MOLECULAR BASIS OF THE TOMATO GREEN SHOULDER PHENOTYPE, ITS ROLE IN FRUIT NUTRITIONAL QUALITY AND CHARACTERIZATION OF MOLECULAR GRADIENTS IN FRUIT DEVELOPMENT

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Cuong Viet Nguyen

January 2015

MOLECULAR BASIS OF THE TOMATO GREEN SHOULDER PHENOTYPE, ITS ROLE IN FRUIT NUTRITIONAL QUALITY AND CHARACTERIZATION OF MOLECULAR GRADIENTS IN FRUIT DEVELOPMENT

Cuong Viet Nguyen, Ph. D.

Cornell University 2015

Abstract

The tomato *uniform ripening* (*u*) mutation confers a light green fruit phenotype absent the wild-type dark green top or "green shoulder" and is widely used in breeding for tomatoes producing evenly ripened fruit, facilitating selection of harvest time and sale of visually appealing fresh fruit. Positional cloning of *u* revealed that *U* encodes a *Golden 2-like (GLK)* transcription factor, *SlGLK2*, influencing chloroplast and chlorophyll accumulation and distribution in developing fruit. Transgene-mediated ectopic expression of two *GLK* genes, *SlGLK1* and *SlGLK2*, affects chlorophyll content, chloroplast number and thylakoid grana stacks. Ectopic *GLK* expression also elevated levels of starch in green fruit and sugars, carotenoids and ascorbic acid in ripe fruit indicating that manipulating *GLK* expression is a means to enhance fruit nutritional quality. Use of the *u* allele thus likely confers an unintended consequence of reduced fruit nutritional value. To test this hypothesis, nearly isogenic lines for the *U* locus in an elite background (cv. *M82*) were developed to assess the effects of the green shoulder phenotype on fruit quality. Although plants with green shoulder fruit have 5-10% more fruit with a yellow shoulder phenotype than u plants, *U* fruit displayed 10-20% increases in soluble solids and carotenoids compared to *u* fruit. The isogenic lines also present tightly linked and novel genetic polymorphisms with potential as markers for molecular assisted selection. A latitudinal gradient

of *GLK2* expression influences the typical uneven coloration of green and ripe wild-type fruit. Transcriptome profiling highlighted a broader gene expression gradient through fruit development revealing additional molecular complexity during ripening. As with the latitudinal gradient, spatial and temporal gene expression differences in different fruit tissues are still poorly understood. Comparing transcriptomes from locular and pericarp tissues during fruit development suggests that climacteric ethylene and overall ripening may initiate in the locules.

BIOGRAPHICAL SKETCH

Cuong Nguyen graduated from Vietnam National University in Hanoi, Vietnam in 2006 with a Bachelor degree in Biology. Before he came to the United States, he worked as a research assistant at the Institute of Biotechnology, Vietnam Academy of Science and Technology. He was awarded a Vietnam Education Foundation Fellowship and admitted to Cornell University to pursue a PhD degree in Plant Breeding in the Department of Plant Breeding and Genetics in the Fall 2008. He joined Dr. Jim Giovannoni's lab at the Boyce Thompson Institute for Plant Research to do his thesis work on the genetics of tomato fruit development and fruit nutritional quality. Both his parents are tomato researchers and he is proud to continue the family tradition. He is living with his wife and his little daughter whom was born in 2012.

I would like to dedicate this dissertation to my family, especially my mother.

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor, Dr. James J. Giovannoni for all the opportunities, support, guidance and encouragement throughout my thesis work. This thesis is not possible without the chance that Jim gave me to freely do research on the topics that I love. I would like to extend my sincere thanks to my other committee members, Dr. Jocelyn Rose and Dr. Chris Watkins for their helpful input and suggestions for my research.

I am very grateful to all members of the Giovannoni laboratory, past and present for the friendship and many years of conversation and critical feedback. I specifically express my thanks to Julia Vrebalov, Ruth White, Nigel Gapper, Yimin Xu, Betsy Ampofo, Je Min Lee, Silin Zhong and Ryan McQuinn for all their contributions to my research.

Special thanks to Dr. Joyce van Eck, Patricia Keen and Lesley Middleton for their work on generation of transgenic plants, to Brian Bell and all the greenhouse staff for taking care of my plants, to Dr. Zhangjun Fei and Yi Zheng for all their help with Illumina sequencing analyses, to Prof. Andre Jargendorf for helping me with sugar assays, and to John Grazul for the guidance on electron microscopy.

I would like to thank my family for all of the support they have given me throughout my life. Mom I love you so much and I wish you could be here to celebrate with me this important milestone in my life. During my stay in the US, I am blessed to be accompanied by my wife, Hanh Nguyen and my little daughter, Minh-Thu Nguyen. I feel so happy to have them along all the time.

Finally, I would like to acknowledge the Vietnam Education Foundation Fellowship grant for the financial support of my research.

TABLE OF CONTENTS

Page

BIOGRAPHICAL SKETCH iii

DEDICATION iv

ACKNOWLEDGEMENTS v

TABLE OF CONTENTS vi

CHAPTER ONE: INTRODUCTION 1

CHAPTER TWO: MOLECULAR BASIS OF THE GREEN SHOULDER/ UNIFORM RIPENING PHENOTYPE IN TOMATO AND ITS ROLE IN FRUIT NUTRITIONAL QUALITY 32

CHAPTER THREE: FRUIT QUALITY ANALYSIS OF TOMATO NEARLY ISOGENIC LINES FOR THE GENE CONFERRING THE GREEN SHOULDER PHENOTYPE 74

CHAPTER FOUR: CHARACTERIZATION OF MOLECULAR GRADIENTS IN FRUIT DEVELOPMENT 96

CHAPTER FIVE: TRANSCRIPTOMIC DIFFERENCES BETWEEN LOCULAR GEL AND PERICARP TISSUES DURING FRUIT DEVELOPMENT 136

SUMMARY AND FUTURE DIRECTIONS 161

CHAPTER I.

INTRODUCTION

Tomato is one of the most important vegetables in the world. Worldwide tomato production was nearly 160 million tons in 2011 and total production as well as total cultivated area has doubled in the last 20 years (FAOSTAT, http://faostat3.fao.org/home/index.html). Tomato is also familiar to the public as a favored garden crop. Since the early 20th century, tomato has increasingly become a model organism for plant research, both for basic and applied purposes. The cultivated tomato (*Solanum lycopersicum*) belongs to the *Solanaceae* family, containing more than 3,000 species including many economically important plants such as potato, peppers (*Capsicum*), eggplant, petunia, tobacco, tomatillo (*Physalis)* in addition to many other species producing toxic or medicinal compounds (Knapp, 2002). Several attributes facilitate the utilization of tomato as a model research system: it is relatively easy to grow under a diversity of conditions; it has a short life cycle with seed to seed generation time of approximately 3 months; tomato is a self-pollinated species, but also easily outcrossed; the genome is diploid, highly homozygous, relatively small (~900 Mb) and has been fully sequenced (Tomato Genome Consortium, 2012); large collections of mutants, wild species, mutation lines, introgression lines cumulating years of development and analysis are readily available for researchers to use without restriction (Tomato Genetics Resource Center, tgrc.ucdavis.edu); tomato can also be propagated asexually by grafting; and methods for transformation by

Agrobacterium are routine. The fleshy fruit of tomato and its dramatic and easily monitored development and ripening have rendered tomato a widely studied model for fruit development and ripening.

Tomato breeding, as for most crops, traditionally focuses on yield and disease resistance, but recently has shifted to increased attention on quality traits including shelf-life and especially taste and nutritional content (Causse, 2007). Over the past 20 years, extensive work in tomato and *Arabidopsis* has provided a plethora of knowledge concerning fruit development and fruit ripening that can be, and often is, readily applied to breeding and horticultural objectives.

Fruit structure and development

The tomato fruit is composed of a pericarp with several different layers/tissue types (exocarp, mesocarp, endocarp), the placental tissues surrounding the seeds (which later give rise to the locular gel) and the septa that divide the fruit/ovary to define the internal locule boundaries (Gillaspy et al., 1993). The number of carpels determine the number of locules of the fruit. The pericarp has five major cell types/tissues from outside to inside: the outer epidermis (one single cell layer), collenchyma, parenchyma, vascular tissue and inner epidermis (a single cell layer) (Matas et al., 2011).

Early fleshy fruit development begins at fertilization, which triggers cell division as the first stage of fruit growth. After cell division ceases (typically 5-10 days depending on variety), the fruit continue to grow primarily through cell expansion until they reach their final size (Gillaspy et al., 1993; Tanksley, 2004). During this process, spatial and temporal variations in concentrations of several hormones, including auxin, cytokinin and gibberellin, play important roles in initiating and maintaining cell division and expansion. Parthenocarpic fruit have higher

levels of auxin and gibberellin compared to normal seed producing fruit and parthenocarpy can be achieved by applying auxin or gibberellin to the ovary prior to pollination.

Tomato fruit size and shape are important agronomic traits and have been subject to constant artificial selection during domestication and subsequent breeding, resulting in highly variable tomato fruit sizes and shapes in modern varieties. As such, fruit size can vary from less than 20 g in cherry tomato types to up to 500 g or more for so-called "beef steak" types. QTL mapping studies have identified major genetic loci that underlie this morphological variation. The *Fw2.2* gene is responsible for up to 30% of the variation in tomato fruit size and encodes a negative repressor of cell division homologous to human oncogenes (Frary et al., 2000). Two additional loci, *fasciated* and *locule number* affect fruit size by altering the number of carpels and hence locules. *FASCIATED* encodes a *YABBY* transcription factor (Cong et al., 2008) and *LOCULE NUMBER* encodes a gene in the *WUSCHEL* homeodomain family, which includes members regulating plant stem cell fate (Rodríguez et al., 2011). *sun* and *ovate* are two major tomato loci influencing fruit elongation. *SUN* encodes an IQD family protein that is thought to alter auxin levels (Xiao et al., 2008), whereas *OVATE* encodes a class of protein acting as a negative modulator resulting in reduced fruit length (Liu et al., 2002).

Hormonal and Genetic Regulation of Maturation and Ripening in Fleshy Fruits

After growth has ceased, tomato fruit undergo a ripening phase involving changes in structure, physiology, and biochemistry that influence fruit appearance, texture, flavor, and aroma facilitating attraction of seed dispersing organisms. These changes typically include modification of color through de-greening (loss of chlorophyll) and accumulation of colored pigments such as carotenoids or flavonoids, textural changes due to cell wall metabolism and cell

turgor alteration, and changes in fruit chemistry such as altered levels of sugars, acids and volatiles (Giovannoni et al, 2004).

Physiologically, fruit can be classified into two major groups, climacteric and non-climacteric (Biale, 1960). Climacteric fruit, like tomato, apple, and banana, display a pronounced increase in respiration coincident with elevated production of the gaseous hormone ethylene at the onset of ripening. Ethylene is often essential for normal fruit ripening in climacteric species, and fruit with defects in either ethylene synthesis or signal transduction typically fail to ripen. Non-climacteric fruit, including citrus, strawberry and grape, can ripen in the absence of elevated ethylene and do not exhibit increased respiration upon initiation of ripening (Klee and Giovannoni, 2011).

Ethylene synthesis and signaling

Ethylene is a gaseous plant hormone contributing to numerous plant responses and developmental processes including biotic and abiotic stress, seed germination, root initiation, floral development, senescence and fruit ripening (Bleecker and Kende, 2000). In plants, ethylene is derived from methionine metabolism (Adams and Yang, 1979). Methionine is converted to SAM (s-adenosyl-l-methionine) by SAM synthetase. SAM is then metabolized to 1-aminocyclopropane-1-carboxylate (ACC) by ACC synthase (ACS). ACC is subsequently converted to ethylene by ACC oxidase (ACO). The intermediate product methylthioadenine from the SAM to ACC reaction is reconverted to methionine by a series of reactions termed the Yang cycle (Adams and Yang, 1979; Dong et al., 1992) insuring a sufficient supply of methionine for ethylene synthesis and additional biological activities.

