences and resulting fruit clusters that develop according to the "sympodial" growth habit [26]. The defining feature of sympodial plants is the shoot apical meristem (SAM) ends growth by differentiating into a terminal flower after producing a set number of leaves, and growth then renews from a specialized axillary (i.e. sympodial) meristem (SYM) that, in tomato, produces just three leaves before undergoing its own flowering transition and termination. Indefinite reiteration of three-leaf sympodial flowering events results in an "indeterminate" plant that continuously produces equally spaced inflorescences (Figure 1A). In homozygous *sft* mutants, reduced florigen signals delay the transition to reproductive growth and cause a substantial loss of flower production and yield due to loss of sympodial growth and conversion of inflorescences into leafy vegetative shoots producing scattered flowers [27]. Counter-intuitively, *sft*/+ heterozygotes generate more inflorescences, flowers, and harvestable ripe fruits compared to parental controls in the same growing period, but these effects are limited to "determinate" tomato types in which sympodial shoot and inflorescence production ends prematurely due to a classical mutation in the gene *SELF PRUNING* (*SP*) (Figure 1A) [25,26]. Notably, *SP* is a flowering repressor and a known florigen antagonist in the *SFT* gene family, implying that *SFT*-dependent yield heterosis is likely directly linked to the flowering transition, and specifically to the opposing functional relationship of *SP* to *SFT*.

Tomato breeding goals are multifaceted and shift according to the needs and desires of growers (e.g. improved pest resistances) and consumers (e.g. better quality), but one unwavering aim is to improve yield. Indeterminate cultivars are grown commercially to enable continuous market delivery of "round," "roma," "cocktail," "grape," and "cherry" tomato types that are eaten fresh and command a premium price. Indeterminate tomatoes are primarily grown in greenhouses where successively ripening clusters are harvested by hand multiple times over an extended period, in some cases up to a year, to maximize yield on plants that must be pruned to one or two main shoots to enable efficient greenhouse growth and maintain fresh market quality [28]. While the necessary pruning of indeterminate tomatoes facilitates agronomic practices that maximize quality, such as size, shape, and flavor, it also limits yield [29]. In contrast, tomatoes grown for sauces, pastes, juices, or other processed can or jar products where fruit quality is less relevant, must be managed agronomically to produce maximum yields (per acre) through once-over mechanical harvests to be economically justified [28]. Maximal yields for processing tomatoes are achieved by growing determinate *sp* mutants in the open field to their full potential, because sequential sympodial shoots transition to flowering progressively faster in *sp* plants, which results in a compact bush-like form where fruits ripen uniformly (Figure 1A) [26]. Thus, *sp* varieties lend themselves to once-over mechanical harvesting and have therefore come to dominate the processing tomato industry, although determinate varieties have also been bred for fresh market production [28]. In a parallel to the physical pruning of indeterminate tomatoes, one drawback of *sp*-imposed determinate growth is that inflorescence and fruit production is restricted, because of a genetic pruning that causes sympodial cycling to stop. Thus, strategies to improve processing tomato yield are limited, primarily because the most logical approach of simply increasing sympodial flowering events would lead back to indeterminate growth and large plants that perform poorly in the field from competition and a loss of uniform ripening. Thus, maximizing inflorescence and fruit production while simultaneously minimizing shoot production for the processing tomato industry has remained a challenging goal. To explore how interactions between mutations in *SP* and *SFT* affect tomato flowering to create a new optimum for fruit yield, we explored tomato *sft* heterosis from a developmental and molecular context of the reproductive transition and its impact on plant architecture and inflorescence production.

Results

sft/+ heterozygosity suppresses sympodial shoot termination in determinate tomatoes

The discovery that *sft*/+ heterozygosity in an *sp* background (*sft*/+ *sp*) dramatically increases fruit production while only modestly increasing plant size was remarkable, but explaining this single gene overdominant effect was limited to showing that the yield boost mostly came from *sft*/+ *sp* plants having altered sympodial architectures that lead to more inflorescences [25]. *sft* mutant phenotypes are epistatic over *sp* [27], leading us to speculate that having only one functional allele of *SFT* might result in a dose-dependent partial suppression of *sp* determinacy. Indeed, heterosis disappears in a functional *SP* background [25]; yet, how the *sft*/+ *sp* genetic constitution affects the flowering process to create a new optimum for yield has not been resolved. To address this, we grew *sp* and *sft*/+ *sp* plants in controlled greenhouse conditions to precisely compare inflorescence production and flowering times of recurring sympodial shoots on the main axis (i.e. derived from the primary shoot; Figure 1A). We found an average of 1.5 more inflorescences and sympodial units on *sft*/+ *sp* plants, confirming a delay in sympodial termination (Figure 1B). To determine whether this was based on a delay in the flowering transition of each sympodial shoot, we measured leaf number in the first three units and observed a modest, but significant, increase in leaf production (Figure 1C). Importantly, and as expected [25], these delays required the *sp* background, as *sft*/+ heterozygosity alone produced three-leaf sympodial units like WT (Figure 1C). Impor-

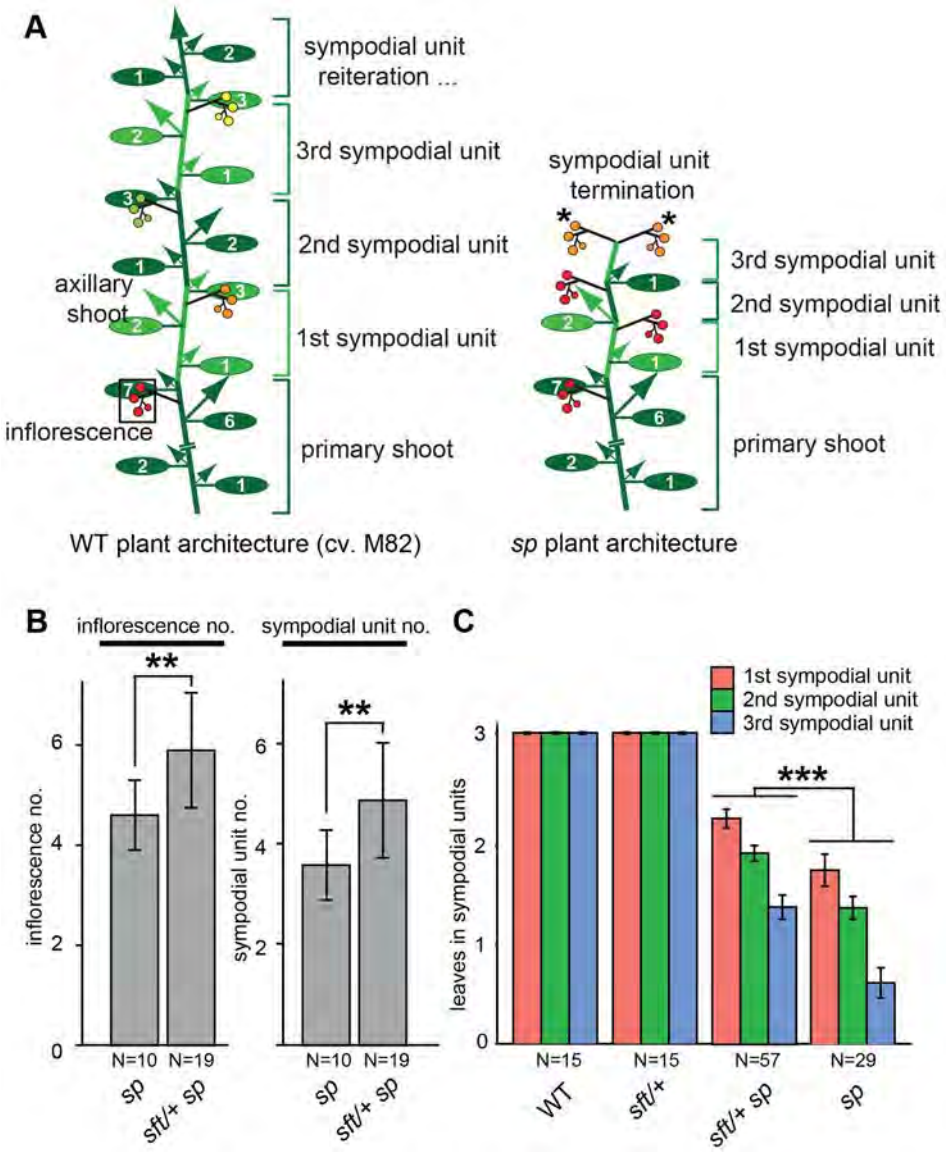


Figure 1. Precocious shoot termination in determinate tomatoes is partially suppressed by *sft*/+ mutant heterozygosity. (A) Schematic diagrams showing shoot architecture of a wild type (WT) indeterminate tomato plant (left) and an *sp* determinate mutant (right). In WT M82 plants the primary shoot meristem (PSM) from the embryo gives rise to 7–9 leaves before terminating in the first flower of the first multi-flowered inflorescence (boxed). A specialized axillary meristem called a sympodial meristem (SYM) in the axil of the last leaf on primary shoot then generates three leaves before terminating in the first flower of the next inflorescence. In indeterminate tomatoes, this process continues indefinitely (left). In *sp* mutants (right), sympodial cycling accelerates progressively on all shoots causing leaf production to decrease in successive units until growth ends in two juxtaposed inflorescences (asterisks). Alternating colored groups of three ovals represent leaves within successive sympodial units numbered at right. Colored circles represent fruits and flowers within each inflorescence (red: fully ripe fruit; orange: ripening fruit; green: unripe fruit; yellow: flowers) and arrows represent canonical axillary shoots. (B) Compared to *sp* mutants alone, *sft*/+ *sp* plants produce more inflorescences (left) and sympodial units (right) before sympodial cycling terminates on the main shoot. Genotypes and sample sizes are shown below, and standard deviations of averages are presented. (C) Compared to *sp* alone, *sft*/+ *sp* plants produce more leaves in the first three sympodial units, indicating a delay in precocious termination. Colored bars indicate average leaf numbers within sympodial units with standard deviations. Statistical significance in B and C was tested by Wilcoxon rank sum test, and significance levels are indicated by asterisks (*P<0.05, **P<0.01, ***P<0.001).
doi:10.1371/journal.pgen.1004043.g001

tantly, delays in flowering time and sympodial termination were also observed on side shoots (Figure S1A–C), indicating a whole plant effect from *sft*/+ heterozygosity that explains the increase in total inflorescence number (Figure S1D) [25]. Thus, postponement of sympodial termination in *sp* mutants from *sft*/+ heterozygosity is based on recurring weak delays of all main and side shoot sympodial flowering transitions.