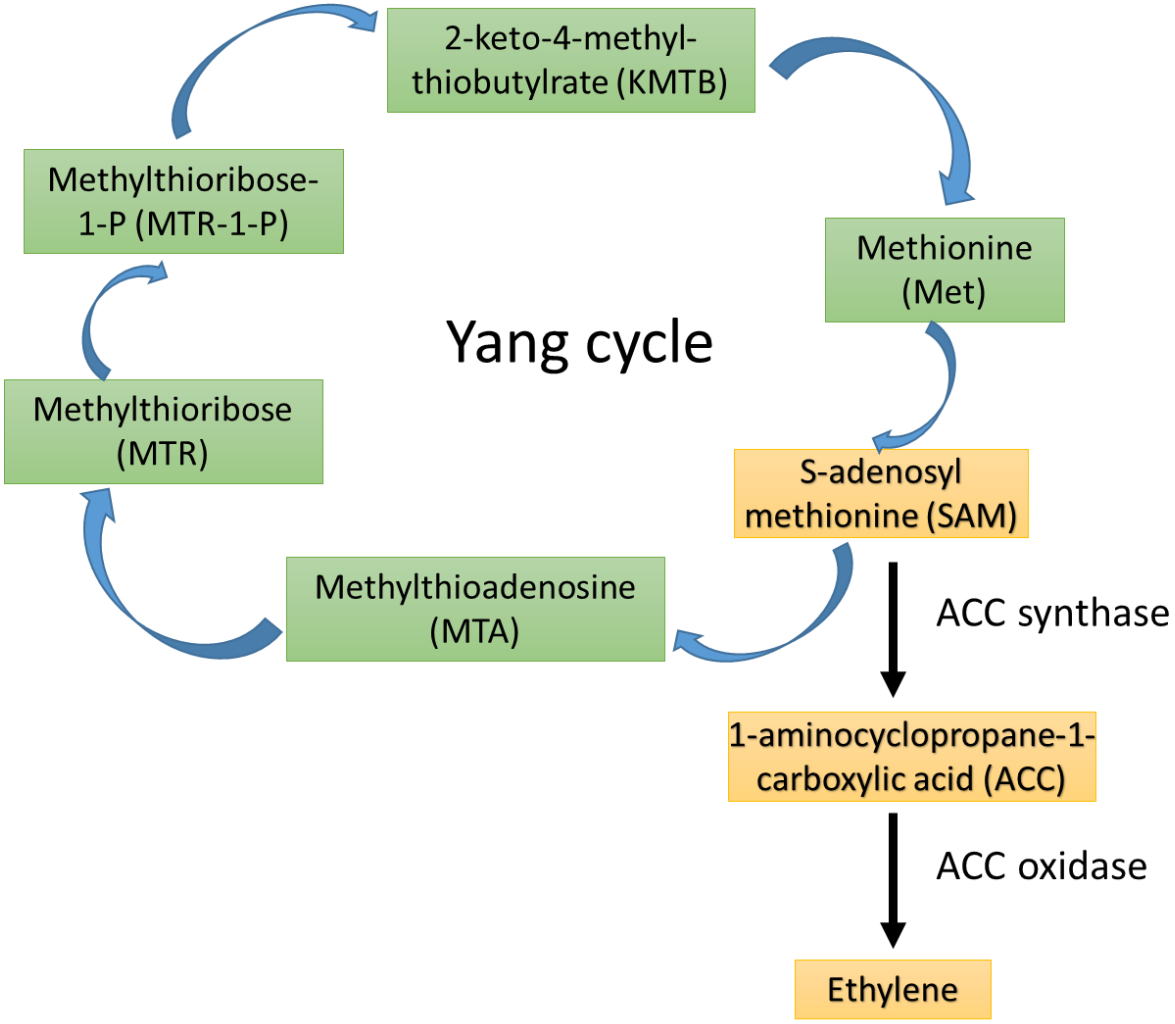


Figure 1.1. Ethylene biosynthesis pathway and the Yang cycle

ACC synthase is often the rate-limiting enzyme in ethylene synthesis and as such transcriptional regulation of ACS is highly regulated including during ripening (Rottmann et al., 1991; Nakatsuka et al., 1998). In tomato, ACS is encoded by a gene family of at least nine members (*LeACS1A*, *LeACS1B* and *LeACS2-8*), with three additional putative genes in the genome sequence (Tomato Genome Consortium, 2012). Among these genes, *LeACS2* and *LeACS4* are induced and contribute the majority of transcripts during ripening (Barry et al., 1996). Fruit with transgene repressed *LeACS2* failed to ripen but can be rescued by treatment with exogenous ethylene (Rottmann et al., 1991). Besides transcriptional regulation, post-translational modification of ACS proteins provides an additional control system over ethylene synthesis. As an example, the LeACS2 protein is stabilized by phosphorylation at the C-terminal region and degraded following de-phosphorylation (Tatsuki and Mori, 2001).

Like ACS, there are multiple *ACO* genes in tomato with five genes functionally defined and another three genes found in the genome sequence (Tomato Genome Consortium, 2012). Most tomato *ACO* genes are expressed in fruit except for *LeACO2*. During ripening, ACO genes are up-regulated and suppression of tomato *ACO1* (as well as *ACO* in several additional fruit species including melon and apple) inhibits normal ripening and ethylene production confirming the function of the genes in ethylene biosynthesis and demonstrating the potential for practical application to extend shelf-life (Hamilton et al., 1990; Ayub et al., 1996; Atkinson et al., 2011).

Ethylene receptors and signal transduction pathway components were discovered first in *Arabidopsis* through elegant characterization of ethylene response mutants and epistasis interactions (Bleecker and Kende, 2000; Guo and Ecker, 2004). Subsequent genetic analyses in tomato showed conservation of the basic signal transduction components defined in *Arabidopsis* (Wilkinson et al., 1995; Leclercq et al., 2002; Adams-Phillips et al., 2004). Plant ethylene

receptors are related to the two component histidine kinases in bacteria, and are found in the endoplasmic reticulum as integral membrane proteins. In tomato, the first ethylene receptor was discovered through cloning of the *Never-ripe (Nr)* dominant ethylene insensitive mutation (Lanahan et al., 1994; Wilkinson et al., 1995; Yen et al., 1995). *Nr* fruit do not ripen even with exogenous ethylene and this lack of *Nr* mediated ripening confirms the essential role of ethylene and its perception in ripening. Tomato has seven ethylene receptor genes and *Arabidopsis* has five (Klee, 2004). While there are differences in expression profiles of each receptor in different tissues and development stages, in general, suppression of receptor expression increases ethylene sensitivity, whereas increased receptor content has the opposite effect confirming their function as negative regulators of ethylene signaling (Hua and Meyerowitz, 1998; Tieman et al., 2000; O'Malley et al., 2005).

Blocking ethylene receptor activity has been a long-standing target for ethylene control in crop production. 1-Methylcyclopropene (1-MCP) can prevent ethylene signaling by binding to the ethylene receptors has been used successfully to delay fruit ripening and delay senescence of ornamental plants (Blankenship and Dole, 2003). A major industry centered on 1-MCP production, application and expansion of use has evolved around this compound, specifically for apple fruit in the last 12 years (Watkins, 2006, 2008; Mahajan et al., 2014). Kevany et al., (2007) further refined the mode of ethylene receptors biology by demonstrated that ethylene binding triggers receptor degradation through the 26S proteosome-dependent pathway. This result indicated that although receptor genes are highly expressed during ripening, the actual levels of receptor proteins remain low throughout ripening.

Immediately downstream of ethylene receptor function in *Arabidopsis* is a single *CONSTITUTIVE TRIPLE RESPONSE1* (*CTR1*) map kinase gene. Loss of function *ctr1* mutants

exhibited consecutive ethylene response phenotypes (e.g. seedling triple response) indicating CTR is a negative ethylene response regulator (Hua and Meyerowitz, 1998). In tomato, there are four *CTR* homologs, three of which are capable of complementing the *Arabidopsis ctr1* loss of function mutant (Leclercq et al., 2002; Adams-Phillips et al., 2004). Zhong et al. (2008) showed that the tomato NR ethylene receptor is able to bind multiple LeCTR proteins, supporting the model where ethylene receptors directly transmit signals to CTR proteins.

Further downstream steps of the pathway are mediated by the positive regulators *ETHYLENE INSENSITIVE 2 (EIN2)* and *ETHYLENE INSENSITIVE 3 (EIN3). EIN2* is negatively controlled by *CTRs*: when ethylene binds to the receptors, CTRs change conformation and EIN2 is released to activate transcription factor *EIN3*. In the absence of ethylene, CTR1 directly phosphorylates the cytosolic C-terminal domain of EIN2 and prevents EIN2 from signaling. Upon ethylene perception, CTR1 no longer interacts with EIN2 and this is a signal for cleavage and nuclear localization of EIN2 (Ju et al., 2012). In tomato there are four *EIN3*-like genes (*EIL1-4*) and suppression of *LeEIL1*, *LeEIL2*, and *LeEIL3* reduced ethylene sensitivity (Tieman et al., 2001; Yokotani et al., 2003). EIN3 is the target of a degradation system (ubiquitin/26S proteasome), mediated by F box proteins called EIN3-binding factor (EBF) F-box proteins. Stabilizing the LeEIL proteins by reducing EBF expression leads to consecutive ethylene response and early fruit ripening (Guo and Ecker, 2003; Potuschak et al., 2003; Chen et al., 2004; Gagne et al., 2004).

The last step of the ethylene signaling pathway includes a large family of transcription factors called *ETHYLENE RESPONSE FACTORS* (*ERFs*) that are activated by *EILs*. While a large gene family, only a few genes have been assigned specific functions. For example, Lee et al. (2012) used a systems biology approach by combining transcriptome and metabolome data to

identify *SlERF6* as a negative regulator of ethylene, carotenoid biosynthesis and overall ripening in maturing tomato fruit.

Tomato GREEN-RIPE and the *Arabidopsis* homolog, REVERSION TO ETHYLENE SENSITIVITY1 (RTE1) define additional proteins contributing to ethylene signaling (Barry and Giovannoni, 2006; Resnick et al., 2006). In the case of tomato, *GR* was discovered through cloning of the locus harboring a dominant non-ripening mutant with fruit phenotypes similar to *Nr*. GR is not an ethylene receptor and its biochemical function remains elusive though its structure suggests a possible role in receptor-copper interaction. Copper is a necessary cofactor for ethylene receptor function (Rodriguez et al., 1999).

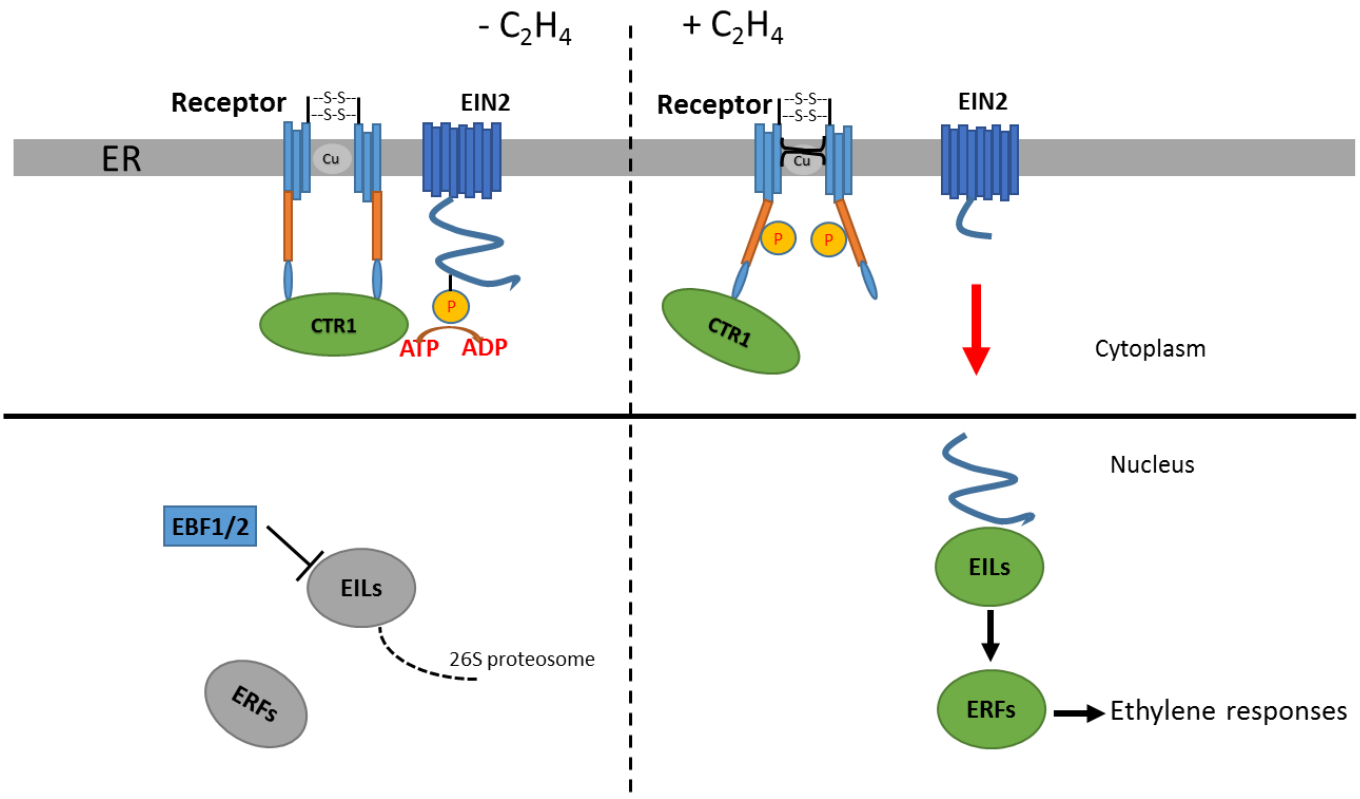


Figure 1.2. Ethylene signal transduction.

Left: Without ethylene, the ethylene receptors at the endoplasmic reticulum (ER) membrane activate the CTR1 protein kinase, which in turn phosphorylates the C-terminal domain of EIN2, preventing its nuclear localization. F-box proteins EBF1/2 target transcription factors EILs for degradation.

Right: With ethylene, ethylene binding causes phosphorylation and inactivates the receptors resulting in inactivation of the CTR1 kinase. Without phosphorylation, the EIN2 C terminus is cleaved and localized to the nucleus where it activates the downstream transcriptional cascade.

Transcriptional regulation of fruit ripening

Several single gene mutations that abolish the normal ripening process have provided key information about ripening control. Fruit from three mutants, *ripening-inhibitor (rin)*, *non-ripening (nor)*, and *Colorless nonripening (Cnr)*, have similar characteristics: they do not advance to ripening; they fail to produce climacteric respiration and ethylene but also do not ripen in response to exogenous ethylene; however, ethylene responsive genes are still induced by ethylene (Vrebalov et al., 2002; Manning et al., 2006; Giovannoni, 2007). These characteristics indicate that these mutants have defects in ripening regulatory systems and reveal additional aspects of ripening control beyond ethylene production and perception.

RIN encodes a *MADS*-box transcription factor that is a member of the *SEP* clade as defined in *Arabidopsis* (Vrebalov et al., 2002). *RIN* gene expression is highly induced at the onset of ripening. The *rin* mutation affects virtually all ripening pathways, which supports its role as one of the critical regulators of the ripening process. In fact, the *rin* mutant allele is widely used in the heterozygous state to produce hybrids with long shelf life and fruit firmness. The nature of the mutation is a spontaneous deletion that removes the C-terminus of the *RIN* gene and approximately 1 kb of the sequence separating *RIN* from the adjacent gene *MACROCALYX* (*MC*), which is the tomato homolog of *APETALA1*, and resulting in loss-of-function of both *RIN* and *MC*. Chromatin immunoprecipitation analyses showed that the gene interacts with CArG-box elements in promoters of genes involved in a wide range of ripening phenomena including ethylene synthesis and perception, affecting genes such as *LeACS2*, *LeACS4*, *NR*, and *E8*; cell wall metabolism, via polygalacturonase (*PG*), galactanase (*TBG4*), and expansins (*LeEXP1*); carotenoid synthesis through phytoene synthase (*PSY1*) and aroma biosynthesis, via lipoxygenase (*Tomlox C*) and alcohol dehydrogenase (*ADH2*) (Fujisawa et al.,

2011; Martel et al., 2011; Kumar et al., 2012; Qin et al., 2012). *RIN* is also involved in suppressing the expression of most *ARF* genes, suggest a link between auxin and ripening control (Kumar et al., 2011).

CNR encodes a transcription factor of the *SQUAMOSA PROMOTER BINDING PROTEIN* family (*SPBP*) (Manning et al., 2006). *Cnr* is a rare example of a heritable epigenetic allele, with high methylation in the promoter of the mutant allele resulting in reduced transcription. *CNR* expression is reduced in the *rin* mutant; however, the relationship among these regulators and ethylene is likely not linear. Interestingly, interaction of the RIN protein with the promoters of target genes required CNR, suggesting a link between these two transcription factors (Martel et al., 2011). Specific interactions of *CNR* with ripening gene targets are still to be tested; however' there is evidence that tomato ortholog of the *Arabidopsis FRUITFUL* (*FUL*) *MADS*-box gene, *TDR4*, is greatly reduced in *Cnr* (Manning et al., 2006). *TDR4/FUL1* repression seemed to have no effect on the fruit ripening phenotype (Jaakola et al., 2010) but both tomato *FUL1* and *FUL2* suppression result in ripening phenotypes including reduced pigment accumulation (Bemer et al., 2012), similar to the antisense repression of an orthologous gene in bilberry (Jaakola et al., 2010).

The nature of the *Cnr* mutation as an epi-allele suggests that epigenetics may contribute to ripening control. The *Cnr* epi-allele confers ripening inhibition through heritable hypermethylation of cytosine residues at the *CNR* promoter (Manning et al., 2006). An intriguing experiment involving treatment of immature fruit with a methyltransferase inhibitor resulted in ripening of immature fruits well in advance of seed maturation and normal ripening (Zhong et al., 2013). Whole genome methylome analysis of tomato fruit from the immature to ripe stages indicated there are distinct regions of DNA cytosine methylation that are de-methylated during

fruit maturation and ripening and these regions are often localized upstream of ripening-related genes, and are adjacent to the binding sites of the RIN protein (Zhong et al., 2013).

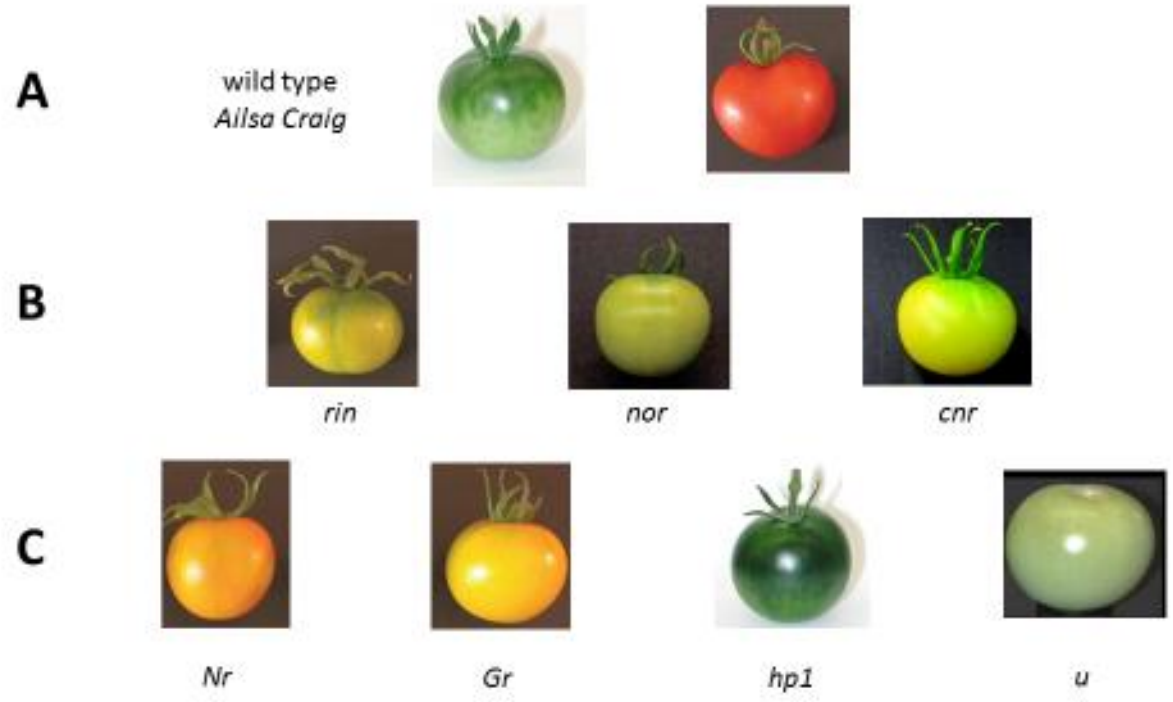


Figure 1.3. Fruit ripening and pigmentation mutants of tomato.

A. Wild-type (cultivar *'Ailsa Craig'*)

B. Near isogenic lines (NILs) homozygous for the *ripening-inhibitor (rin)*, *non-ripening (nor)* and *Colorless non-ripening* (*Cnr*). The *rin*, *nor* and *Cnr* loci act upstream in the ripening regulatory pathway.

C. Near isogenic lines homozygous for *Never-ripe (Nr)*, *Green-ripe (Gr)*, *high pigment 1* (*hp1*) and *uniform ripening* (*u*) loci. The non-ripening phenotypes of *Nr* and *Gr* are caused by reduced ethylene responsiveness. *hp1* fruit have elevated chlorophylls and carotenoids. *u* fruit do not display the "green shoulder".

In addition to cloning of natural ripening mutants like *rin* and *Cnr*, reverse genetics strategies involving transcriptome profiling and transgenic analysis of differentially expressed putative regulatory genes have identified additional ripening mediators. A HD-zip homeobox protein (*LeHB1*) was found to directly interact with *LeACO1*, the most highly expressed ACO gene during fruit ripening (Lin et al., 2008). *LeHB1* repression using virus induced gene silencing (VIGS) resulting in delayed ripening and reduced LeACO1 transcripts. Interestingly, ChIP assays showed that although RIN does not interact directly with *LeACO1*, it binds to the promoter of *LeHB1*, indirectly controlling this gene in the ethylene biosynthetic pathway as well as other through direct interaction (Martel et al., 2011).

Tomato *AGAMOUS-like 1* (*TAGL1*) is one of the genes targeted for functional analysis following transcriptome studies (Alba et al., 2005). It is a *MADS*-box gene of the *AGAMOUS* clade, and is expressed highly both early in carpel development and later at the onset of ripening. *TAGL1* repression by RNAi resulted in ripening inhibition and a reduction of carpel thickness due to reduced cell division (Itkin et al., 2009; Vrebalov et al., 2009). The ripening phenotype occurs partly through inhibition of ethylene biosynthesis, as *ACS2* expression is significantly reduced in *TAGL1*-silenced lines and up-regulation of *LYC-B* and *CYC-B* (converting lycopene to β-carotene) in *TAGL1* repression lines led to low lycopene accumulation. *TAGL1* overexpression resulted in fleshy sepals displaying ripening characteristics including lycopene accumulation (Vrebalov et al., 2009). It is important to note that *TAGL1* is the tomato ortholog of the *Arabidopsis SHATTERPROOF* (*SHP*) gene. *SHP* is responsible for development of valve margin tissue that leads to silique dehiscence and pod shattering culminating in seed dispersal (Liljegren et al., 2000). The functions of TAGL1 in tomato suggest its role is also involved in

seed dispersal through the development of fleshy tissue and later ripening, making fruit attractive and palatable for consumption by animal seed dispersers.

SlAP2a, a transcription factor in the *AP2/ETHYLENE RESPONSE FACTOR (ERF)* gene family, was also selected for functional studies following microarray analysis. Repression of *SlAP2a* via RNAi resulted in accelerated ripening, elevated ethylene production, and altered carotenoid accumulation indicating that AP2 is a negative regulator of ripening in tomato (Chung et al., 2010; Karlova et al., 2011). CNR has been demonstrated to bind to the promoter of *AP2 in vitro* (Karlova et al., 2011).

In summary, there are multiple transcription factors contributing to the regulation of fruit ripening and derived from different gene families, suggesting a complex and layered regulatory hierarchy with multiple avenues of effect and interaction. There is also evolutionary conservation of the regulatory circuits between different fruit types and species as MADS-box genes, for example, play a central ripening function in different species including strawberry and possibly the fruit of the monocot, banana (Elitzur et al., 2010; Seymour et al., 2011). These regulators likely facilitate fruit development, maturation and ripening-related changes in many fruit species.

Cell wall metabolism

Textural changes and fruit softening during ripening make the fruit more palatable to seed dispersers. However, softening reduces fruit shelf life and makes them more susceptible to opportunistic pathogens and less resistant to storage and transport (Watkins, 2006, 2008). These changes are mainly due to cell wall metabolism, a very complex process catalyzed by many different groups of cell wall modifying enzymes and structural proteins. (Brummell, 2006; Watkins, 2006; Vicente et al., 2007).

In general, primary plant cell walls are composed of cellulose microfibrils and cross-linked hemicelluloses such as xyloglucan embedded in a mixture of hydrated pectins (for example, homogalacturonan and rhamnogalacturonans) and structural proteins (Cosgrove, 2005). Polygalacturonase (PG) is a major cell wall polyuronide degrading enzyme that is highly induced during tomato ripening and which catalyzes the degradation of polygalacturonan by hydrolysis of the glycosidic bonds that link galacturonic acid residues. PG is regulated by ethylene and was thought to play a key role in fruit softening. However, PG suppression only led to slightly firmer fruit suggesting that PG is not the only enzyme responsible for fruit softening (Giovannoni et al., 1989; Smith et al., 1990). Nevertheless, PG suppression was the target for the first transgenic crop to be marketed, Flavr Savr, with the objective of increased shelf life and firmness allowing marketing of more fully ripe and flavorful fruit (Martineau, 2001).

Several additional classes of cell wall enzymes and proteins function during ripening and have been characterized in multiple fruit species but predominantly and most thoroughly in tomato. These include endo-1,4-beta-glucanase (EGases), β-galactosidases (GAL), pectin methyl esterases (PME), pectate lyases (PL), xyloglucan endotransglucosylase/hydrolases (XTH) and expansins (EXP) (reviewed in Vicente et al., 2007). In most cases, each class of enzymes acts on substrates that are part of the cell wall. Expansins are unique in that they are nonenzymatic proteins facilitating disruption of the hydrogen bonds between cellulose microfibrils and cross linking glycans to modify cell wall structure and integrity (Rose et al., 1997; Cosgrove, 2000).

The cuticle is a proportionally thin lipophilic layer that covers the aerial organs of all plants. Cuticles are composed of two main components: cutin, a polymer of mainly C16 and C18 hydroxyl fatty acids or diacids, and waxes (Isaacson et al., 2009). In tomato pericarp, there is a thick hydrophobic cuticle coating the outer epidermis and tomato mutants with low levels of

cutin showed increase water loss and pathogen susceptibility (Yeats et al., 2012). The cuticle provides an efficient barrier against water loss, maintains fruit skin integrity and limits microbial infection. Flavonoids (e.g. anthocyanins) are also found embedded in the cuticle matrix, facilitating protection against microbial infection and UV irradiation (Treutter, 2005). Cloning of cuticle deficient mutants and comparative expression profiling helped identify important genes involved in the synthesis of this important layer of plant tissues. Examples include transcription factors such as a homeodomain gene (*HD-GL2),* and a *SHINE* gene (*SlSHN3),* in addition to a cutin synthase (*CD1*) and *GDSL* lipase (*GDSL1*) (Isaacson et al., 2009; Yeats et al., 2012; Shi et al., 2013).

Sugars, pigments and aroma volatile compounds

During ripening various metabolic changes increase in soluble solids or sugars (glucose and fructose), carotenoids (especially lycopene and β-carotene), ascorbic acid (vitamin C) and aroma volatiles contribute to the nutritional quality of the fruit.

Pigment accumulation is one of the hallmark changes during fruit ripening. In tomato, the color change is due to increased carotenoid synthesis and accumulation, mainly as related to lycopene and to a lesser extent, β-carotene. In higher plants, carotenoids are synthesized from the MEP (methylerythritol 4-phosphate) pathway in the plastids. The carotenoid pathway has been well characterized (Hirschberg, 2001). As isoprenoids, carotenoids are synthesized from the five carbon unit isopentenyl diphosphate (IPP) and its double-bond isomer dimethylallyl diphosphate (DMAPP). GGPP (geranylgeranyl diphosphate), a 20 C molecule is formed by adding three molecules of IPP to DMAPP catalyzed by GGPP synthase (GGPS). The condensation of two molecules of GGPP to produce phytoene is catalyzed by phytoene synthase (PSY). This is the

first committed step as well as the often rate limiting step in carotenoid biosynthesis. A series of desaturation and isomerization steps of uncolored phytoene by ZDS (ζ-carotene desaturase) and PDS (phytoene desaturase) and two isomerases (ζ-carotene isomerase, Z-ISO; and carotene isomerase, CRTISO), results in the production of predominantly trans-lycopene, a red carotenoid. The cyclization of the ends of the lycopene polyene chain can be catalyzed by either lycopene β or ε- cyclase and leads to β - or α-carotene, respectively. Xanthophylls such as lutein (from α-carotene) and zeaxanthin (from β carotene) are generated by the hydroxylation of the carotene rings (Stigliani et al., 2011).