sft/+ heterozygosity weakly delays the primary flowering transition

Initiation and perpetuation of tomato sympodial growth depends on a gradual flowering transition culminating in PSM termination in a process mediated in part by accumulating florigen product from *SFT* counterbalancing repressive signals from *SP*. Regardless of whether *SP* is mutated, mutations in *SFT* cause late

flowering and produce vegetative inflorescences, and strong alleles fail to initiate sympodial growth (Figure 2A) [27]. Our observation that precocious sympodial termination was delayed in *sft*/+ *sp* plants beginning with the first sympodial shoot (Figure 1C) led us to ask whether the flowering delay might commence in the PSM where *sft* homozygous mutant phenotypes first manifest. Surprisingly, whereas flowering time of *sft*/+ heterozygotes alone was not significantly different from *sp* mutants and WT, *sft*/+ *sp* plants were slightly later flowering (Figure 2A). We pinpointed this weak semi-dominant effect more precisely by evaluating developmental progression (ontogeny) of meristems. Like vegetative shoots, multi-flowered inflorescences of tomato are based on sympodial growth [26]. Just before the PSM transitions to a terminal floral meristem (FM), a sympodial inflorescence meristem (SIM) initiates perpendicularly, and this process reiterates several times to produce the characteristic zigzag inflorescence [30]. At 20 days after germination (DAG), we quantified SIM production in the primary inflorescence and found that *sft*/+ *sp* plants were on average one SIM behind *sp* mutants (Figure 2B–D). At this same point, while the first SYM of *sp* plants had already given rise to the first or second FM-SIM pair of the second inflorescence, most *sft*/+ *sp* SYMs were still in the reproductive transition (no FM evident morphologically) or starting the development of the first SIM-FM pair (Figure 2E–G). Thus, having only one fully functional allele of *SFT* delays the flowering transitions of both primary and sympodial shoots in *sp* mutants.

sft/+ heterozygosity delays seedling development and primary shoot meristem maturation

Our developmental findings suggested that *sft*/+ overdominance and yield increases might commence with a semi-dominant delay of the primary flowering event. The flowering transition is paralleled by a maturation of seedlings marked by changes in morphological complexity and molecular states (e.g. transcriptomes) of leaves [27,31]. As leaves of *sft*/+ *sp* plants are indistinguishable from those of WT and *sp*, we captured global gene expression patterns of the 6th expanding (3 cm) leaf, which is when differences in meristem ontogeny first appear (Figure 2B–G, Figure 3A and Dataset S1). In comparing *sp* single and *sft sp* double mutant leaf mRNA-Seq generated transcriptomes with those of *sft*/+ *sp* plants, we found 838 differentially expressed genes among all genotypes. Previous studies comparing gene expression between hybrids and parents involved whole genome heterozygosity and reported thousands of differentially expressed genes representing all modes of gene action (e.g. dominant, recessive, additive, overdominant, etc.) [6,8,32]. Surprisingly, despite having heterozygosity at only a single gene in an otherwise homozygous background, we observed expression changes in all directions (Dataset S2). One possible explanation among many for this complexity is that *SFT* is involved in multiple feedback loops and regulates major signaling cascades [33]. However, our primary interest was not to classify and compare these expression differences to whole genome heterozygotes or to dissect transcriptional regulatory networks controlled by *SP* or *SFT*, but rather to use the RNA-Seq data as a quantitative molecular phenotyping tool to determine if there are changes in seedling maturation caused by *sft*/+ heterozygosity before gross morphological differences in shoot architecture become apparent.

The Digital Differentiation Index (DDI) algorithm identifies transcriptional marker genes whose expressions peak at chosen reference stages to identify stage-enriched marker genes and then queries these marker genes from transcriptomes of "unknown" tissues to predict their maturation states relative to the references

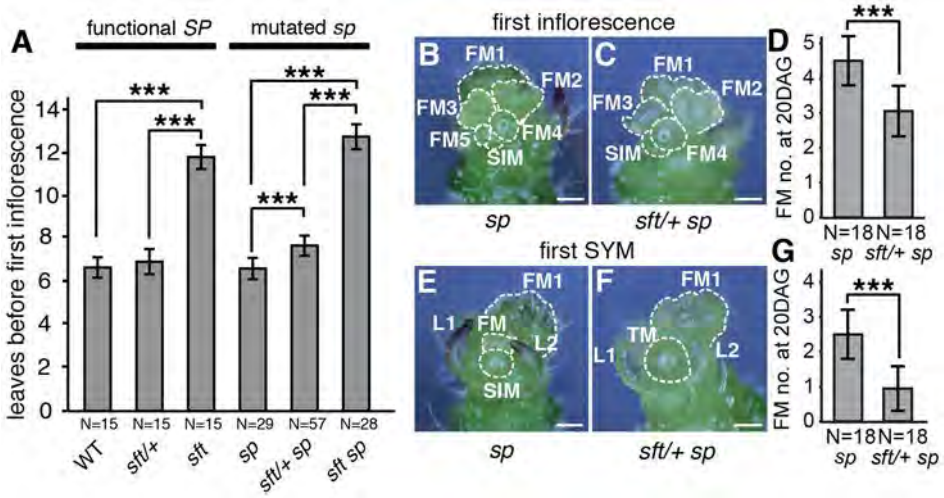


Figure 2. *sft*/+ heterozygosity induces weak semi-dominant delays in both primary and sympodial flowering transitions. (A) *sft*/+ *sp* plants show slightly delayed primary shoot flowering time compared to *sp* as measured by leaf production before formation of the first inflorescence. Note the extremely delayed flowering of *sft sp* double mutants, indicating a weak semi-dominant effect for *sft*/+ heterozygosity. Bars indicate average leaf numbers with standard deviations. Genotypes and sample sizes are shown below. Statistical differences were tested by Wilcoxon rank sum tests and significance levels are marked by asterisks (***P<0.001). (B–G) Representative images and quantification of developmental progression (ontogeny) of meristems in the first inflorescence and sympodial shoot meristems (SYM) of *sp* (left images) and *sft*/+ *sp* plants (right images) at 20th DAG. Both *sp* (B) and *sft*/+ *sp* (C) PSMs have completed the primary flowering transition and generated a series of floral meristems (FM) and sympodial inflorescence meristems (SIM) [26,30]. *sft*/+ *sp* plants are consistently one SIM behind ontogenically, consistent with a weak delay in flowering from *sft*/+ heterozygosity (D). Developmental progression of the first SYM in *sp* (E) and *sft*/ + *sp* (F) plants at the same time point as in B–C. While the SYM of *sp* mutants has already completed the flowering transition and differentiated into the first or second FM and initiated the next SIM, the SYM of *sft*/+ *sp* plants is still transitioning or initiating the first SIM, indicating a developmental delay parallel to the PSM of *sft*/+ *sp* plants (G). In D and G, bars indicate average numbers of initiated FMs with standard deviations. Genotypes and sample sizes are shown below. Statistical differences were tested by Wilcoxon rank sum tests and significance levels are marked by asterisks (***P<0.001). Scale bar: 100 um.
doi:10.1371/journal.pgen.1004043.g002

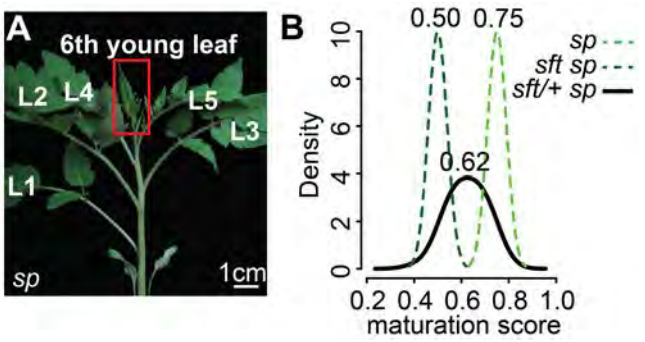


Figure 3. Transcriptome profiling reveals an early semi-dominant delay on seedling development from *sft*/+ heterozygosity. (A) Representative 6th expanding leaf from *sp* mutants. The same leaf and stage (3 cm long) was profiled by RNA-Seq for *sft*/+ *sp* and *sft sp* genotypes. (B) Molecular quantification of leaf maturation using the DDI algorithm [31]. Given that seedling development of *sft sp* is delayed compared to *sp* based on extreme late flowering, the *sft sp* 6th expanding leaf was designated an early leaf calibration point. Dark and light green curves indicate *sft sp* and *sp* maturation score distributions based on 124 DDI-defined marker genes. The black curve for the *sft*/+ *sp* 6th leaf indicates an intermediate maturation state. Numbers above indicate average maturation scores.
doi:10.1371/journal.pgen.1004043.g003

[31]. DDI revealed that *sft*/+ *sp* 6th leaf maturity was in between *sft sp* and *sp*, indicating that *sft*/+ heterozygosity delays maturation of *sp* plants already as young seedlings (Figure 3B, Dataset S3). We next asked whether the change in *SFT* dosage might be sensed in the PSM before it transitioned to flowering. We previously captured and quantified transcriptomes of five developmental stages of PSM maturation, which revealed a meristem maturation clock underlies a gradual transition of the PSM to a reproductive state [34]. The transition meristem (TM) stage of this clock is marked by increasing expression of flowering transition genes [34], and we therefore chose this stage for molecular phenotyping and comparison (Dataset S2 and S3). Importantly, TMs can be collected at precisely matched ontogenetic points, defined by initiation of the last leaf and indistinguishable meristem morphologies of tall round domes (Figure 4A–C) [34]. As expected based on the primary inflorescence of *sft* mutants reverting into a vegetative shoot, and consistent with *sft* epistatic over *sp*, DDI revealed that the TM of *sft sp* double mutants exhibited a severely delayed maturation, most closely matching a vegetative meristem state (Figure 4D). In contrast, whereas *sp* TM maturity was indistinguishable from WT, the *sft*/+ *sp* TM was delayed relative to *sp* and therefore intermediate between *sp* single and *sft sp* double mutants (Figure 4D). Importantly, we also profiled the first SYM from *sp* and *sft*/+ *sp* plants (*sft sp* plants fail to form a SYM) (Figure 4E and F), and found that, like in the PSM, the *sft*/+ *sp* SYM was also delayed relative to *sp* (Figure 4G). Altogether, these expression data suggest an early semi-dominant effect on the PSM flowering transition is the triggering event for *sft*/+ yield increases, and that all subsequently formed vegetative meristems in *sp* plants become equally sensitive to reduced dosage of *SFT* as they transition to a reproductive state.

Suppression of *SFT* by artificial microRNA phenocopies the dosage effects of *sft*/+ heterozygosity

Our findings that *sft* single gene overdominance traced back to cumulative delays on recurring flowering transitions led us to reason that the dosage effects of *sft*/+ heterozygosity might be recapitulated by simply partially reducing levels of functional *SFT* transcripts. We tested this by over-expressing artificial microRNAs against *SFT* (*35S::amirSFT*) in the *sp* background [35,36]. In addition to *SFT*, the artificial microRNAs were designed to target the *Arabidopsis thaliana SFT* ortholog, *FLOWERING LOCUS T* (*FT*), to assess their broad efficacy, and were incorporated into two different *Arabidopsis* pre-microRNA templates, *At* pre-mir164b and *At* pre-mir319a, to guard against differential amir backbone efficiencies (Figure 5A). In *Arabidopsis*, *35S:amiR-SFT/FT*^{At164b} and *35S:amiR-SFT/FT*^{At319a} transformants exhibited late flowering phenotypes equivalent to *ft* mutants (Figure S2B–C). In tomato, six of eight first generation (T1) transformants showed *sp* suppression phenotypes, and we selected three lines representing the range of observed suppression for further analysis. *SFT* transcript abundance was evaluated in these lines by quantitative RT-PCR, revealing a range of knockdown levels by the artificial microRNAs (Figure 5B). We evaluated progenies from two *35S:amiR-SFT/FT*^{At164b} (referred to as *amirSFTa* and *amirSFTb*) and one *35S:amiR-SFT/FT* ^{At319a} (referred to as *amirSFTc*) transformants, and found that the *amirSFTa* produced an average of one additional sympodial unit and inflorescence compared to non-transformed *sp* mutants, closely resembling the dosage effects of *sft*/+ heterozygosity (Figure 5C). *amirSFTc* showed greater suppression, terminating sympodial growth after producing often more than two additional units, while *amirSFTb* fully suppressed *sp* to indeterminacy like WT plants (Figure 5C). Notably, the level of suppression of *sp* determinacy corresponded with the level of knockdown of *SFT*; e.g. the indeterminate line, *amirSFTb*, showed the greatest reduction of *SFT* transcripts (Figure 5B–C). In all six lines, we failed to find strong *sft sp* double mutant phenotypes of reverted inflorescences or loss of sympodial growth, suggesting only weak alleles of *SFT* were created with the *35S::amirSFT* transgene – an effect that is also consistent with often observed weak target knockdown by artificial microRNAs [35,36]. Importantly, we found delayed flowering time in successive sympodial units like in *sft*/+ *sp* heterozygotes, and all three *amirSFT* progeny populations exhibited delayed primary shoot flowering time (Figure 5D). Thus, tuning *SFT* dosage transgenically mimics the effects of *sft*/+ heterozygosity, further illustrating that a classical epistasis relationship between the *sft* and *sp* mutants is ultimately responsible for the overdominant effect on yield.

A dosage effect from florigen mutant heterozygosity is conserved in *Arabidopsis*, but does not cause heterosis

As florigen is a universal inductive signal for flowering that several flowering pathways converge upon [37,38], we wondered if and how florigen mutant heterozygosity in a different system might affect growth, and specifically whether heterosis would result. We tested this by creating orthologous mutant combinations in *Arabidopsis thaliana*, which is a monopodial plant in which a single flowering event converts the SAM into a continuously growing inflorescence meristem (IM) that produces flowers laterally, in contrast to the tomato sympodial growth habit in which multiple flowering transitions occur. Despite this difference, *Arabidopsis ft* (*sft*) mutants are likewise late flowering [39] and completely epistatic over the early flowering and precocious termination of inflorescence meristems of *tfl* (*sp*) mutants [40]. To evaluate potential dosage effects of *ft*/+ heterozygosity, we phenotyped progeny from *ft-2*/+ *tfl1-2* plants, in which the *ft-2* mutation, a strong allele, segregates in the *tfl1* background (Figure 6A). We measured flowering time by counting rosette leaves and found a clear dosage effect in *ft-2*/+ *tfl1* plants compared to *tfl1* single and *ft-2 tfl1* double mutants (Figure S3A). We next tested for heterosis by quantifying yield related traits, including plant height, number of axillary shoots, and, as a parallel

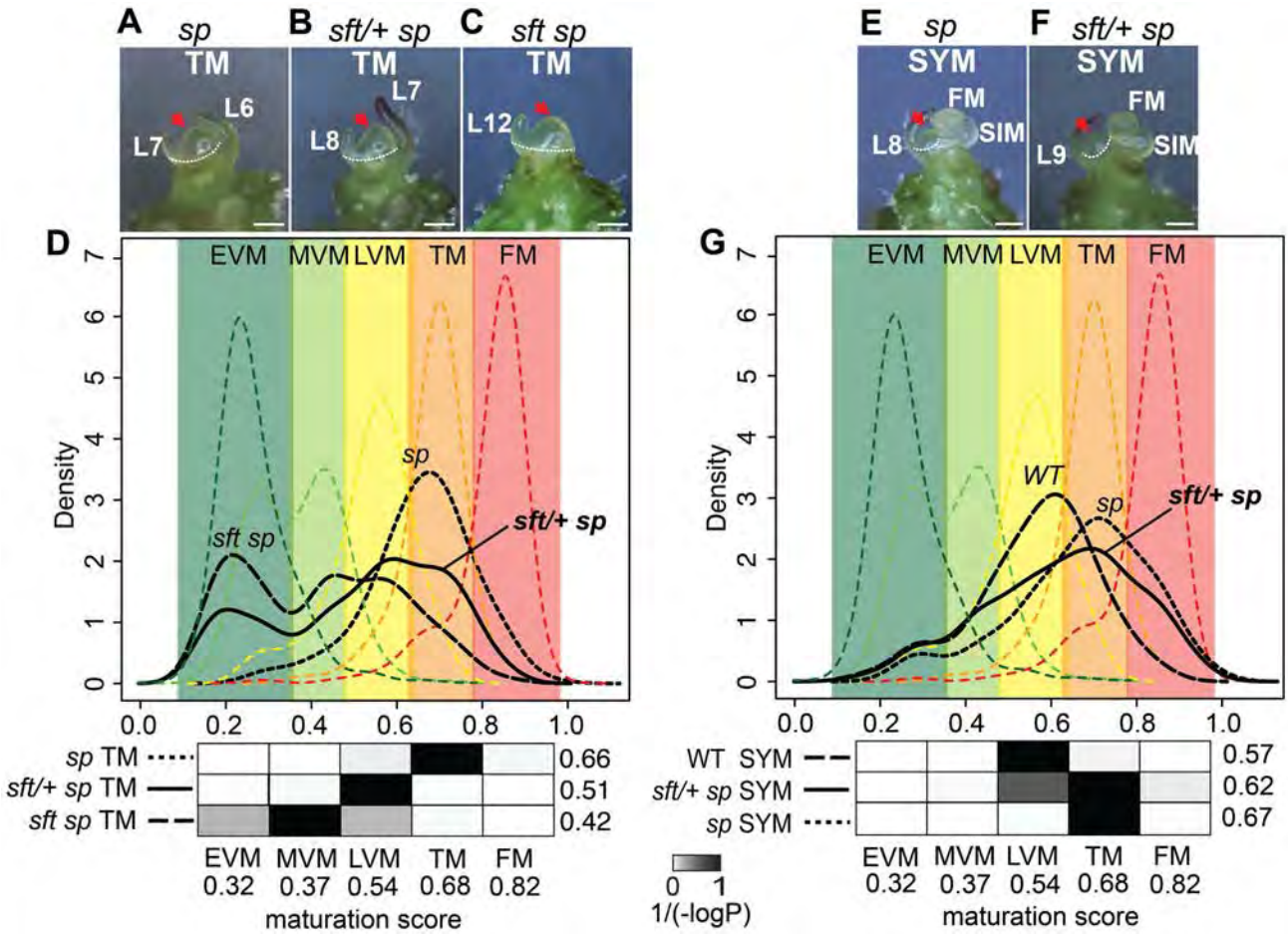


Figure 4. Transcriptome profiling reveals a semi-dominant delay in meristem maturation from *sft*/+ heterozygosity. (A–C) Stereoscope images showing morphology and dissection (white dashed line) of the TM stage used for mRNA-Seq from *sp* (A), *sft*/+ *sp* (B) and *sft sp* (C) genotypes. Scale bar: 100 um. Red arrows highlight identical TM morphologies. L: leaf primordium number. The additional leaf primordium at the *sft*/+ *sp* TM is consistent with the one leaf delay in primary shoot flowering time (Figure 2). (D) DDI quantification of maturation scores for *sp*, *sft sp*, and *sft*/+ *sp* predicted from the WT PSM meristem maturation atlas [34]. Colored dashed curves indicate maturation stages for the 5 PSM stages used for calibration EVM, MVM, LVM, TM and FM [the Early, Middle, and Late Vegetative Meristems, Transition Meristem and the Flower Meristem]. Colored areas define boundaries of these stages estimated from the curves. Maturation scores are derived from 637 DDI-selected marker genes (Dataset S3). Student's t-tests are presented as heat-maps of scaled $1/(-\log_{10}P)$ values below each graph, and associated numbers to the right indicate average maturation scores for the predicted meristems. Darker color indicates greater similarity in maturation state. Note the statistically intermediate TM maturation state of *sft*/+ *sp* relative to *sft sp* and *sp*, indicating *sft*/+ heterozygosity causes a semi-dominant delay in the primary flowering transition. The presence of more than one peak along the curves of the *sft sp* and *sft*/+ *sp* genotypes reflect mixed maturation states for these TMs, as different subsets of marker genes are driving different maturation stage estimates that translate to less uniform maturation patterns. (E–F) Stereoscope images showing morphology and dissection of the first sympodial shoot meristem (SYM) used for mRNA-Seq profiling in *sp* (E) and *sft*/+ *sp* genotypes (F). Meristems and leaf primordia are marked as in Figure 2. (G) DDI quantification of SYM maturation scores from *sp*, *sft*/+ *sp*, and WT using the PSM stages as calibrations. Maturation scores for *sft*/+ *sp*, *sp* and WT indicate an intermediate maturation state for the SYM of *sft*/+ *sp* plants, mirroring the delay in the PSM. P-value heat maps are shown below along with average maturation scores to the right.
doi:10.1371/journal.pgen.1004043.g004

to tomato yield, the number of siliques, flowers, and flower buds (Figure 6B–C and Figure S3B–D). Surprisingly, *ft*/+ *tfl* plants showed semi-dominance for plant height and total yield (Figure 6B and C), and similar effects were observed for a moderate second allele of *ft* (Figure S3E). Thus, whereas the dosage effect on flowering time from florigen mutant heterozygosity is conserved in the monopodial growth habit of *Arabidopsis*, it does not translate to heterosis.