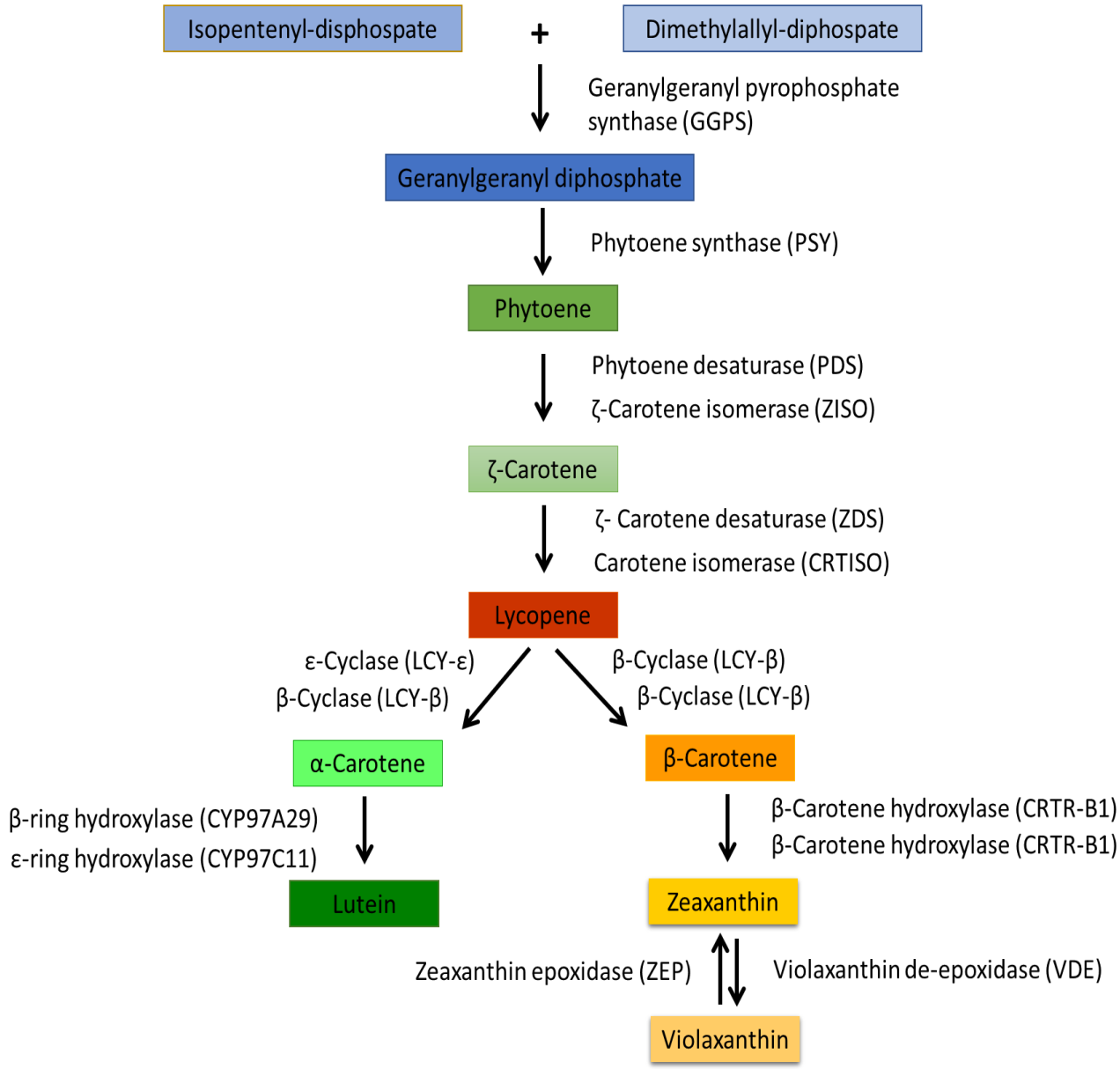


Figure 1.4. Carotenoid biosynthetic pathway in tomato chloroplasts. The enzyme prevalently involved in each step is indicated near the reaction arrow. CYP97A29, Cyt P450-type monooxygenase 97A29; CYP97C11, Cyt P450-type monooxygenase 97C11 (Stigliani et al., 2011).

As noted above, the synthesis of phytoene by phytoene synthase (PSY) is a rate-limiting step (Giuliano et al., 1993). During ripening, the expression of *PSY1* is significantly increased in response to ethylene (Vrebalov et al., 2002). *PSY1* is also a direct binding target of RIN (Martel et al., 2011). Predominantly lycopene accumulates in red tomato fruit due to the fact that the lycopene cyclases are repressed as the fruit matures, thus blocking further metabolism of lycopene (Ronen et al., 2000). This cyclase repression is controlled by ethylene (Alba et al., 2005).

Carotenoids in red fruit are stored in specialized plastids termed chromoplasts. During ripening, chloroplasts in green fruit are broken down and converted to chromoplasts and plastoglobules in these organelles are the site for carotenoid accumulation (Ytterberg et al., 2006). Studies in several mutants suggest that there may be a link between the chloroplast to chromoplast transition and ripening control beyond changes in pigmentation. For example, the *lutescent 1* (*l1*) and *lutescent 2* (*l2*) mutants of tomato are deficient in chlorophyll synthesis. Positional cloning revealed a mutation in a chloroplast targeted zinc metalloprotease that is homologous to the *Arabidopsis ETHYLENE-DEPENDENT GRAVITROPISM DEFICIENT AND YELLOW-GREEN1 (EGY1)* gene (Barry et al., 2012). The delayed ripening phenotypes in fruit may arise from an indirect consequence of an altered chloroplast derived signal that promotes the onset of ripening.

Light signal transduction also contributes to the regulation of carotenoid synthesis. In tomato, there are several mutants (*hp1*, *hp2*) with elevated levels of chlorophylls in green fruit and carotenoids in red fruit. Functional analysis of the mutants showed that the *high pigment1 (hp1)* mutation results from a lesion in a homolog of *UV-DAMAGED DNA BINDING PROTEIN1 (DDB1)* of *Arabidopsis,* while *hp2* is due to a mutation in tomato *DE-ETIOLATED1*

(*DET1*) (Mustilli et al., 1999; Liu et al., 2004). These genes are involved in the suppression of light responses in the absence of light by a molecular mechanism involving chromatin remodeling (Davuluri et al., 2005).

The taste of ripe tomatoes reflects a combination of sugars, acids and volatile compounds. α-Tomatine, a bitter taste compound found in green tomato that generally disappears during ripening is influenced by ethylene (Iijima et al., 2009). At the onset of ripening, starch is hydrolyzed to glucose and fructose by invertase. These two sugars alone can represent up to 5% of the fresh weight of a fruit and contribute substantially to final fruit taste attributes. Organic acids, principally citric and malic acids, contribute to the pH of the fruit. Ascorbic acid (vitamin C) is also abundant in many ripe fruits including tomato and plays an important role in the maintenance of the redox balance during fruit ripening through the removal of ROS. Ascorbic acid is synthesized mainly by the Wheeler–Smirnoff pathway from mannose-6-phosphate via GDP-mannose and GDP-l-galactose (Wheeler et al., 1998). A balance of sugars and acids contribute to flavor, although different individuals, cultures and cuisines reflect differing preferences for varying sweetness and acid balances.

Aroma and flavor volatiles are derived mainly from amino acids (e.g. Leu, Phe, Ile.), fatty acids (e.g. linolenic acid) and carotenoids (e.g. lycopene and β-carotene). Approximately 20-30 volatile compounds contribute to tomato flavor and consumer liking and some telegraph the presence of essential nutrients (Tieman et al., 2012). Ripening plays a key role in the synthesis of these compounds that usually increase at the onset of ripening and peak at full ripening, suggesting a highly regulated process. Perturbations in ethylene synthesis and perception, in addition to transcription factors such as RIN, affect volatile levels suggesting their production is ethylene and transcription factor dependent (Kovács et al., 2009).

While the biosynthetic pathways of many volatiles are increasingly being defined, regulation of volatile production is still poorly understood. Some important genes are known to participate in key steps in several pathways. For example, phenylalanine derived volatile accumulation is dependent on a small family of aromatic amino acid decarboxylases (*AADCs*). Expression of *AADCs* is upregulated during ripening, and increased expression of AADC enzymes results in increased flux into this volatile synthesis pathway (Tieman et al., 2006). Multiple six-carbon volatiles, e.g. cis-3-hexenal and hexanal, are synthesized from linoleic and linolenic acids. Several genes that involve in their synthesis, like 13-lipoxygenase (*LOXC*), hydroperoxide lyase, and alcohol dehydrogenase 2 (*ADH2*), are elevated in expression during ripening (Speirs et al., 1998; Kovács et al., 2009). In the case of some apocarotenoid volatiles, such as geranylacetone, pseudoionone, and β-ionone, although their synthesis significantly increases during ripening, expression of the biosynthetic genes including carotenoid cleavage dioxygenase genes (*CCD1A* and *CCD1B)* does not increase, suggesting that additional factors contribute to volatile accumulation (Simkin et al., 2004; Klee and Giovannoni, 2011).

In conclusion, tomato is an important fruit crop as well as a model species in plant biology. Basic and applied research in this system are often intertwined to shed light on aspects of fruit biology, including ethylene synthesis and perception, genetic regulation of ripening and mechanisms of fruit softening and accumulation of flavor and nutrients. However, there are many areas that remain poorly understood, especially the regulation of ripening and nutritional quality traits. Additional regulators remain to be identified and insights into specific ripening phenomenology, including the nature of obvious but minimally studied ripening gradients, are missing.

In this thesis, I focus on the regulation of pigmentation and fruit nutritional quality as a function of ripening. Chlorophyll and chloroplasts are present in green fruit and can influence attributes, including those important for consumption, in ripe fruit. For example, photosynthesis occurs in green tomato fruit and can contribute up to 20% of fruit photosynthate (Hetherington et al., 1998). During tomato ripening, there is a conversion of chloroplasts to chromoplasts, which are carotenoid-accumulating plastids in ripe fruit. Hence, chloroplast levels can affect chromoplast formation as well as storage capacity of nutritionally important carotenoids. This regulation of chloroplast development and its role in fruit nutritional quality is still not fully understood. The green shoulder phenotype in wild-type and the uniform green color due to the *uniform ripening* mutation (*u*) in unripe tomato fruit (Figure 2.1) present a genetic system to study these important phenomena. In wild-type, there is more green color at the shoulder (top) of the fruit compared to the bottom (stylar end) of the fruit. In *u* mutant fruit, this green shoulder is absent and fruit ripen more uniformly. This is important agronomically and for more than 70 years breeders have selected tomato varieties with uniformly light green fruit before ripening through breeding with *uniform* selection (Yeager 1935). The molecular basis of the trait as well as its function during fruit development and ripening are not known. Here I report the pursuit of the *u* mutation, its positional cloning and characterization, in additional to transcriptional profiling of distinct fruit tissues to shed light on outstanding regulatory mysteries of fruit development and maturation.

RERERENCES

Adams-Phillips, L., Barry, C., Kannan, P., Leclercq, J., Bouzayen, M., and Giovannoni, J. (2004). Evidence that CTR1-mediated ethylene signal transduction in tomato is encoded by a multigene family whose members display distinct regulatory features. Plant Molecular Biology 54: 387-404.

Adams, D., and Yang, S. (1979). Ethylene biosynthesis: identification of 1-aminocyclopropane-1-carboxylic acid as an intermediate in the conversion of methionine to ethylene. Proceedings of the National Academy of Sciences 76: 170-174.

Alba, R., Payton, P., Fei, Z.J., McQuinn, R., Debbie, P., Martin, G.B., Tanksley, S.D., and Giovannoni, J.J. (2005). Transcriptome and selected metabolite analyses reveal multiple points of ethylene control during tomato fruit development. Plant Cell 17: 2954-2965.

Atkinson, R.G., Gunaseelan, K., Wang, M.Y., Luo, L., Wang, T., Norling, C.L., Johnston, S.L., Maddumage, R., Schröder, R., and Schaffer, R.J. (2011). Dissecting the role of climacteric ethylene in kiwifruit (Actinidia chinensis) ripening using a 1-aminocyclopropane-1-carboxylic acid oxidase knockdown line. Journal of Experimental Botany 62: 3821-3835.

Ayub, R., Guis, M., Amor, M.B., Gillot, L., Roustan, J.-P., Latché, A., Bouzayen, M., and Pech, J.-C. (1996). Expression of ACC oxidase antisense gene inhibits ripening of cantaloupe melon fruits. Nature Biotechnology 14: 862-866.

Barry, C.S., and Giovannoni, J.J. (2006). Ripening in the tomato Green-ripe mutant is inhibited by ectopic expression of a protein that disrupts ethylene signaling. Proceedings of the National Academy of Sciences 103: 7923-7928.

Barry, C.S., Blume, B., Bouzayen, M., Cooper, W., Hamilton, A.J., and Grierson, D. (1996). Differential expression of the 1-aminocyclopropane-1-carboxylate oxidase gene family of tomato. Plant Journal 9: 525-535.

Barry, C.S., Aldridge, G.M., Herzog, G., Ma, Q., McQuinn, R.P., Hirschberg, J., and Giovannoni, J.J. (2012). Altered chloroplast development and delayed fruit ripening caused by mutations in a zinc metalloprotease at the lutescent2 locus of tomato. Plant Physiology 159: 1086-1098.

Bemer, M., Karlova, R., Ballester, A.R., Tikunov, Y.M., Bovy, A.G., Wolters-Arts, M., Rossetto Pde, B., Angenent, G.C., and de Maagd, R.A. (2012). The tomato FRUITFULL homologs TDR4/FUL1 and MBP7/FUL2 regulate ethylene-independent aspects of fruit ripening. Plant Cell 24: 4437-4451.