Discussion

Crop yields derive from a complex integration of fitness-related traits founded on developmental and physiological mechanisms for organ production and biomass accumulation. Thus, studying heterosis inevitably involves a broad analysis of the myriad mechanisms controlling plant growth. It is therefore perhaps not surprising that recent gathering of vast genetic, phenotypic, and molecular data on cases of heterosis from diverse systems has suggested that multiple non-mutually exclusive system-specific mechanisms are likely at work [8–10,41]. Looking at heterosis from the developmental perspective, it would be reasonable to assume *a priori* that flowering would have a major role given that selection of allelic variation for flowering time regulators has been a major contributor to adaptation, domestication, and maximizing crop yields through classical and modern breeding [42]. In rice, for example, alleles of strong effect from various flowering regulators, many showing epistatic interactions, were selected to enable growth at different climates and day lengths [43–45]. The

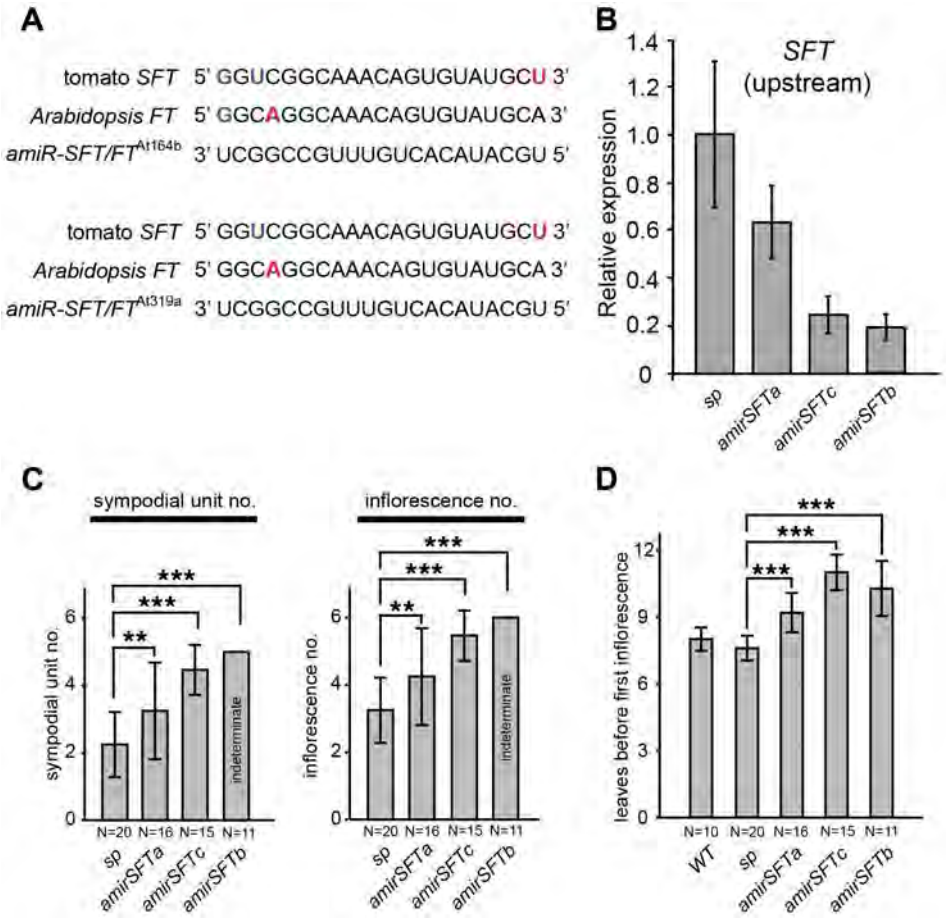


Figure 5. Reducing *SFT* transcripts with artificial microRNAs mimics the dosage effects of *sft*/+ heterozygosity. (A) Artificial microRNAs targeting tomato *SFT* and *Arabidopsis FT*. Shown are alignments of amiR-SFT/FT^{At164b} and amiR-SFT/FT^{At319a} with the complementary region of *SFT* and *FT*. G–U wobbles and mismatches between the two amiR-SFT/FTs and the target are highlighted in the target sequence with bold blue and red, respectively. (B) Quantitative RT-PCR measurements of tomato *SFT* transcript levels in *amirSFT* plants showing knock down. Results shown are from using primers targeting *SFT* transcripts 5′ to the amiRNA binding site, consistent with reports of primer-dependent transitivity occurring at the 3′ to 5′ direction upon the initial target cleavage, resulting in degradation of the 5′ cleaved product of the target but not the 3′ product [80,81] (Figure S2). Bars indicate relative expression level and error bars indicate standard deviation among replicates. (C) Depending on the strength of suppression, *amirSFT* plants produce at least one additional sympodial unit and inflorescence compared to *sp* alone, indicating that reducing *SFT* transcript levels by artificial microRNA partially suppresses *sp* sympodial termination, mimicking the dosage effect of *sft*/+ heterozygosity. Note that some *amirSFTc* progeny plants showed indeterminacy, whereas *amirSFTb* progeny plants were always indeterminate, indicating that a stronger suppression of *SFT* completely suppresses the *sp* phenotype and reverts the plants to normal sympodial cycling. Differences in sympodial unit and inflorescence numbers between *amirSFT* and *sp* plants were tested by Wilcoxon rank sum test and significance levels are marked by asterisks (* $P<0.05$, ** $P<0.01$, *** $P<0.001$). (D) *amirSFT* plants have delayed primary shoot flowering time compared to *sp* and WT controls, similar to *sft*/+ heterozygosity. Bars indicate average leaf numbers with standard deviations. Genotypes and sample sizes are shown below. Differences in leaf numbers between *amirSFT* and *sp* plants were tested by Wilcoxon rank sum test and significance levels are marked by asterisks (* $P<0.05$, ** $P<0.01$, *** $P<0.001$).
doi:10.1371/journal.pgen.1004043.g005

same was achieved in maize, but, instead, dozens of loci of small additive effect were found to be involved [46]. In both rice and maize, and as occurred during the domestication and breeding of many crops, this selection enabled a shift from an extended period of flowering in wild populations to uniform flowering, which provided sudden bursts of yield that facilitated agronomic practices, particularly harvesting [42]. Interestingly, the genetic path leading to high yielding tomatoes has differed from other major crops in that domestication has mostly acted on fruit size to increase yield with little evidence for selection on flowering [47–50]. Indeed, while there is certainly flowering time and architectural variation among distantly related wild tomato species [51], cultivated tomatoes and their wild progenitor, *S. pimpinellifolium*, share nearly identical flowering times and indeterminate growth habits, suggesting there was little or no standing genetic variation for artificial selection to act upon [52]. Only with the relatively recent discovery of *sp* did a change in flowering provide a major agronomic shift in how tomato was grown in the field, enabling a burst of flower production and yield on compact plants grown at high density, which gave rise to the processing tomato industry [26]. In this regard, in contrast to maize where altered flowering times are frequently observed in hybrids [10,53], cultivated tomato hybrids do not differ substantially from their parental inbreds for flowering time, inflorescence production, or overall plant architectures. Only upon introgressing quantitative trait loci (QTL) from distantly related wild species are heterotic effects on yield observed, a subset of which have been tied to changes in flowering and plant architecture, but the causative genes have not been identified [54]. Thus, our dissection of *sft* heterosis is the first to expose a direct link to flowering and resolve the underlying mechanism.

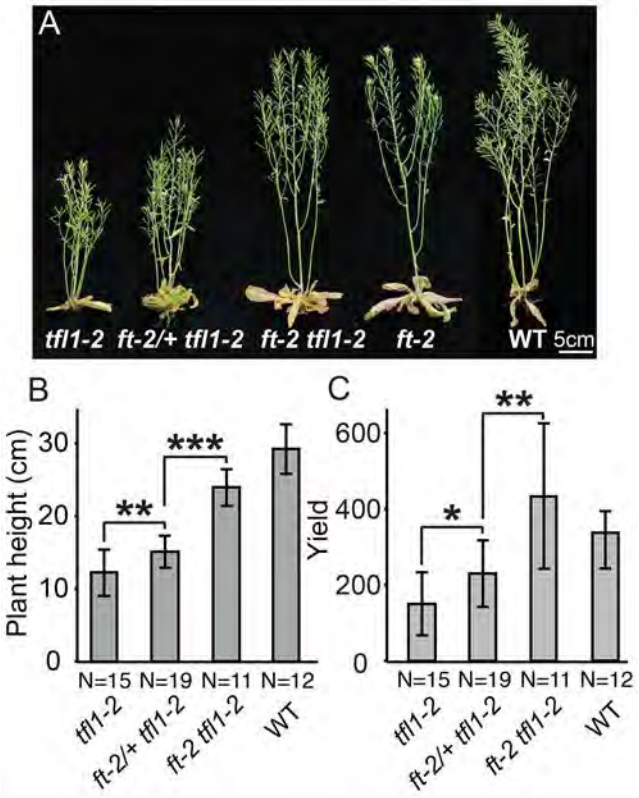


Figure 6. Dose-dependent suppression of *tfl1* (*sp*) by *ft*/+ (*sft*/+) heterozygosity is conserved in *Arabidopsis thaliana*. (A) Representative plants from left to right of: *tfl1-2* single mutants, *ft-2*/+ *tfl1-2*, *ft-2 tfl1-2* double mutants, *ft-2* single mutants and wild type *Ler-0* (WT) showing the intermediate height of *ft-2*/+ *tfl1-2* plants compared to *tfl1-2* and *ft-2 tfl1-2* genotypes. (B–C) Statistical comparisons among all genotypes for plant height and flower/fruit yield showing semi-dominant effects from *ft-2*/+heterozygosity in the *tfl1-2* background. Bars indicate average values with standard deviation. Genotypes and sample size are shown below. Differences between genotypes were tested by a Wilcoxon rank sum test and significance levels are marked by asterisks (*P<0.05, **P<0.01, ***P<0.001). (B) *ft-2* heterozygosity in a *tfl1-2* mutant background partially suppresses the early flowering and early termination phenotype of the *tfl1-2* mutation in a semi-dominant manner, resulting in plant height in between *tfl1-2* and *ft-2 tfl1-2* mutant parental lines. (C) Unlike tomato, *ft*/+ heterozygosity in a *tfl1-2* mutant background does not drive heterosis for yield (number of total siliques and floral buds) in *Arabidopsis*. Rather, yield in the *ft-2*/+ *tfl1-2* plants is intermediate to *tfl1-2* and *ft-2 tfl1-2* double mutants.
doi:10.1371/journal.pgen.1004043.g006

Our combined developmental and molecular phenotyping of *sft*/+ overdominance has exposed a novel principle for how tomato plant architecture and yields might be further optimized by taking advantage of the surprising and remarkable level of dosage sensitivity within florigen and the florigen pathway. The genetically induced reduction in dosage of florigen from *sft*/+ heterozygosity causes a slight delay in the transition to reproductive growth that, in the context of recurring flowering events of the sympodial habit and the *sp* mutant background, translates to cumulative overdominance. Indeed, this heterosis example, like many others [20–24,55], is conditional. Yet, it is this genetic and developmental conditionality that suggests *sft*/+ heterosis could be considered less about heterozygosity and heterosis *per se* and more about the potential to genetically fine-tune *SFT* expression levels to manipulate yield in a way that domestication and breeding efforts have not yet capitalized on, perhaps because standing allelic diversity for florigen and members of its pathway is limited. In this respect, we propose that additional directed quantitative manipulation of the relative doses of *SFT* to *SP* might enable further fine-tuning of flowering, sympodial cycling, and inflorescence production. For example, as yet undiscovered, or artificially created [56,57], transcriptional or loss-of-function alleles of *SFT* and *SP* of various strengths could be combined in different genetic constitutions to pinpoint an even higher optimum of plant architecture to maximize yield. In an even simpler scenario, homozygosity for very weak mutant alleles of *SFT* in a strong *sp* background, or homozygosity for weaker mutant alleles of *SP* alone, could potentially match or exceed fruit production of the *sft*/+ *sp* genotype. Finally, beyond tweaking *SP* and *SFT*, partial suppression of *sp* determinacy by generating mutations in other pathway genes, especially those encoding components of the florigen activating complex [58], could provide novel alleles and breeding germplasm that natural variation might not be able to provide.

Importantly, although there is tremendous diversity among angiosperms in when and where inflorescences and flowers form, the *SFT/SP* system is highly conserved [38,59,60], suggesting the aforementioned concepts could be applicable to other plants. Yet, our findings in *Arabidopsis* imply that while dosage effects on flowering time from florigen mutant heterozygosity will be broadly conserved, yield benefits might not be, and species-specific outcomes will likely trace back to differences in growth habits. The lack of meristem termination and recurring flowering events in the monopodial growth habit of *Arabidopsis* means that florigen mutant heterozygosity is sensed only once during development, and that no compounding of the semi-dominant dosage effect is possible. Indeed, increasing yield in *Arabidopsis* simply requires a larger plant, which can be achieved by delaying and prolonging flowering either environmentally through short day growth conditions or genetically through mutations in flowering regulators like *FT*. Consistent with this, we found that homozygous *ft* mutants were the highest yielding of all genotypes (Figure 6). At first glance, this would suggest limited possibilities for exploiting our findings beyond tomato; however, for some breeding goals, such as improving biomass, delaying flowering quantitatively and predictably through an allelic series of florigen mutants in either the homozygous or heterozygous condition could prove valuable to customize plant architecture and size for particular agronomic needs. Remarkably, yield benefits from heterozygous mutations in florigen orthologs have been found in at least one plant that lacks sympodial growth. In a strikingly similar example to tomato, a major domestication QTL for flowering time in sunflower traces back to a deletion in a duplicated paralogous *FT* gene that causes heterosis for both seed size and weight when heterozygous under short day conditions [61]. In another example, a classical report of overdominance for sorghum yield involves heterozygosity for an as yet uncharacterized late flowering mutant that has all the hallmarks of being defective in florigen or a florigen pathway component [62]. Thus, heterozygosity for florigen mutants holds potential for broadly improving crop yields, which, in hindsight, is perhaps not surprising given that selection for beneficial alleles of various strengths in florigen family genes, especially orthologs of *SFT* and *SP*, was key for the domestication of barley [63], beets [64], beans [65,66], grape [67], potatoes [68], roses [69], soybeans [70,71], sunflower [61], tobacco [72], and likely many other plants. With these examples in mind, and considering our findings in *Arabidopsis*, we suggest that *sft*/+ heterozygosity in a dose-dependent epistatic relationship with *sp* may represent only one of several ways to genetically tailor florigen levels, and that hunting for new alleles in existing germplasm or engineering custom alleles could allow an optimal fine-tuning of florigen and its pathway to

maximize flowering, inflorescence production, and other yield components in these and other crops. The potential to broadly manipulate agronomic traits by florigen and its family members in diverse plant species stems not only from roles in flowering time, but also as general coordinators of diverse physiological processes affecting multiple aspects of plant growth and fertility [38]. Thus, parallel to how mutations in biosynthesis genes for the hormone gibberellin created the dwarf mutants that propelled the Green Revolution [73], our findings provide compelling evidence that manipulating florigen family genes can provide a new path to meet current breeding challenges associated with a rapidly changing climate.

Materials and Methods

Tomato plant growth conditions, genotyping, and phenotyping

The *sp* mutant was first reported more than 80 years ago and arose spontaneously, and the strong *sft* mutant allele used in this study, *sft-7187*, was isolated from a fast neutron mutagenesis screen performed in tomato cultivar M82, and has a two nucleotide deletion that truncates the C-terminal portion of the protein [26,27,74]. All mutants were backcrossed to M82 at least four times to eliminate background mutations prior to the original yield trials [25]. For all experiments in this study, plants were grown in controlled greenhouse conditions at Cold Spring Harbor Laboratory. Greenhouses were supplemented with artificial light from high-pressure sodium bulbs (50 µmol/m2/sec; 16 h/8 h) and daytime temperature was 78°F and nighttime temperature was 65°F, with a relative humidity of 40–60%. Tomato F2 generation seeds derived from self fertilization of an *sft*/+ *sp* F1 plant were grown in 72-cell insert flats and transplanted after four weeks into 2 gallon pots (three plants per pot) for quantitative phenotyping. Young leaf tissue was collected from each F2 individual at the time of transplanting for DNA extraction and genotyping. Total genomic DNA was extracted using a standard cetrimonium bromide (CTAB) DNA extraction protocol. Genomic fragments of the *SFT* locus were amplified using the PCR primers: "sft-7187 full exon F2" 5′-GGGCAAGAAATAGTGAGCTAT-3′ and "sft-7187 full exon R2" 5′-TTCAAATAAATTGAGAGGAAGA-3′ and the following PCR program: initial denaturation at 94°C for 3 minutes, then 35 cycles at 94°C for 30 seconds, annealing at 52°C for 30 seconds, extension at 72°C for 1 minute, and a final extension at 72°C for 10 minutes. The PCR products were subjected to enzyme digestion with *Tse*I at 60°C for 6 hours, resulting in two bands for wild type, one band for *sft* mutant and three bands for *sft*/+ after running on a 3% agarose gel at 150 V for 40 minutes. The number of leaves in the primary shoot prior to the first inflorescence and leaves within three successive sympodial units were counted for each individual at 8–12 weeks after germination. This same phenotyping scheme was applied to two axillary shoots: the lower (basal) axillary shoot originating from the axil of the first leaf on the primary shoot and the uppermost (proximal) axillary shoot originating from the axil of the last leaf formed before the first inflorescence. Quantitative measurements for inflorescence number, sympodial unit number, primary and lateral shoot flowering time, and leaf number in three sympodial units were evaluated for the shape of each phenotype's distribution and subjected to two-tailed Wilcoxon rank sum tests between genotypes and Kruskal–Wallis one-way analysis of variance across all genotypes. To quantitatively compare the progression of sympodial inflorescence meristem (SIM) and floral meristem (FM) initiation on the first developing inflorescence of *sp* and *sft*/+ *sp* plants, we germinated 18 plants for both genotypes at the same time and counted the number of differentiated FMs on both primary and sympodial shoots at 20th days after germination (DAG). The FM numbers were subjected to two-tailed Wilcoxon rank sum tests between genotypes. To image live meristems, shoot apices were dissected from seedlings, and older leaf primordia (>150 µm) were removed under a Nikon SMZ1500 stereomicroscope. The meristem images were taken immediately after dissection with an integrated Nikon digital camera, recaptured by Z-series manually, and merged to create focused images.

Arabidopsis plant growth conditions, genotyping, and phenotyping

Arabidopsis thaliana plants were grown in the greenhouse under long day (16 h light, 8 hr dark) conditions in 32-cell flats with two plants per cell. Individual seeds were delivered to the corner of each cell to avoid growth competition during germination. The seeds were stratified at 4°C for 4 days before transferring to a long day greenhouse maintained at 21°C. All mutant lines were acquired from the Arabidopsis Biological Resource Center (ABRC) and originated from EMS mutagenesis in the Landsberg *erecta* (L*er*) background. Homozygous *tfl1-2* mutant plants were crossed to a moderate (*ft-1*) and strong (*ft-2*) allele of *ft*. Individual F1 plants from each cross were self-fertilized to generate F2 populations segregating for both *tfl1-2* and *ft* mutants. Plants homozygous for the *tfl1-2* mutation and heterozygous for the *ft-2* mutation were self-fertilized to generate F3 populations fixed for the *tfl1-2* mutation and segregating for the *ft-2* mutation. Tissue was harvested from young rosette leaves and DNA was extracted using a standard CTAB DNA extraction protocol. The *tfl1-2* and *ft-2* mutations were detected using derivative CAPS (dCAPS) assays. A fragment of *TFL1* was amplified by PCR using the primers "tfl1-2 dCAPS-F" 5′- AAACGTCTCACTTCCTTTTCCTC-3′ and "tfl1-2 dCAPs-R2" 5′- AAATGAAAAGAAAGAATAAATAAATTAAA**G**GTAC-3′ and a fragment of *FT* was amplified using "ft-2 dCAPS-F2" 5′- CCCTGCTACAACTGGAACAACCTTTGGTG-3′ and "ft-2 dCAPS-R2" 5′- AAACTCGCGAGTGTTGAAGTTCTGG**G**GC-3′. Both *TFL1* and *FT* fragments were amplified using a touchdown PCR program: initial denaturation at 95°C for 3 minutes, then 10 cycles at 95°C for 20 seconds, 65°C for 30 seconds (decreased by −0.5°C/cycle), 72°C for 30 seconds followed by an additional 30 cycles at 95°C for 20 seconds, 52°C for 30 seconds, 72°C for 30 seconds and ending with a final extension at 72°C for 10 minutes. Underlined nucleotides in the aforementioned sequences introduce a new restriction site in the wild type PCR amplicons. *TFL1* PCR amplicons were digested using *Kpn*I for 3 hours at 37°C, which cuts wild type but not mutant sequences. *FT* PCR amplicons were digested using *Hae*III for 3 hours at 37°C, which cuts wild type but not the *ft-2* mutant sequences. Wild type versus mutant banding patterns was resolved on a 3% half MetaPhor agarose-half regular agarose gel. Phenotyping was completed in the F3 generation, and we compared *tfl1-2 ft-2* double, *tfl1-2 ft-2*/+ and *tfl1-2* single mutants. Homozygous single mutants and wild type L*er*-0 were grown at the same time for comparison. Phenotyping and imaging was performed when the plants completed flowering and inflorescence meristems stopped growing (6–8 weeks after germination). The height of each plant was measured along the main shoot of the plant from where the base emerged from the rosette to top of the shoot. The number of rosette leaves, axillary shoots, siliques, open flowers, and floral buds were also recorded as measures of flowering time and yield. For each measured trait, the mean and standard deviation was calculated for each genotype. The means were compared using a

Student's t-test (Wilcoxon rank sum test when the phenotypic distribution was not normal).