Biale, J.B. (1960). Respiration of fruits. Encyclopedia of Plant Physiology 12: 536-586.

Blankenship, S.M., and Dole, J.M. (2003). 1-Methylcyclopropene: a review. Postharvest Biology and Technology 28: 1-25.

Bleecker, A.B., and Kende, H. (2000). Ethylene: a gaseous signal molecule in plants. Annual Review of Cell and Developmental Biology 16: 1-18.

Brummell, D.A. (2006). Cell wall disassembly in ripening fruit. Functional Plant Biology 33, 103-119.

Causse, R.D., P. Rousselle (2007). Traditional and enhanced breeding for quality traits in tomato. In Genetic improvement of Solanaceous crops, Tomato, A.K.M. M.K. Razdan, ed (Enfield, NH: Science Publishers), pp. 153-192.

Chen, G., Alexander, L., and Grierson, D. (2004). Constitutive expression of EIL-like transcription factor partially restores ripening in the ethylene-insensitive Nr tomato mutant. Journal of Experimental Botany 55: 1491-1497.

Chung, M.Y., Vrebalov, J., Alba, R., Lee, J., McQuinn, R., Chung, J.D., Klein, P., and Giovannoni, J. (2010). A tomato (Solanum lycopersicum) APETALA2/ERF gene, SlAP2a, is a negative regulator of fruit ripening. Plant Journal 64: 936-947.

Cong, B., Barrero, L.S., and 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.

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

Cosgrove, D.J. (2005). Growth of the plant cell wall. Nature Reviews Molecular Cell Biology 6: 850-861.

Davuluri, G.R., van Tuinen, A., Fraser, P.D., Manfredonia, A., Newman, R., Burgess, D., Brummell, D.A., King, S.R., Palys, J., Uhlig, J., Bramley, P.M., Pennings, H.M., and Bowler, C. (2005). Fruit-specific RNAi-mediated suppression of DET1 enhances carotenoid and flavonoid content in tomatoes. Nature Biotechnology 23: 890-895.

Dong, J.G., Fernandez-Maculet, J.C., and Yang, S.F. (1992). Purification and characterization of 1-aminocyclopropane-1-carboxylate oxidase from apple fruit. Proceedings of the National Academy of Sciences 89: 9789-9793.

Elitzur, T., Vrebalov, J., Giovannoni, J.J., Goldschmidt, E.E., and Friedman, H. (2010). The regulation of MADS-box gene expression during ripening of banana and their regulatory interaction with ethylene. Journal of Experimental Botany 61: 1523-1535.

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

Fujisawa, M., Nakano, T., and Ito, Y. (2011). Identification of potential target genes for the tomato fruit-ripening regulator RIN by chromatin immunoprecipitation. BMC Plant Biology 11: 26.

Gagne, J.M., Smalle, J., Gingerich, D.J., Walker, J.M., Yoo, S.-D., Yanagisawa, S., and Vierstra, R.D. (2004). Arabidopsis EIN3-binding F-box 1 and 2 form ubiquitin-protein ligases that repress ethylene action and promote growth by directing EIN3 degradation. Proceedings of the National Academy of Sciences 101: 6803-6808.

Gillaspy, G., Ben-David, H., and Gruissem, W. (1993). Fruits: A Developmental Perspective. Plant Cell 5: 1439-1451.

Giovannoni, J.J. (2007). Fruit ripening mutants yield insights into ripening control. Current Opinion in Plant Biology 10: 283-289.

Giovannoni, J.J., DellaPenna, D., Bennett, A.B., and Fischer, R.L. (1989). Expression of a chimeric polygalacturonase gene in transgenic rin (ripening inhibitor) tomato fruit results in polyuronide degradation but not fruit softening. Plant Cell 1: 53-63.

Giuliano, G., Bartley, G.E., and Scolnik, P.A. (1993). Regulation of carotenoid biosynthesis during tomato development. Plant Cell 5: 379-387.

Guo, H., and Ecker, J.R. (2003). Plant Responses to Ethylene Gas Are Mediated by SCF< sup> EBF1/EBF2</sup>-Dependent Proteolysis of EIN3 Transcription Factor. Cell 115: 667-677.

Guo, H., and Ecker, J.R. (2004). The ethylene signaling pathway: new insights. Current Opinion in Plant Biology 7: 40-49.

Hamilton, A., Lycett, G., and Grierson, D. (1990). Antisense gene that inhibits synthesis of the hormone ethylene in transgenic plants. Nature 346: 284-287.

Hetherington, S.E., Smillie, R.M., and Davies, W.J. (1998). Photosynthetic activities of vegetative and fruiting tissues of tomato. Journal of Experimental Botany 49: 1173-1181.

Hirschberg, J. (2001). Carotenoid biosynthesis in flowering plants. Current Opinion in Plant Biology 4: 210-218.

Hua, J., and Meyerowitz, E.M. (1998). Ethylene Responses Are Negatively Regulated by a Receptor Gene Family in Arabidopsis thaliana. Cell 94: 261-271.

Iijima, Y., Fujiwara, Y., Tokita, T., Ikeda, T., Nohara, T., Aoki, K., and Shibata, D. (2009). Involvement of ethylene in the accumulation of esculeoside A during fruit ripening of tomato (Solanum lycopersicum). Journal of Agricultural and Food Chemistry 57: 3247-3252.

Isaacson, T., Kosma, D.K., Matas, A.J., Buda, G.J., He, Y., Yu, B., Pravitasari, A., Batteas, J.D., Stark, R.E., and Jenks, M.A. (2009). Cutin deficiency in the tomato fruit cuticle consistently affects resistance to microbial infection and biomechanical properties, but not transpirational water loss. Plant Journal 60: 363-377.

Itkin, M., Seybold, H., Breitel, D., Rogachev, I., Meir, S., and Aharoni, A. (2009). TOMATO AGAMOUS-LIKE 1 is a component of the fruit ripening regulatory network. Plant Journal 60: 1081-1095.

Jaakola, L., Poole, M., Jones, M.O., Kämäräinen-Karppinen, T., Koskimäki, J.J., Hohtola, A., Häggman, H., Fraser, P.D., Manning, K., and King, G.J. (2010). A SQUAMOSA MADS box gene involved in the regulation of anthocyanin accumulation in bilberry fruits. Plant Physiology 153: 1619-1629.

Ju, C., Yoon, G.M., Shemansky, J.M., Lin, D.Y., Ying, Z.I., Chang, J., Garrett, W.M., Kessenbrock, M., Groth, G., Tucker, M.L., Cooper, B., Kieber, J.J., and Chang, C. (2012). CTR1 phosphorylates the central regulator EIN2 to control ethylene hormone signaling from the

ER membrane to the nucleus in Arabidopsis. Proceedings of the National Academy of Sciences 109: 19486-19491.

Karlova, R., Rosin, F.M., Busscher-Lange, J., Parapunova, V., Do, P.T., Fernie, A.R., Fraser, P.D., Baxter, C., Angenent, G.C., and de Maagd, R.A. (2011). Transcriptome and metabolite profiling show that APETALA2a is a major regulator of tomato fruit ripening. Plant Cell 23: 923-941.

Kevany, B.M., Tieman, D.M., Taylor, M.G., Cin, V.D., and Klee, H.J. (2007). Ethylene receptor degradation controls the timing of ripening in tomato fruit. Plant Journal 51: 458-467.

Klee, H.J. (2004). Ethylene signal transduction. Moving beyond Arabidopsis. Plant Physiology 135: 660-667.

Klee, H.J., and Giovannoni, J.J. (2011). Genetics and control of tomato fruit ripening and quality attributes. Annual Review of Genetics 45: 41-59.

Kovács, K., Fray, R.G., Tikunov, Y., Graham, N., Bradley, G., Seymour, G.B., Bovy, A.G., and Grierson, D. (2009). Effect of tomato pleiotropic ripening mutations on flavour volatile biosynthesis. Phytochemistry 70: 1003-1008.

Kumar, R., Tyagi, A.K., and Sharma, A.K. (2011). Genome-wide analysis of auxin response factor (ARF) gene family from tomato and analysis of their role in flower and fruit development. Molecular Genetics and Genomics 285: 245-260.

Kumar, R., Sharma, M.K., Kapoor, S., Tyagi, A.K., and Sharma, A.K. (2012). Transcriptome analysis of rin mutant fruit and in silico analysis of promoters of differentially regulated genes provides insight into LeMADS-RIN-regulated ethylene-dependent as well as ethylene-independent aspects of ripening in tomato. Molecular Genetics and Genomics 287: 189-203.

Lanahan, M.B., Yen, H.-C., Giovannoni, J.J., and Klee, H.J. (1994). The never ripe mutation blocks ethylene perception in tomato. Plant Cell 6: 521-530.

Leclercq, J., Adams-Phillips, L.C., Zegzouti, H., Jones, B., Latché, A., Giovannoni, J.J., Pech, J.-C., and Bouzayen, M. (2002). LeCTR1, a tomato CTR1-like gene, demonstrates ethylene signaling ability in Arabidopsis and novel expression patterns in tomato. Plant Physiology 130: 1132-1142.

Lee, J.M., Joung, J.G., McQuinn, R., Chung, M.Y., Fei, Z., Tieman, D., Klee, H., and Giovannoni, J. (2012). Combined transcriptome, genetic diversity and metabolite profiling in tomato fruit reveals that the ethylene response factor SlERF6 plays an important role in ripening and carotenoid accumulation. Plant Journal 70: 191-204.

Liljegren, S.J., Ditta, G.S., Eshed, Y., Savidge, B., Bowman, J.L., and Yanofsky, M.F. (2000). SHATTERPROOF MADS-box genes control seed dispersal in Arabidopsis. Nature 404: 766-770.

Lin, Z., Hong, Y., Yin, M., Li, C., Zhang, K., and Grierson, D. (2008). A tomato HD-Zip homeobox protein, LeHB-1, plays an important role in floral organogenesis and ripening. Plant Journal 55: 301-310.

Liu, J., Van Eck, J., Cong, B., and Tanksley, S.D. (2002). A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. Proceedings of the National Academy of Sciences 99: 13302-13306.

Liu, Y., Roof, S., Ye, Z., Barry, C., van Tuinen, A., Vrebalov, J., Bowler, C., and Giovannoni, J. (2004). Manipulation of light signal transduction as a means of modifying fruit nutritional quality in tomato. Proceedings of the National Academy of Sciences 101: 9897-9902.

Mahajan, P.V., Caleb, O.J., Singh, Z., Watkins, C.B., and Geyer, M. (2014). Postharvest treatments of fresh produce. Philosophical Transactions of the Royal Society A 372:20130309.

Manning, K., Tör, M., Poole, M., Hong, Y., Thompson, A.J., King, G.J., Giovannoni, J.J., and Seymour, G.B. (2006). A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. Nature Genetics 38: 948-952.

Martel, C., Vrebalov, J., Tafelmeyer, P., and Giovannoni, J.J. (2011). The tomato MADS-box transcription factor RIPENING INHIBITOR interacts with promoters involved in numerous ripening processes in a COLORLESS NONRIPENING-dependent manner. Plant Physiology 157: 1568-1579.

Martineau, B. (2001). First fruit: the creation of the Flavr Savr tomato and the birth of genetically engineered food. (McGraw-Hill Companies).

Matas, A.J., Yeats, T.H., Buda, G.J., Zheng, Y., Chatterjee, S., Tohge, T., Ponnala, L., Adato, A., Aharoni, A., Stark, R., Fernie, A.R., Fei, Z.J., Giovannoni, J.J., and Rose, J.K.C. (2011). Tissue- and Cell-Type Specific Transcriptome Profiling of Expanding Tomato Fruit Provides Insights into Metabolic and Regulatory Specialization and Cuticle Formation. Plant Cell 23: 3893-3910.

McMurchie, E., McGlasson, W., and Eaks, I. (1972). Treatment of fruit with propylene gives information about the biogenesis of ethylene. Nature 237: 235-236

Mustilli, A.C., Fenzi, F., Ciliento, R., Alfano, F., and Bowler, C. (1999). Phenotype of the tomato high pigment-2 mutant is caused by a mutation in the tomato homolog of DEETIOLATED1. Plant Cell 11: 145-157.

Nakatsuka, A., Murachi, S., Okunishi, H., Shiomi, S., Nakano, R., Kubo, Y., and Inaba, A. (1998). Differential expression and internal feedback regulation of 1-aminocyclopropane-1-carboxylate synthase, 1-aminocyclopropane-1-carboxylate oxidase, and ethylene receptor genes in tomato fruit during development and ripening. Plant Physiology 118: 1295-1305.

O'Malley, R.C., Rodriguez, F.I., Esch, J.J., Binder, B.M., O'Donnell, P., Klee, H.J., and Bleecker, A.B. (2005). Ethylene-binding activity, gene expression levels, and receptor system output for ethylene receptor family members from Arabidopsis and tomato. Plant Journal 41: 651-659.

Potuschak, T., Lechner, E., Parmentier, Y., Yanagisawa, S., Grava, S., Koncz, C., and Genschik, P. (2003). EIN3-Dependent Regulation of Plant Ethylene Hormone Signaling by Two Arabidopsis F Box Proteins: EBF1 and EBF2. Cell 115: 679-689.

Qin, G., Wang, Y., Cao, B., Wang, W., and Tian, S. (2012). Unraveling the regulatory network of the MADS box transcription factor RIN in fruit ripening. Plant Journal 70: 243-255.

Resnick, J.S., Wen, C.-K., Shockey, J.A., and Chang, C. (2006). REVERSION-TO-ETHYLENE SENSITIVITY1, a conserved gene that regulates ethylene receptor function in Arabidopsis. Proceedings of the National Academy of Sciences 103: 7917-7922.

Rodriguez, F.I., Esch, J.J., Hall, A.E., Binder, B.M., Schaller, G.E., and Bleecker, A.B. (1999). A copper cofactor for the ethylene receptor ETR1 from Arabidopsis. Science 283: 996-998.