Global gene expression profiling (mRNA-Seq) of tomato leaves and meristems

Tomato homozygous *sp* mutants, *sft sp* double mutants and F1 single gene heterozygotes of *sft*/+ *sp* plants were used for leaf and meristem expression profiling experiments. All *sft*/+ *sp* plants originated from F1 seeds of direct crosses between the *sp* and *sft sp* parents, and a subset of F1 plants were confirmed by PCR genotyping to ensure 100% *sft*/+ heterozygosity. Seeds were germinated in petri plates on water-soaked Whatman paper at 28°C for 72 hours until the root radicles emerged. The germinated seeds were then transplanted to 72-cell insert flats with pre-wet soil and placed in the greenhouse. The plants used for leaf expression profiling were transplanted to two-gallon pots (three plants per pot), and tissue from the 6th young expanding leaf from each plant was collected and immediately frozen in liquid Nitrogen when the leaves reached 3 cm in length. Total RNA was extracted using a Qiagen RNeasy mini total RNA extraction kit according to the manufacturer's protocol. Growth of seedlings for meristem expression profiling was monitored daily under a dissecting microscope using the meristem morphological cues marking previously defined maturation stages [34]. At the transition maturation (TM) stage, the cotyledons and leaves were removed from seedlings and the shoot apices with 3 cm hypocotyl attached were collected and stored in 100% acetone followed by vacuum infiltration for 30 minutes. Meristem tissue was dissected from the fixed stems using a surgical blade following the lines shown in Figure 4A–C and E–F under a dissecting microscope after confirming the morphology that marks the TM stage. Total RNA was extracted from the dissected meristem tissues with an Arcturus PicoPure total RNA extraction kit (Life Technologies). Except for the *sp* SYM, which is difficult to capture in high numbers because of a rapid termination, for all genotypes, tissue was harvested and prepared for mRNA-Seq construction for two biological replicates, and *sp* SYM was subjected to two technical replicates. As reported previously [34], two replicates were sufficient to quantify meristem maturation states using the DDI algorithm, which was our primary goal in the expression analysis.

RNA-Seq library preparation

For all tissues, poly-A containing mRNA was purified from total RNA using Invitrogen oligo-dT DynaBeads for mRNA-Seq library construction using the ScriptSeq v2 RNA library preparation kit (Epicentre). The maximum amount of mRNA input (50 ng) was used when possible to maximize the library output. The final PCR enrichment step was carried out following the standard protocol with 15 cycles and primers with barcode indices supplied by Epicentre to create barcoded mRNA-Seq libraries. The quantity and size distribution of each individual barcoded mRNA-Seq library was detected with a High Sensitivity DNA Chip on a Bioanalyzer 2100 machine (Agilent). The final concentration of each library was verified by qPCR using a KAPA library quantification kit and based on these results, four to six barcoded libraries were pooled together with equal concentration for one lane of Illumina paired-end (PE) 100 bp sequencing on an Illumina HiSeq sequencing machine (Dataset S1). All reads files were deposited to SGN (ftp://ftp.solgenomics.net/transcript_sequences/by_species/Solanum_lycopersicum/libraries/illumina/LippmanZ/) and the mean RPKM values of meristems are visualized on an eFP browser (http://tomatolab.cshl.edu/efp/cgi-bin/efpWeb.cgi, *SFT* heterosis panel).

Read mapping and analysis

All mRNA-Seq reads were trimmed to 50 bp to remove the bases with low qualities and mapped using Bowtie [75] to the tomato reference CDS [76] with paired-end relationships maintained. Trimming the reads to 50 bp also made the libraries comparable to our previous mRNA-Seq libraries [34] for combined DDI analyses. The lack of size selection step in the Epicentre ScriptSeq v2 mRNA-Seq library preparation protocol allowed lower initial mRNA input but produced a larger insert size range (150 bp~1000 bp), which lowered the successful mapping with proper distance between paired-end reads. Mapping to predicted CDS also reduced the mapping rate due to failed mapping of reads coming from 5′ and 3′ UTR regions. However, the higher total read number from Illumina HiSeq compensated for the relatively lower mapping rates, yielding comparable mapped read numbers and sequencing depth to previous mRNA-Seq libraries that allowed for differential expression analysis and molecular phenotyping by DDI [34]. The resulting bam alignments were sorted and indexed by SAMtools [77], and the number of reads mapped to each CDS was counted to calculate the raw counts for all libraries. The raw counts from leaf and TM tissues across three genotypes were normalized using the TMM method. The distribution of gene expression levels were modeled following a negative-binomial distribution and tag-wise dispersion were estimated based on two replicates. Finally, exact tests for differential expression were conducted based on the replicates in pairwise comparisons. All normalization and differential expression tests were conducted using the edgeR package [78,79]. Although only two replicates were performed, we classified gene expression patterns from comparing *sft*/+ *sp* heterozygotes and homozygous parents into 12 categories belonging to five major classes: additive, recessive, dominant, overdominant and underdominant (Dataset S2) using a threshold of two-fold change and P-value$<=0.01$. Numbers of genes in each category were counted and their proportions in each category relative to all differential expressed genes were calculated for the 6th young leaf and TM, respectively, revealing all categories of gene expression changes were detected (Dataset S2).

Digital Differentiation Index analyses

Raw counts for the leaf expression profiles (including *sp*, *sft*/+*sp* and *sft sp* 6th young leaves) were incorporated into a master leaf data set. Raw counts for the meristem expression profiles (including *sp* and *sft*/+*sp* TM and SYM) were incorporated into a master meristem data set that includes all raw counts from our previous meristems profiling experiments [34]. For both master data sets, all raw counts were then summarized over replicates and normalized against number of mapped reads and CDS lengths to calculate RPKM values for DDI analyses [31]. DDI selects samples with known or pre-determined maturation states in the whole data set as calibration points, and then identifies marker genes that show maximum expression at each calibration point. These genes characterize the calibration points molecularly. DDI checks the marker gene expressions in the samples that are submitted to query (the 'unknown' samples) and quantifies the 'unknown' samples' maturation states relative to the calibration points. For each marker gene, DDI compares expression levels between 'unknown' samples and each calibration point and calculates a 'maturation score'. Collectively, all marker genes generate a distribution of maturation scores for the 'unknown' sample [31]. Importantly, curves showing multiple 'peaks' reflect a mixed molecular maturation state for the queried tissue, as different marker genes give different maturation estimates. This is most evident in *sft sp* double mutants that still transition to

flowering, but at a much slower rate compared to wild type and with vegetative reversion of the inflorescence, indicative of a mixed vegetative-reproductive state. At the same time, a Student's t-test of average maturation score difference between calibration and unknown samples was conducted for each unknown meristem sample, yielding a P-value for the significance of the maturation state difference. For each prediction, this P-value was obtained for comparisons between the unknown sample and temporarily successive calibration points, in order to generate a 'gradient' of meristem similarity (plotted in heat-maps in the form of scaled 1/(−log10P)). For example, to predict the maturation state of *sft*/+ *sp* SYM using the first replicate of WT EVM, MVM, LVM, TM and FM [the Early, Middle, and Late Vegetative Meristems (EVM: 5th leaf initiated; MVM: 6th leaf initiated; LVM: 7th leaf initiated), the Transition Meristem (TM: 8th leaf initiated), and the Flower Meristem (FM)] as calibration points, P-values were calculated for maturation state comparisons SYM vs. EVM, SYM vs. MVM, SYM vs. LVM, SYM vs. TM and SYM vs. FM, respectively. The P-values were then transformed into 1/(−log10P) and scaled across five values into a zero to one range (scaling was done for each prediction independently). Because smaller P-values indicate larger differences in maturation scores, the scaled 1/(−log10P) values quantify the relative similarity of the *sft*/+ *sp* SYM to each of the five calibration points. With the master leaf data set, DDI analyses were conducted using *sft sp* and *sp* 6th young leaves as two calibration points to predict maturation stages of *sft*/+*sp* leaf maturation. With the master meristem data set, DDI analyses were conducted using five WT primary shoot meristem (PSM) stages as calibration points to predict maturation stages of *sp*, *sft*/+*sp* and *sft sp* meristems. As in [34], one replicate of calibration samples was used for marker gene identification (Dataset S3), a second replicate of calibration samples treated as unknowns was predicted and plotted to set the boundaries of maturation stages (colored curves and boxes in Figure 4D and Figure 4G), and averaged RPKM values of predicting leaves and meristems were used to generate and plot the predicted distribution of maturation scores. All parameters for DDI analyses were as previously described [34]. All DDI analyses were carried out using modified R scripts as described previously [34].

Artificial microRNA construction and transformation

Artificial microRNAs were designed to repress both tomato *SFT* and *Arabidopsis FT* with two different backbones (Figure S1) [35]. The artificial microRNA amiR-SFT/FT^{At164b} and amiR-SFT/FT^{At319a} were synthesized by DNA2.0 and Bio S&T, respectively, and transformed into both tomato and *Arabidopsis* plants and phenotyped for repression of *SFT* and *FT*, respectively (Figure 5B, Figure S2). Tomato plants carrying mirSFT transgenes were measured for sympodial unit and inflorescence number, and phenotyping stopped after counting five or more sympodial units with two or more leaves in each unit and classified as indeterminate. The means of phenotypes were compared using a Student's t-test (Wilcoxon rank sum test when phenotype distribution is not a normal distribution).

For quantitative RT-PCR of *SFT* transcript abundance in the amirRNA lines, cotyledon tissue was collected from two-week old seedlings for total RNA extraction with Qiagen RNeasy mini total RNA extraction kit including DNase treatment with RNase-free DNase (Qiagen) according to the manufacturer's instructions. First-strand cDNA was then synthesized using the SuperScript III First-Strand Synthesis System with oligo dT (Invitrogen). *Ubiquitin* mRNA (Solyc01g056940) was used as the reference for normalization in quantifying cDNA. 5′ mRNA (upstream, Figure 5B) and 3′ mRNA (downstream, Figure S2A) of *SFT* transcript (Solyc03g063100) from the amirSFT binding site were quantified with 1 ul of cDNA using Phusion High-fidelity DNA polymerase (NEB), iQTM SYBR Green Supermix (Bio-Rad). A loss of transcripts was detected 5′ to the amiRNA binding site, consistent with reports of primer-dependent transitivity occurring at the 3′ to 5′ direction upon the initial target cleavage, resulting in degradation of the 5′ cleaved product of the target but not the 3′ product [80,81]. Primers pairs used were: 5′-CGTGGTGGTGCTAAGAAGAG-3′ and 5′- ACGAAGCCTCTGAACCTTTC-3′ for *Ubiquitin* (*UBI*); 5′-GCTTAGGCCTTCCCAAGTTA-3′ and 5′-GGGTCCACCATAACCAAAGT-3′ for *5′ mSFT* (upstream); 5′-GACAATTAGGTCGGCAAACA-3′ and 5′-AGCAGCAACAGGTAAACCAA- 3′ for *3′mSFT* (downstream). Two biological replicates of qRT-PCR were performed on the CFX96TM Real-time PCR System (Bio-Rad). qRT-PCR data were calculated from the number of PCR cycles needed to reach the linear phase for each *SFT* transcript from amirSFT lines and normalized against *Ubiquitin* using the qbase PLUS Data-Analysis Software.

Supporting Information

Dataset S1 Design of the mRNA-Seq expression profiling experiments, including genotypes, tissues, replicates, total read numbers and mapping rates.
(XLSX)

Dataset S2 Global gene expression profiling from two tissue types, 6th young expanding leaf and TM, grouped as percentages of differentially expressed genes in 12 possible gene action categories when comparing *sp*, *sft*/+ *sp* and *sft sp*. There are five major classes of gene action: additive (semi-dominant), dominant, recessive, overdominant and underdominant. Subcategories for each major class of gene action are represented by cartoon bar graphs. The first sheet shows the summary statistics of classification in two tissues and results of Fisher's exact tests for significant differences between the percentages in each gene action category. The following sheets show detailed information of the genes, including gene IDs, mean RPKM values, log fold changes for three pairwise comparisons, and P-values from differential expression tests. The 12 gene expression categories are classified based on a threshold of two-fold change and P-value$<=0.01$ between genotypes. All possible modes of gene action were observed in both tissues.
(XLSX)

Dataset S3 Marker genes selected by DDI and used in maturation score estimations all meristem DDI analyses involving the 6th expanding leaf, TM stage, and SYM stage. Included are gene IDs and functional annotations from tomato gene annotation iTAG version 2.3 [76].
(XLSX)

Figure S1 *sft*/+ mutant heterozygosity delays precocious axillary shoot termination in determinate tomato. (A) Compared to *sp* mutants, *sft*/+ *sp* plants show delayed primary flowering time on both basal and proximal axillary shoots similar to the main shoot (Figure 2A). Although no statistically significant ($P=0.11$), there is a trend towards a delay on the proximal lateral shoots of *sft*/+ *sp* plants (B) *sft*/+ *sp* plants produce more sympodial units before sympodial cycling terminates on both basal and proximal axillary shoots, similar to the main shoot (Figure 1B). (C) On both axillary shoots, *sft*/+ *sp* plants produce more leaves in the first three sympodial units, indicating a delay in precocious termination similar to the main shoot (Figure 1C). (D) Compared to *sp* mutants, *sft*/+ *sp* plants produce more inflorescences on each plant.

Genotypes and sample sizes are shown below, and error bars indicate standard deviations of averages. Statistical significance was tested by Wilcoxon rank sum test, and significance levels are indicated by asterisks (*P<0.05; **P<0.01; ***P<0.001).
(TIF)

Figure S2 Artificial microRNAs (amiRNA) targeting the *SFT* and *FT* genes. (A) Quantitative RT-PCR measurements of tomato *SFT* transcript levels using primers targeting 3′ to the amiRNA binding site. Note that transcript levels show little or no reduction compared to 5′ of the amiRNA binding site (Figure 5B), consistent with reports of primer-dependent transitivity occurring at the 3′ to 5′ direction upon the initial target cleavage, resulting in degradation of the 5′ cleaved product of the target but not the 3′ product [80,81]. Bars indicate relative expression level and error bars indicate standard deviation among replicates. (B) The *At* pre-amiR-SFT/FT ^{At164b} and pre-amiR-SFT/FT ^{At319a} sequences that were introduced into the plants along with theoretical representations of the RNA secondary structure. The fold-back structure in each of the sequences is emboldened and the miRNA sequence is highlighted. (C) 43-day old, long day (18 hours daylight, six hours night) grown *Arabidopsis thaliana* (Landsberg *erecta*) demonstrating the phenotypic effect of *amiR-SFT/FT* ^{At164b} and *amiR-SFT/FT* ^{At319a} on *FT* activity and flowering. *35S:amiR-SFT/FT* ^{At164b} and *35S:amiR-SFT/FT* ^{At319a} transformants exhibit delayed flowering equivalent to *ft* mutant plants.
(TIF)

Figure S3 Dose-dependent suppression of *Arabidopsis thaliana tfl1* mutant flowering time and yield-associated traits when either strong or moderate mutant alleles of *ft* are heterozygous. (A–D) Statistic analyses of *Arabidopsis* phenotypes caused by *ft-2*/+ heterozygosity in the *tfl1-2* mutant background. Bars indicate average values with standard deviation. Genotypes and sample size are shown below. Statistical significance was tested by Wilcoxon rank sum test, and significance levels are indicated by asterisks (*P<0.05; **P<0.01; ***P<0.001). (A) Total number of rosette leaves; (B) Total number of axillary shoots; (C) Total number of siliques; (D) Total number of floral buds; Note that number of rosette leaves and siliques showed semi-dominance caused by *ft*/+ heterozygosity. (E) Representative plants from left to right of wild type Ler-0 (WT), *tfl1-2* single mutants, *ft-1*/+ *tfl1-2*, and *ft-1* single mutants. Like for *ft-2*, *ft-1* mutants are completely epistatic over *tfl1-2* mutants, and therefore *ft tfl* double mutants (not shown) are not significantly different from *ft* single mutants (Figure 6).
(TIF)

Acknowledgments

We thank Dani Zamir, Eliezer Lifschitz, and Yuval Eshed for discussions spanning yield, heterosis, florigen, and plant development. We thank Oron Gar for comments on text related to tomato agronomic practices. We also thank Yuval Eshed for comments on the manuscript, and Tim Mulligan and Paul Hanlon for plant care.

Author Contributions

Conceived and designed the experiments: KJ KLL SJP JPA ZBL. Performed the experiments: KJ KLL SJP JPA. Analyzed the data: KJ KLL SJP JPA. Wrote the paper: KJ KLL ZBL.

References

1. Darwin C (1868) The Variation of Animals and Plants under Domestication. 2.
2. Charlesworth D, Willis JH (2009) The genetics of inbreeding depression. Nat Rev Genet 10: 783–796.
3. Shull GH (1908) The composition of a field of maize. Am Breed Assn Rep 4: 269–301.
4. Crow JF (1948) Alternative Hypotheses of Hybrid Vigor. Genetics 33: 477–487.
5. Hochholdinger F, Hoecker N (2007) Towards the molecular basis of heterosis. Trends Plant Sci 12: 427–432.
6. Lippman ZB, Zamir D (2007) Heterosis: revisiting the magic. Trends Genet 23: 60–66.
7. Springer NM, Stupar RM (2007) Allelic variation and heterosis in maize: how do two halves make more than a whole? Genome Res 17: 264–275.
8. Chen ZJ (2013) Genomic and epigenetic insights into the molecular bases of heterosis. Nat Rev Genet 14: 471–482.
9. Schnable PS, Springer NM (2013) Progress toward understanding heterosis in crop plants. Annu Rev Plant Biol 64: 71–88.
10. Birchler JA, Yao H, Chudalayandi S, Vaiman D, Veitia RA (2010) Heterosis. The Plant cell 22: 2105–2112.
11. McMullen MD, Kresovich S, Villeda HS, Bradbury P, Li H, et al. (2009) Genetic properties of the maize nested association mapping population. Science 325: 737–740.
12. Stuber CW, Lincoln SE, Wolff DW, Helentjaris T, Lander ES (1992) Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. Genetics 132: 823–839.
13. Li ZK, Luo LJ, Mei HW, Wang DL, Shu QY, et al. (2001) Overdominant epistatic loci are the primary genetic basis of inbreeding depression and heterosis in rice. I. Biomass and grain yield. Genetics 158: 1737–1753.
14. Luo LJ, Li ZK, Mei HW, Shu QY, Tabien R, et al. (2001) Overdominant epistatic loci are the primary genetic basis of inbreeding depression and heterosis in rice. II. Grain yield components. Genetics 158: 1755–1771.
15. Stupar RM, Springer NM (2006) Cis-transcriptional variation in maize inbred lines B73 and Mo17 leads to additive expression patterns in the F1 hybrid. Genetics 173: 2199–2210.
16. Ishikawa A (2009) Mapping an overdominant quantitative trait locus for heterosis of body weight in mice. J Hered 100: 501–504.
17. Schuler JF (1954) Natural Mutations in Inbred Lines of Maize and Their Heterotic Effect. I. Comparison of Parent, Mutant and Their F(1) Hybrid in a Highly Inbred Background. Genetics 39: 908–922.
18. Mukai T, Burdick AB (1959) Single Gene Heterosis Associated with a Second Chromosome Recessive Lethal in Drosophila Melanogaster. Genetics 44: 211–232.
19. Redei GP (1962) Single Locus Heterosis. Zeitschrift Fur Vererbungslehre 93: 164–&.
20. Efron Y (1973) Specific differences in maize alcohol dehydrogenase: possible explanation of heterosis at the molecular level. Nat New Biol 241: 41–42.
21. Hall JG, Wills C (1987) Conditional overdominance at an alcohol dehydrogenase locus in yeast. Genetics 117: 421–427.
22. Grobet L, Martin LJ, Poncelet D, Pirottin D, Brouwers B, et al. (1997) A deletion in the bovine myostatin gene causes the double-muscled phenotype in cattle. Nat Genet 17: 71–74.
23. Mosher DS, Quignon P, Bustamante CD, Sutter NB, Mellersh CS, et al. (2007) A mutation in the myostatin gene increases muscle mass and enhances racing performance in heterozygote dogs. PLoS Genet 3: e79.
24. Delneri D, Hoyle DC, Gkargkas K, Cross EJ, Rash B, et al. (2008) Identification and characterization of high-flux-control genes of yeast through competition analyses in continuous cultures. Nat Genet 40: 113–117.
25. Krieger U, Lippman ZB, Zamir D (2010) The flowering gene SINGLE FLOWER TRUSS drives heterosis for yield in tomato. Nat Genet 42: 459–463.
26. Pnueli L, Carmel-Goren L, Hareven D, Gutfinger T, Alvarez J, et al. (1998) The SELF-PRUNING gene of tomato regulates vegetative to reproductive switching of sympodial meristems and is the ortholog of CEN and TFL1. Development 125: 1979–1989.
27. Lifschitz E, Eviatar T, Rozman A, Shalit A, Goldshmidt A, et al. (2006) The tomato FT ortholog triggers systemic signals that regulate growth and flowering and substitute for diverse environmental stimuli. Proc Natl Acad Sci U S A 103: 6398–6403.
28. Saltveit ME (2005) Post harvest biology and handling. In: Heuvelink E, editor. Tomatoes. Wallingford, U.K.: CABI Publishing. pp. 305–325.
29. Peet MM, Welles G (2005) Greenhouse tomato production. In: Heuvelink E, editor. Tomatoes. Wallingford, U.K.: CABI Publishing. pp. 257–304.
30. Lippman ZB, Cohen O, Alvarez JP, Abu-Abied M, Pekker I, et al. (2008) The making of a compound inflorescence in tomato and related nightshades. PLoS Biol 6: e288.
31. Efroni I, Blum E, Goldshmidt A, Eshed Y (2008) A protracted and dynamic maturation schedule underlies Arabidopsis leaf development. Plant Cell 20: 2293–2306.
32. Swanson-Wagner RA, Jia Y, DeCook R, Borsuk LA, Nettleton D, et al. (2006) All possible modes of gene action are observed in a global comparison of gene expression in a maize F1 hybrid and its inbred parents. Proc Natl Acad Sci U S A 103: 6805–6810.
33. Jaeger KE, Pullen N, Lamzin S, Morris RJ, Wigge PA (2013) Interlocking feedback loops govern the dynamic behavior of the floral transition in Arabidopsis. Plant Cell 25: 820–833.
34. Park SJ, Jiang K, Schatz MC, Lippman ZB (2012) Rate of meristem maturation determines inflorescence architecture in tomato. Proc Natl Acad Sci U S A 109: 639–644.