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

Ronen, G., Carmel-Goren, L., Zamir, D., and Hirschberg, J. (2000). An alternative pathway to β-carotene formation in plant chromoplasts discovered by map-based cloning of Beta and old-gold color mutations in tomato. Proceedings of the National Academy of Sciences 97: 11102-11107.

Rose, J.K., Lee, H.H., and Bennett, A.B. (1997). Expression of a divergent expansin gene is fruit-specific and ripening-regulated. Proceedings of the National Academy of Sciences 94: 5955-5960.

Rottmann, W.H., Peter, G.F., Oeller, P.W., Keller, J.A., Shen, N.F., Nagy, B.P., Taylor, L.P., Campbell, A.D., and Theologis, A. (1991). 1-Aminocyclopropane-1-carboxylate synthase in tomato is encoded by a multigene family whose transcription is induced during fruit and floral senescence. Journal of Molecular Biology 222: 937-961.

Seymour, G.B., Ryder, C.D., Cevik, V., Hammond, J.P., Popovich, A., King, G.J., Vrebalov, J., Giovannoni, J.J., and Manning, K. (2011). A SEPALLATA gene is involved in the development and ripening of strawberry (Fragaria× ananassa Duch.) fruit, a non-climacteric tissue. Journal of Experimental Botany 62: 1179-1188.

Shi, J.X., Adato, A., Alkan, N., He, Y., Lashbrooke, J., Matas, A.J., Meir, S., Malitsky, S., Isaacson, T., and Prusky, D. (2013). The tomato SlSHINE3 transcription factor regulates fruit cuticle formation and epidermal patterning. New Phytologist 197: 468-480.

Simkin, A.J., Schwartz, S.H., Auldridge, M., Taylor, M.G., and Klee, H.J. (2004). The tomato carotenoid cleavage dioxygenase 1 genes contribute to the formation of the flavor volatiles β-ionone, pseudoionone, and geranylacetone. Plant Journal 40: 882-892.

Smith, C., Watson, C., Bird, C., Ray, J., Schuch, W., and Grierson, D. (1990). Expression of a truncated tomato polygalacturonase gene inhibits expression of the endogenous gene in transgenic plants. Molecular and General Genetics 224: 477-481.

Speirs, J., Lee, E., Holt, K., Yong-Duk, K., Scott, N.S., Loveys, B., and Schuch, W. (1998). Genetic manipulation of alcohol dehydrogenase levels in ripening tomato fruit affects the balance of some flavor aldehydes and alcohols. Plant Physiology 117: 1047-1058.

Stigliani, A.L., Giorio, G., and D'Ambrosio, C. (2011). Characterization of P450 carotenoid beta- and epsilon-hydroxylases of tomato and transcriptional regulation of xanthophyll biosynthesis in root, leaf, petal and fruit. Plant and Cell Physiology 52, 851-865.

Tanksley, S.D. (2004). The genetic, developmental, and molecular bases of fruit size and shape variation in tomato. Plant Cell 16: S181-S189.

Tatsuki, M., and Mori, H. (2001). Phosphorylation of tomato 1-aminocyclopropane-1-carboxylic acid synthase, LE-ACS2, at the C-terminal region. Journal of Biological Chemistry 276: 28051-28057.

Tieman, D., Taylor, M., Schauer, N., Fernie, A.R., Hanson, A.D., and Klee, H.J. (2006). Tomato aromatic amino acid decarboxylases participate in synthesis of the flavor volatiles 2-phenylethanol and 2-phenylacetaldehyde. Proceedings of the National Academy of Sciences 103: 8287-8292.

Tieman, D., Bliss, P., McIntyre, L.M., Blandon-Ubeda, A., Bies, D., Odabasi, A.Z., Rodríguez, G.R., van der Knaap, E., Taylor, M.G., and Goulet, C. (2012). The chemical interactions underlying tomato flavor preferences. Current Biology 22: 1035-1039.

Tieman, D.M., Taylor, M.G., Ciardi, J.A., and Klee, H.J. (2000). The tomato ethylene receptors NR and LeETR4 are negative regulators of ethylene response and exhibit functional compensation within a multigene family. Proceedings of the National Academy of Sciences 97: 5663-5668.

Tieman, D.M., Ciardi, J.A., Taylor, M.G., and Klee, H.J. (2001). Members of the tomato LeEIL (EIN3-like) gene family are functionally redundant and regulate ethylene responses throughout plant development. Plant Journal 26: 47-58.

Tomato Genome, C. (2012). The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485: 635-641.

Treutter, D. (2005). Significance of flavonoids in plant resistance and enhancement of their biosynthesis. Plant Biology 7: 581-591.

Vicente, A.R., Saladie, M., Rose, J.K., and Labavitch, J.M. (2007). The linkage between cell wall metabolism and fruit softening: looking to the future. Journal of the Science of Food and Agriculture 87: 1435-1448.

Vrebalov, J., Ruezinsky, D., Padmanabhan, V., White, R., Medrano, D., Drake, R., Schuch, W., and Giovannoni, J. (2002). A MADS-box gene necessary for fruit ripening at the tomato ripening-inhibitor (rin) locus. Science 296: 343-346.

Vrebalov, J., Pan, I.L., Arroyo, A.J., McQuinn, R., Chung, M., Poole, M., Rose, J., Seymour, G., Grandillo, S., Giovannoni, J., and Irish, V.F. (2009). Fleshy fruit expansion and ripening are regulated by the Tomato SHATTERPROOF gene TAGL1. Plant Cell 21: 3041-3062.

Watkins, C.B. (2006). The use of 1-methylcyclopropene (1-MCP) on fruits and vegetables. Biotechnology Advances 24, 389-409.

Watkins, C.B. (2008). Overview of 1-methylcyclopropene trials and uses for edible horticultural crops. Hortscience 43, 86-94.

Wheeler, G.L., Jones, M.A., and Smirnoff, N. (1998). The biosynthetic pathway of vitamin C in higher plants. Nature 393: 365-369.

Wilkinson, J.Q., Lanahan, M.B., Yen, H.C., Giovannoni, J.J., and Klee, H.J. (1995). An ethylene-inducible component of signal transduction encoded by never-ripe. Science 270: 1807-1809.

Xiao, H., Jiang, N., Schaffner, E., Stockinger, E.J., and van der Knaap, E. (2008). A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit. Science 319: 1527-1530.

Yeager, T. (1935). The uniform fruit color gene in the tomato. Proceedings of the American Society for Horticultural Science 33: 512.

Yeats, T.H., Martin, L.B., Viart, H.M., Isaacson, T., He, Y., Zhao, L., Matas, A.J., Buda, G.J., Domozych, D.S., and Clausen, M.H. (2012). The identification of cutin synthase: formation of the plant polyester cutin. Nature Chemical Biology 8: 609-611.

Yen, H.-C., Lee, S., Tanksley, S.D., Lanahan, M.B., Klee, H.J., and Giovannoni, J.J. (1995). The tomato Never-ripe locus regulates ethylene-inducible gene expression and is linked to a homolog of the Arabidopsis ETR1 gene. Plant Physiology 107: 1343-1353.

Yokotani, N., Tamura, S., Nakano, R., Inaba, A., and Kubo, Y. (2003). Characterization of a novel tomato EIN3-like gene (LeEIL4). Journal of Experimental Botany 54: 2775-2776.

Ytterberg, A.J., Peltier, J.-B., and Van Wijk, K.J. (2006). Protein profiling of plastoglobules in chloroplasts and chromoplasts. A surprising site for differential accumulation of metabolic enzymes. Plant Physiology 140: 984-997.

Zhong, S., Lin, Z., and Grierson, D. (2008). Tomato ethylene receptor–CTR interactions: visualization of NEVER-RIPE interactions with multiple CTRs at the endoplasmic reticulum. Journal of Experimental Botany 59: 965-972.

Zhong, S., Fei, Z., Chen, Y.R., Zheng, Y., Huang, M., Vrebalov, J., McQuinn, R., Gapper, N., Liu, B., Xiang, J., Shao, Y., and Giovannoni, J.J. (2013). Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening. Nature Biotechnology 31: 154-159.

CHAPTER II.

MOLECULAR BASIS OF GREEN SHOULDER/UNIFORM RIPENING PHENOTYPE IN TOMATO AND ITS ROLE IN FRUIT NUTRITIONAL QUALITY

Abstract

The tomato *uniform ripening* (*u*) mutation confers a light green fruit phenotype absent the wild-type green shoulder and is widely used in breeding for tomato varieties producing evenly ripened fruit to facilitate determination of harvest time and sale of visually appealing fresh fruit. Positional cloning of u revealed that *U* encodes a *Golden 2-like (GLK)* transcription factor, SlGLK2, influencing chlorophyll accumulation and distribution in developing fruit. Two GLKs could be identified in the tomato genome sequence — *SlGLK1* and *SlGLK2*. Both genes are expressed in leaves, but only *SlGLK2* is expressed in fruit. Transgene-mediated ectopic expression of either *SlGLK1* or *SlGLK2* increased fruit chlorophyll content, while suppressing *SlGLK2* mimicked the *u* mutant fruit phenotype and *SlGLK1* suppression yielded plants with pale leaves and no notable impact on fruit. These results reflect a degree of functional redundancy of SlGLK1 and SlGLK2 and the more important role of SlGLK1 in leaves and SlGLK2 in fruit. Further characterization of our transgenic lines indicated that *GLK* over-expression up-regulated fruit photosynthesis related genes increasing chlorophyll content, chloroplast numbers and thylakoid grana stacks. Ectopic *GLK* expression also elevated starch in green fruit and sugars, carotenoids and ascorbate in ripe fruit. I crossed my *SlGLK2* over-

expression lines with the tomato *high pigment 1* (*hp1*) mutant, which has elevated chlorophyll and carotenoid content in fruit due to a loss of tomato *DDB1 (DAMAGED DNA BINDING PROTEIN 1)* activity and resulting *SlGLK2*;*hp1/hp1* lines indicate that SlGLK2 acts independently of DDB1. Combining *SlGLK2* overexpression and *hp1* elevated ripe fruit carotenoids. *SlGLK2* influences photosynthesis in developing fruit, contributes to mature fruit quality characteristics and the results presented here suggest that manipulating *GLK* expression is a means to enhance fruit quality and nutritional value.

I. Introduction

The fruit is an important agricultural plant organ, representing an outstanding source of vitamins, minerals, antioxidants, and fiber for humans and animals. Biologically, fruits develop from carpels (true fruit) or other floral-associated tissues (false or accessory fruit) to protect developing seeds and serve as vehicles for seed dispersal. For fleshy fruits, the ripening process renders the fruit more attractive and palatable for seed dispersing organisms. During ripening, physiological and biochemical changes that make the fruit more attractive occur; these typically include changes in color (loss of chlorophyll and accumulation of pigments), texture (tissue softening), sugar content (conversion of starch to sugar), organic acid and flavor/aroma volatiles contents (Giovannoni, 2004; Klee and Giovannoni, 2011; Gapper et al., 2013).

Uniform green color of unripe tomato fruit, as contrasted with darker green stem end fruit color was first described by (MacArthur, 1928). The uniform fruit color phenotype results from a single, recessive mutation called *u* (for *uniform ripening*) that removes the typical green shoulder at the stem end of the fruit and allows fruit to ripen more uniformly and mitigates the appearance of yellow or white patches of color in some otherwise red ripe tomato fruit (Figure 2.1). The economic value of uniform color in promoting even ripening of fruit was first discussed by

Yeager (1935) who promoted introduction of the *u* gene into commercial varieties as a means of improving market quality. This phenotype also facilitates the determination of optimal timing for destructive harvest, which is common for processing cultivars (Figure 2.1). Another mutation conferring fruit uniform color is called *ug* for *uniform globe* and was first described by Bohn and Scott (1945). Genetic test crosses showed that both genes are recessive. Crosses between homozygous *u/u* and *ug/ug* yield F1 plants with the normal green shoulder phenotype and F2 plants segregated 9:7 for the green shoulder phenotype indicating that the two genes are genetically unlinked (expected ratio 1:1).

Though introduction of the *u* gene into commercial varieties was initiatied nearly a century ago, light green fruit may ripen with reduced sugars, compromising traits that are valuable for processed products and the flavor of fresh fruit (Tanksley and Hewitt, 1988).

Fruit are typically "sink" tissues accumulating high energy reduced carbon compounds derived and transported primarily from leaf "source" tissue. Photosynthesis occurs in green tomato fruit as well and up to 20% of fruit photosynthate is endogenously derived (Hetherington et al., 1998). While fruits from numerous species develop as green photosynthetic tissues (Blanke and Lenz, 1989; Gillaspy et al., 1993; Carrara et al., 2001), the regulation of chloroplast development and photosynthesis in developing and maturing fruit is not fully understood.

The nature of the *u* mutant phenotype suggests that the underlying gene may participate in chlorophyll and/or chloroplast distribution. Two Arabidopsis *GARP* family *Myb* transcription factors [*Golden 2-like* (*GLK1*) and *GLK2*] determine the capacity for light-stimulated photosynthesis by controlling chloroplast formation (Rossini et al., 2001; Fitter et al., 2002; Nakamura et al., 2009). Arabidopsis *glk1 glk2* double mutants are pale green and contain chloroplasts with non-stacked thylakoids and reduced levels of photosynthetic electron transport

complexes (Fitter et al, 2002). Transgenic and chromatin immunoprecipitation analyses showed that GLK transcription factors act directly on the promoters of genes encoding LHC proteins, especially those of LHCII, and key enzymes of the chlorophyll biosynthetic pathway (Waters et al, 2009). Proteins that promote chlorophyll synthesis and LHC assembly like GLKs represent logical candidates for the *u* gene.

Tomato is a major source of vitamins, minerals, antioxidants (e.g. carotenoids) and folic acid in human diets. However, tomato remains a commodity where most growers do not receive added value for elevated nutritional or taste/aroma quality. As such, breeding programs focus on primary production concerns such as yield, disease resistance and appearance while neglecting organoleptic and nutritional properties. This trend is changing in recent years as consumer preferences have become more discerning, such that when breeders are presented with molecular tools that permit selection for superior flavor and nutrition quality, they certainly use them.

Although many national and international dietary improvement programs focus on caloric intake, an equally pervasive problem worldwide is micronutrient malnutrition. Chronic diseases like diabetes, heart disease and cancer are primary public health concerns. Diet can play important roles in these diseases and fruits are the main source of micronutrients in many human diets. Tomato is the highest value and one of the most important vegetable crops in the world (FAOSTAT, http://faostat3.fao.org/home/index.html). As such, any increase in the nutritional quality of tomato will have a major impact world-wide.

This chapter describes the positional cloning of the *uniform ripening* (*u*) locus as well as molecular analysis of the underlying gene and its mutant allele and assesses their roles in fruit development, appearance and nutritional quality.

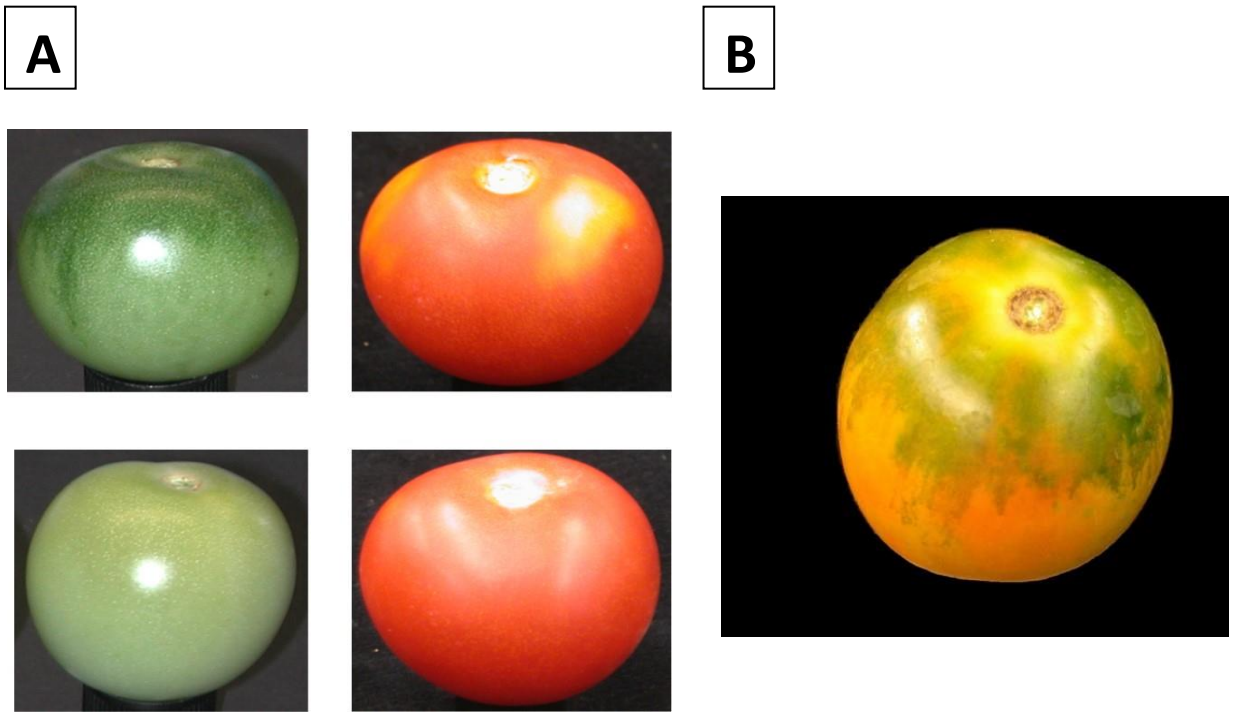


Figure 2.1. Green shoulder *(U/U)* and uniform ripening *(u/u)* fruit phenotype.

A. Top row: Wild type green shoulder *U/U* green and red ripe fruit. *U/U* fruit are more prone to development of yellow spots at the shoulder when they turn red. Bottom row: *uniform u/u* mutant green and red fruit

B. *U/U* early ripening fruit displaying significant chlorophyll on the shoulder. This pigmentation may obscure the determination of optimal timing for destructive harvest.

II. Results

Fine mapping of u locus on chromosome 10

Classical genetic studies in tomato placed *u* on the short arm of chromosome 10 (Tanksley et al., 1990). The *S. lycopersicum* cv. *'M82"* x *S. pennellii* acc. 'LA0716' introgression line (IL) population (Eshed and Zamir, 1995) and an F_2 population of 1100 individuals derived from a backcross of *S. pennellii* IL 10-1 (*U/U*) to its recurrent parent *'M82"* (*u/u*) were used for low and higher resolution mapping, respectively. Initial analysis positioned *u* between markers TG303 (SL2.40ch10:1773625) and CT234 (SL2.40ch10:2641027). Additional CAPS markers more precisely mapped *U* to between markers B (SL2.40ch10:2275056) and C (Sl2.40ch10:2358687) (Figure 2.2).

An independent mapping effort from our collaborators at the University of Valencia (A. Granell) using a BC_2S_1 population of 40 individuals and a recombinant inbred line (RIL) population of 110 individuals from *S. lycopersicum* cv. 'Moneymaker' x *S. pimpinellifolium* (acc. 'TO-937') were also phenotyped for fruit with dark green shoulders (Figure 2.2) and they were genotyped with markers saturating the end of chromosome 10. Linkage analysis using the overlapping minimal regions flanking the *u* locus from the three populations narrowed *u* to the region of chromosome 10 between markers B (SL2.40ch10:2275056) and solcap_snp_sl_17859 (SL2.40ch10:2335463) (Figure 2.2).

This 60,507 bp region on the short arm of chromosome 10 contains eight predicted genes, including a *Golden 2-like* (*GLK2*) transcription factor located at position Sl2.40chr10:2291209-2295578 (Figure 2.2, Table 2.1). GLK transcription factors belong to the GARP family, members of the Myb superfamily, and function in chlorophyll biosynthesis and chloroplast development (Rossini et al., 2001; Fitter et al., 2002; Nakamura et al., 2009; Waters

et al., 2009). As noted above, the green shoulder phenotype likely results from elevated chlorophyll and/or chloroplasts in the shoulder. Thus the tomato *GLK2* presented an excellent candidate for the gene underlying the *uniform ripening* mutation. Indeed, DNA sequencing revealed presence of a single allele harboring a strong mutation in the *GLK* gene in all *u* lines. Sequencing full-length *SlGLK2* transcripts predicted that in *U* genotypes, *SlGLK2* encodes a 310 amino acid protein, but in *u, Slglk2* encodes a truncated 80 amino acid protein,

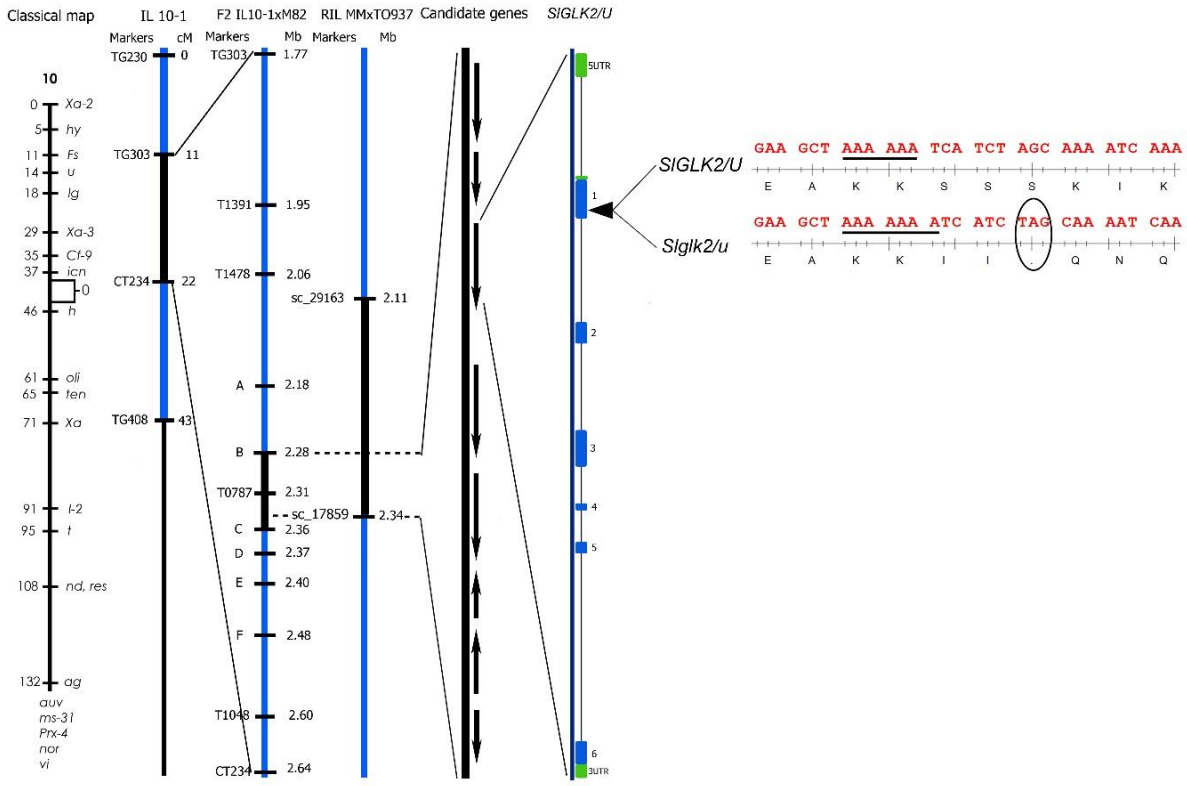


Figure 2.2. Map position and sequence of *SlGLK2* alleles. A. Left to right, the classical morphological map, IL map, maps derived from the IL 10-1 F2 and the MM x 'T0937' populations, candidate genes in the 60,507 bp region defined by markers and gene model of SlGLK2 with the additional A underlined and the resulting stop codon in the *Slglk2/u* allele circled.

Table 1. Eight predicted genes in the 60,507 bp region of *S. lycopersicum* chromosome 10 (ITAG tomato genome annotation release 2.4; http://solgenomics.net) the region defined genetically as containing the *u* locus spans SL2.40ch10:2275056 and SL2.40ch10:2335463. *SlGLK2* (yellow highlight) is within this region, specifically between Sl2.40chr10:2291209 and Sl240chr10:2295578.

Start	Stop	Gene ID	Identifier
2276303	2278610	Solyc10g008140	Unidentified, length=2308
2281545	2281874	Solyc10g008150	Glutaredoxin (AHRD V1 ***- B9MYC1_POPTR); B contains Interpro domain(s) IPR011905 Glutaredoxin-like plant II
2293088	2295945	Solyc10g008160	Transcription factor (Fragment) *Slglk2/u* (AHRD V1 *--- D6MK15_9ASPA)
2300710	2304760	Solyc10g008170	26S proteasome regulatory subunit (AHRD V1 ***- C6HL17_AJECH)
2305283	2311789	Solyc10g008180	26S proteasome regulatory subunit (AHRD V1 ***- A8J3A4_CHLRE); B contains Interpro domain(s) IPR016643 26S proteasome regulatory complex non-ATPase subcomplex Rpn1 subunit
2311949	2314392	Solyc10g008190	OB-fold nucleic acid binding domain containing protein (AHRD V1 ***- B6SHT0_MAIZE); B contains Interpro domain(s) IPR012340 Nucleic acid-binding OB-fold
2315211	2319689	Solyc10g008200	Tyrosine aminotransferase (AHRD V1 **** D3K4J1_PAPSO); B contains Interpro domain(s) IPR005958 Tyrosinenicotianamine aminotransferase
2325581	2332352	Solyc10g008210	Os07g0507200 protein (Fragment) homolog (AHRD V1 *-*- C7J530_ORYSJ); B contains Interpro domain(s) IPR009675 Targeting for Xklp2

due to a single base insertion causing a frame-shift and premature stop codon (Figure 2.2). A panel of diverse germplasm was used to identify polymorphisms in the *u* locus. Sequences of the *u* locus of cultivated tomato varieties from US, Europe and Asia which carry either the *u* mutation or present the wild type green shoulder phenotype together with tomato wild species such as *S. pimpinellifolium*, *S. pennellii* and *S. lycopersicoides* were analyzed. Those with the light green uniform phenotype had the additional adenine (A) between SL2.40ch10:2292260-2292267 and is the only sequence polymorphism observed in the *SlGLK2* sequence that is common to all light green *u/u* varieties and absent in all dark green shouldered *U/U* varieties (Figure 2.3). This result suggests that all *u* allele in modern tomato varieties is derived from a single mutation event identified early in the 20th century. The amino acid sequences of the *SlGLK2* gene and a GLK homolog (*SlGLK1)* are similar to other plant GLKs containing two conserved domains: a *myb*-like DNA binding domain and a novel C-terminal domain. *SlGLKs* are approximately 45% identical to their *Arabidopsis* counterparts (Figure 2.4).

SlGLK2 mRNA is more abundant in the pedicel (shoulder) than in the blossom (stylar) end of the fruit, suggesting that *SlGLK2* contributes to the pattern and intensity of chlorophyll accumulation (Figure 2.5). In *u/u* fruit, *Slglk2* expression was significantly lower than *SlGLK2* levels in wild type (*U/U*) (Figure 2.5). Wild type fruit that developed in the dark or were shaded are pale (Figure 3.3) suggesting that light is essential for fruit chloroplast development and chlorophyll synthesis, but their pattern and intensity is determined by *SlGLK2*.