35. Alvarez JP, Pekker I, Goldshmidt A, Blum E, Amsellem Z, et al. (2006) Endogenous and synthetic microRNAs stimulate simultaneous, efficient, and localized regulation of multiple targets in diverse species. Plant Cell 18: 1134–1151.
36. Schwab R, Ossowski S, Riester M, Warthmann N, Weigel D (2006) Highly specific gene silencing by artificial microRNAs in Arabidopsis. Plant Cell 18: 1121–1133.
37. Wigge PA (2011) FT, a mobile developmental signal in plants. Curr Biol 21: R374–378.
38. Shalit A, Rozman A, Goldshmidt A, Alvarez JP, Bowman JL, et al. (2009) The flowering hormone florigen functions as a general systemic regulator of growth and termination. Proc Natl Acad Sci U S A 106: 8392–8397.
39. Koornneef M, Hanhart CJ, van der Veen JH (1991) A genetic and physiological analysis of late flowering mutants in Arabidopsis thaliana. Molecular & general genetics : MGG 229: 57–66.
40. Shannon S, Meeks-Wagner DR (1991) A Mutation in the Arabidopsis TFL1 Gene Affects Inflorescence Meristem Development. The Plant cell 3: 877–892.
41. Goff SA (2011) A unifying theory for general multigenic heterosis: energy efficiency, protein metabolism, and implications for molecular breeding. New Phytol 189: 923–937.
42. Doebley JF, Gaut BS, Smith BD (2006) The molecular genetics of crop domestication. Cell 127: 1309–1321.
43. Izawa T, Takahashi Y, Yano M (2003) Comparative biology comes into bloom: genomic and genetic comparison of flowering pathways in rice and Arabidopsis. Current opinion in plant biology 6: 113–120.
44. Zhao K, Tung CW, Eizenga GC, Wright MH, Ali ML, et al. (2011) Genome-wide association mapping reveals a rich genetic architecture of complex traits in Oryza sativa. Nature communications 2: 467.
45. Tsuji H, Taoka K, Shimamoto K (2011) Regulation of flowering in rice: two florigen genes, a complex gene network, and natural variation. Current opinion in plant biology 14: 45–52.
46. Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ, et al. (2009) The genetic architecture of maize flowering time. Science 325: 714–718.
47. Tanksley SD (2004) The genetic, developmental, and molecular bases of fruit size and shape variation in tomato. The Plant cell 16 Suppl : S181–189.
48. Cong B, Barrero LS, Tanksley SD (2008) Regulatory change in YABBY-like transcription factor led to evolution of extreme fruit size during tomato domestication. Nature genetics 40: 800–804.
49. Doganlar S, Frary A, Ku HM, Tanksley SD (2002) Mapping quantitative trait loci in inbred backcross lines of Lycopersicon pimpinellifolium (LA1589). Genome/National Research Council Canada = Genome/Conseil national de recherches Canada 45: 1189–1202.
50. Grandillo S, Tanksley SD (1996) QTL analysis of horticultural traits differentiating the cultivated tomato from the closely related species Lycopersicon pimpinellifolium. Theoretical and Applied Genetics 92: 935–951.
51. Jimenez-Gomez JM, Alonso-Blanco C, Borja A, Anastasio G, Angosto T, et al. (2007) Quantitative genetic analysis of flowering time in tomato. Genome/National Research Council Canada = Genome/Conseil national de recherches Canada 50: 303–315.
52. Peralta IE, Spooner DM (2005) Morphological characterization and relationships of wild tomatoes (Solanum L. Section Lycopersicon). Monogr Syst Bot Missouri Bot Gard: 227–257.
53. Flint-Garcia SA, Buckler ES, Tiffin P, Ersoz E, Springer NM (2009) Heterosis is prevalent for multiple traits in diverse maize germplasm. PLoS One 4: e7433.
54. Semel Y, Nissenbaum J, Menda N, Zinder M, Krieger U, et al. (2006) Overdominant quantitative trait loci for yield and fitness in tomato. Proceedings of the National Academy of Sciences of the United States of America 103: 12981–12986.
55. Schuelke M, Wagner KR, Stolz LE, Hubner C, Riebel T, et al. (2004) Myostatin mutation associated with gross muscle hypertrophy in a child. N Engl J Med 350: 2682–2688.
56. Jiang W, Zhou H, Bi H, Fromm M, Yang B, et al. (2013) Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. Nucleic acids research 41(20):e188. doi: 10.1093/nar/gkt780.
57. Shan Q, Wang Y, Li J, Zhang Y, Chen K, et al. (2013) Targeted genome modification of crop plants using a CRISPR-Cas system. Nature biotechnology 31: 686–688.
58. Taoka K, Ohki I, Tsuji H, Kojima C, Shimamoto K (2013) Structure and function of florigen and the receptor complex. Trends in plant science 18: 287–294.
59. Lifschitz E, Eshed Y (2006) Universal florigenic signals triggered by FT homologues regulate growth and flowering cycles in perennial day-neutral tomato. Journal of experimental botany 57: 3405–3414.
60. Meng X, Muszynski MG, Danilevskaya ON (2011) The FT-like ZCN8 Gene Functions as a Floral Activator and Is Involved in Photoperiod Sensitivity in Maize. Plant Cell 23: 942–960.
61. Blackman BK, Strasburg JL, Raduski AR, Michaels SD, Rieseberg LH (2010) The role of recently derived FT paralogs in sunflower domestication. Curr Biol 20: 629–635.
62. Quinby JR, Karper RE (1946) Heterosis in sorghum resulting from the heterozygous condition of a single gene that affects duration of growth. Am J Bot 33: 716–721.
63. Comadran J, Kilian B, Russell J, Ramsay L, Stein N, et al. (2012) Natural variation in a homolog of Antirrhinum CENTRORADIALIS contributed to spring growth habit and environmental adaptation in cultivated barley. Nat Genet 44: 1388–1392.
64. Pin PA, Benlloch R, Bonnet D, Wremerth-Weich E, Kraft T, et al. (2010) An antagonistic pair of FT homologs mediates the control of flowering time in sugar beet. Science 330: 1397–1400.
65. Kwak M, Velasco D, Gepts P (2008) Mapping homologous sequences for determinacy and photoperiod sensitivity in common bean (Phaseolus vulgaris). J Hered 99: 283–291.
66. Repinski SL, Kwak M, Gepts P (2012) The common bean growth habit gene PvTFL1y is a functional homolog of Arabidopsis TFL1. Theor Appl Genet 124: 1539–1547.
67. Fernandez L, Torregrosa L, Segura V, Bouquet A, Martinez-Zapater JM (2010) Transposon-induced gene activation as a mechanism generating cluster shape somatic variation in grapevine. Plant J 61: 545–557.
68. Navarro C, Abelenda JA, Cruz-Oro E, Cuellar CA, Tamaki S, et al. (2011) Control of flowering and storage organ formation in potato by FLOWERING LOCUS T. Nature 478: 119–122.
69. Iwata H, Gaston A, Remay A, Thouroude T, Jeauffre J, et al. (2012) The TFL1 homologue KSN is a regulator of continuous flowering in rose and strawberry. Plant J 69: 116–125.
70. Tian Z, Wang X, Lee R, Li Y, Specht JE, et al. (2010) Artificial selection for determinate growth habit in soybean. Proc Natl Acad Sci U S A 107: 8563–8568.
71. Liu B, Watanabe S, Uchiyama T, Kong F, Kanazawa A, et al. (2010) The soybean stem growth habit gene Dt1 is an ortholog of Arabidopsis TERMINAL FLOWER1. Plant Physiol 153: 198–210.
72. Harig L, Beinecke FA, Oltmanns J, Muth J, Muller O, et al. (2012) Proteins from the FLOWERING LOCUS T-like subclade of the PEBP family act antagonistically to regulate floral initiation in tobacco. Plant J doi: 10.1111/j.1365-313X.2012.05125.x. [Epub ahead of print].
73. Hedden P (2003) The genes of the Green Revolution. Trends in genetics : TIG 19: 5–9.
74. Menda N, Semel Y, Peled D, Eshed Y, Zamir D (2004) In silico screening of a saturated mutation library of tomato. The Plant journal : for cell and molecular biology 38: 861–872.
75. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10: R25.
76. Tomato Genome C (2012) The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485: 635–641.
77. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078–2079.
78. RDC T (2011) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.
79. Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26: 139–140.
80. Allen E, Howell MD (2010) miRNAs in the biogenesis of trans-acting siRNAs in higher plants. Seminars in cell & developmental biology 21: 798–804.
81. Moissiard G, Parizotto EA, Himber C, Voinnet O (2007) Transitivity in Arabidopsis can be primed, requires the redundant action of the antiviral Dicer-like 4 and Dicer-like 2, and is compromised by viral-encoded suppressor proteins. RNA 13: 1268–1278.