We tested whether the *u* phenotype is altered by *GLK* expression. Expression of a full-length *SlGLK2* cDNA in either genotype resulted in homogenously dark green unripe fruit (Figure 2.6). Co-suppression of *SlGLK2* in four *U/U* transgenic lines converted the dark green shoulder *U* trait to light green, confirming that *SlGLK2* is *U* (Figure 2.6). The

Sl2.40ch10	*Slglk2/u (light fruit)*											*SlGLK2/U (dark fruit)*						
position	HEI	M82	TC	MM	CER	CSM	E6	FB	uLR	N93	CRG	AC	73X	T91	CAU	Spi	Sha	Spe
2292143	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	T	A
2292146	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	G	G
2292151	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T
2292203	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	T
2292218	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	T
2292260	A	A	A	A	A	A	A	A	A	A	A	-	-	-	-	-	-	-
2293093	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	T	T
2293727	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A	A
2293793	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	N	A
2293816	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	N	C
2293824	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	G	A
2293851	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	T	T
2293870	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	C
2293929	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	T	T
2293951	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	C	A
2294213	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	N	C
2294233	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	C	C
2294250	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	C	C
2294475	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	A
2294524	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A	A	A
2294541	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	T
2295824	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	G	G

Figure 2.3. *SlGLK2/U* and *Slglk2/u* coding sequences. Nucleic acid calls between SL2.40ch10:2292143-2295824 (SGN tomato genome release 2.4) from cDNA and genomic sequencing of the *u/u S. lycopersicum* varieties '*Heinz* 1706' (HEI), '*'M82''*, '*T63'*, '*Moneymaker'* (MM), '*Castlemart'* (CSM), '*E6203'* (E6), '*Fireball'* (FB), 'Long Red' (uLR), 'N93', *S. cerasiforme* PI114490 (Cer), and '*Craigella'* (CRG), and the *U/U S. pimpinellifolium* (Spi), *S. pennellii* (Spe) and *S. lycopersicoides (Sha)* wild tomato relatives and *S. lycopersicum* varieties '*'Ailsa Craig'*' (AC), '*73X'*, '*T91'*, and a '*Cuatomate'* (CAU) landrace. Translated RT-PCR products and BLAST searches predicted the start codon at 2292050; no differences were detected until position 2292143. The sequence difference at 2292260-2292267 of all *u/u* varieties compared to all *U/U* varieties is boxed in yellow.

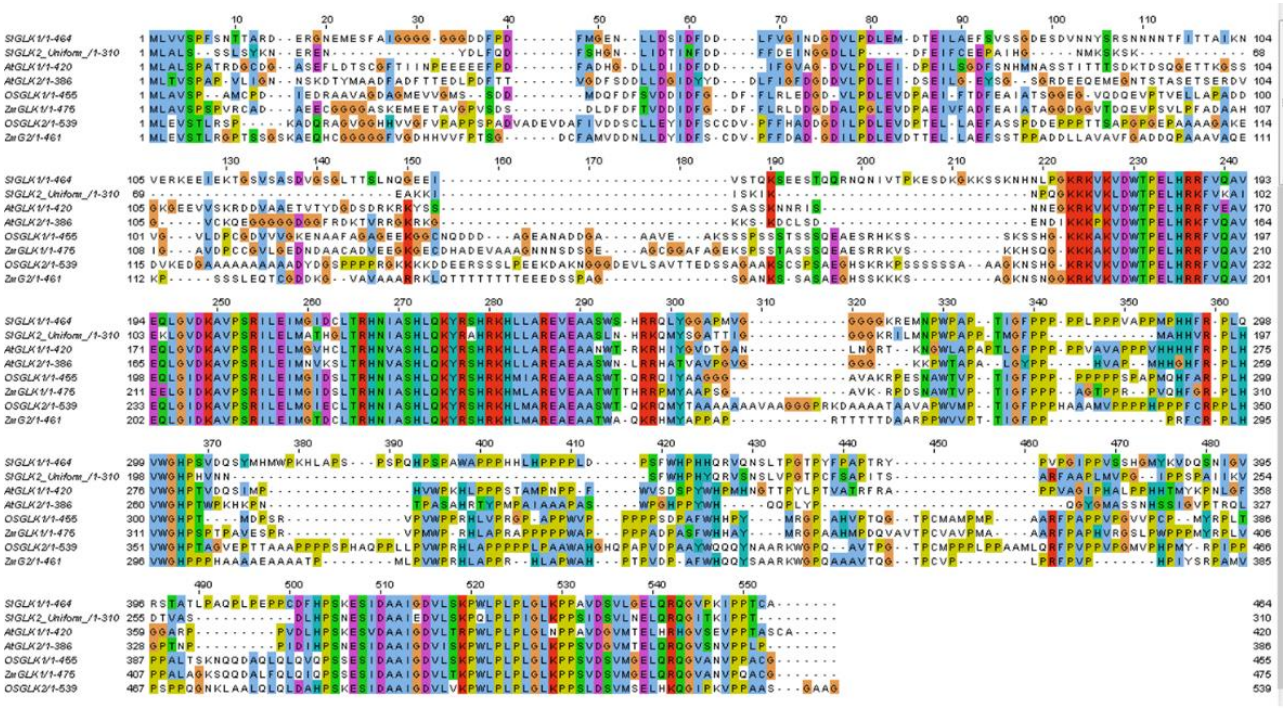


Figure 2.4. ClustalW alignment of *GLK* sequences from different species showing two conserved domains: DNA binding domain and C-terminal domain. Sl: tomato, At: *Arabidopsis*; Os: rice, Zm:maize.

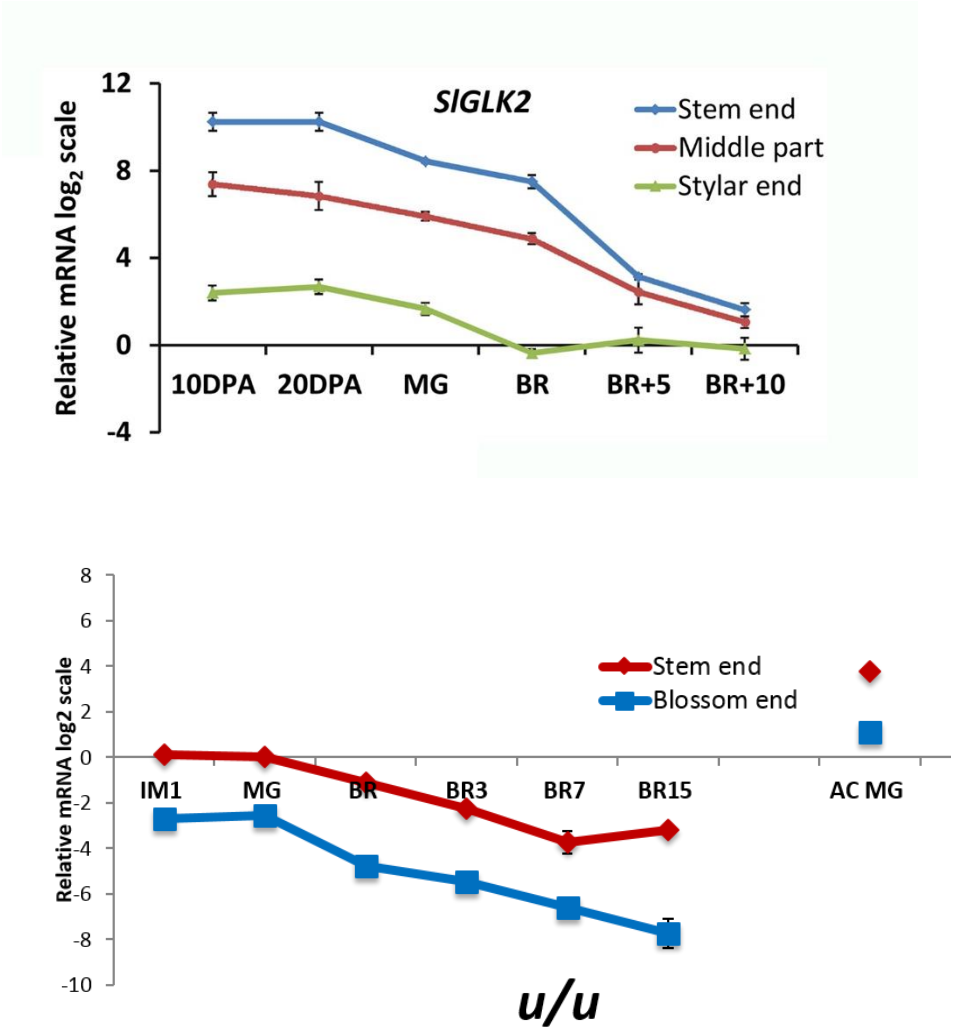


Figure 2.5. Gradient expression levels of *GLK2* as determined by qRT-PCR

Top: *SlGLK2* expression levels of wild type (*U/U*) *'Ailsa Craig'* (AC) fruit during six stages of tomato fruit development. DPA=days post anthesis

Bottom: *SlGLK2* expression level of *uniform* (*u*) fruit during development, which is lower than corresponding *U/U* AC Mature green fruit

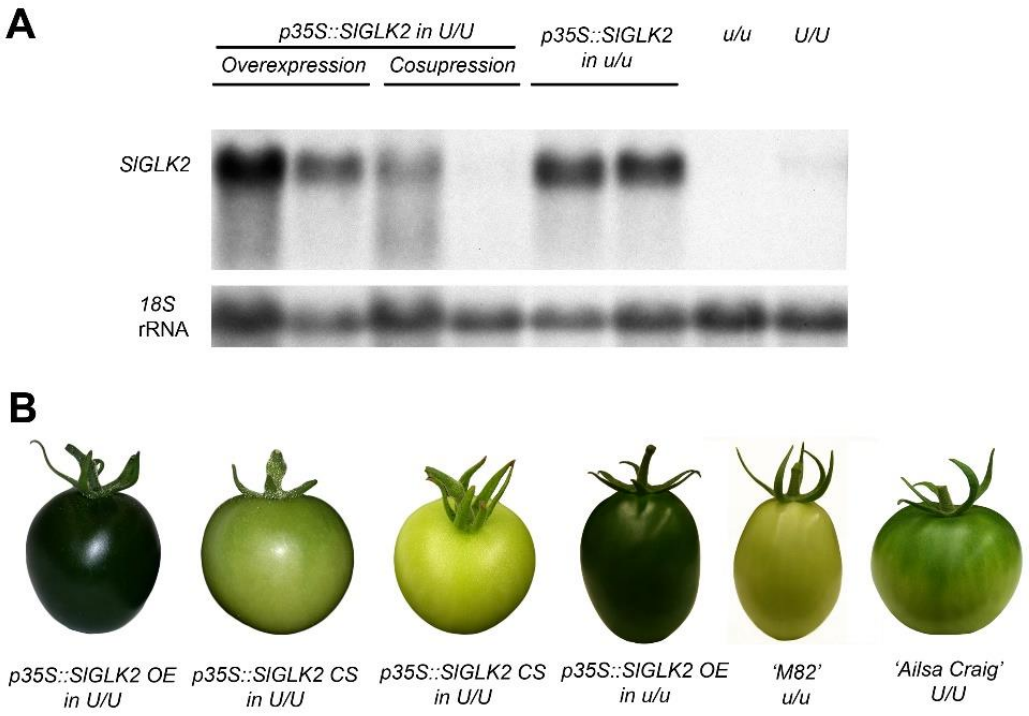


Figure 2.6. Expression and phenotypes of fruit expressing *p35S::SlGLK2*. **A.** *SlGLK2* expression as determined by hybridization of *SlGLK2* specific probes to gel blots of RNA from '*'Ailsa Craig'' U/U* or '*'M82'' u/u* transformed with *p35S::SlGLK2*. **B.** Fruit phenotypes from representative plants of lines overexpressing (OE) or with co-suppressed expression (CS) of *SlGLK2*. Five transformed *U/U* lines and more than 5 transformed *u/u* lines showed the overexpression fruit phenotype. Four other transformed *U/U* lines showed the co-suppression phenotype.

dark green fruit phenotype is confined to the shoulder region where *SlGLK2* is more highly expressed in *U/U* varieties (Figure 2.5) and all lines expressing *GLK*s with promoters expressed throughout the fruit produced homogenously dark green fruit, affirming that the manifestation and intensity of the phenotype depends on the spatial pattern and level of *GLK* expression.

SlGLK1 **and** ***SlGLK2*** **are functionally similar but differentially expressed, resulting in tissue-specific outcomes**

We showed that tomato *uniform ripening* (*u*) mutant fruit uniform light green phenotype is due to a nonsense mutation in tomato *GLK2*. As is the case in other plant species (*Arabidopsis*, maize, rice) (Rossini et al., 2001; Fitter et al., 2002; Nakamura et al., 2009), there are two *GLK* genes in tomato, *SlGLK1* and *SlGLK2*. In *Arabidopsis*, the two genes have redundant functions in regulating a set of light harvesting complex and chlorophyll biosynthetic genes (Waters et al., 2009). In tomato, elevated chlorophyll in the shoulder is positively regulated by *GLK2*, which is more highly expressed at this end of the fruit. To test the function and specificity of tomato *SlGLK1* and *SlGLK2* activities, we generated constructs to over-express each gene independently in tomato plants (in both the *U/U* and *u/u* genotypes) and driven by the 35S promoter. RNA gel-blot analysis (Figure 2.7) indicates up-regulation of each gene in the respective transgenic lines. Over-expression of either *SlGLK1* or *SlGLK2* using the 35S promoter results in darker green unripe fruit along the entirety of the latitudinal axis (Figure 2.8A). Dark green and uniform coloration was consistent with significantly elevated chlorophyll levels (Figure 2.9A) and increased expression of the *GLKs* throughout the fruit as a result of 35S promoter activity (Figure 2.9E). In wild-type tomato fruit, only *SlGLK2* is expressed at high levels and more so in the top than the bottom of the fruit resulting in the green shoulder phenotype. *SlGLK1* is expressed in the fruit but at very low levels compared to *SlGLK2* (Figure 2.7).

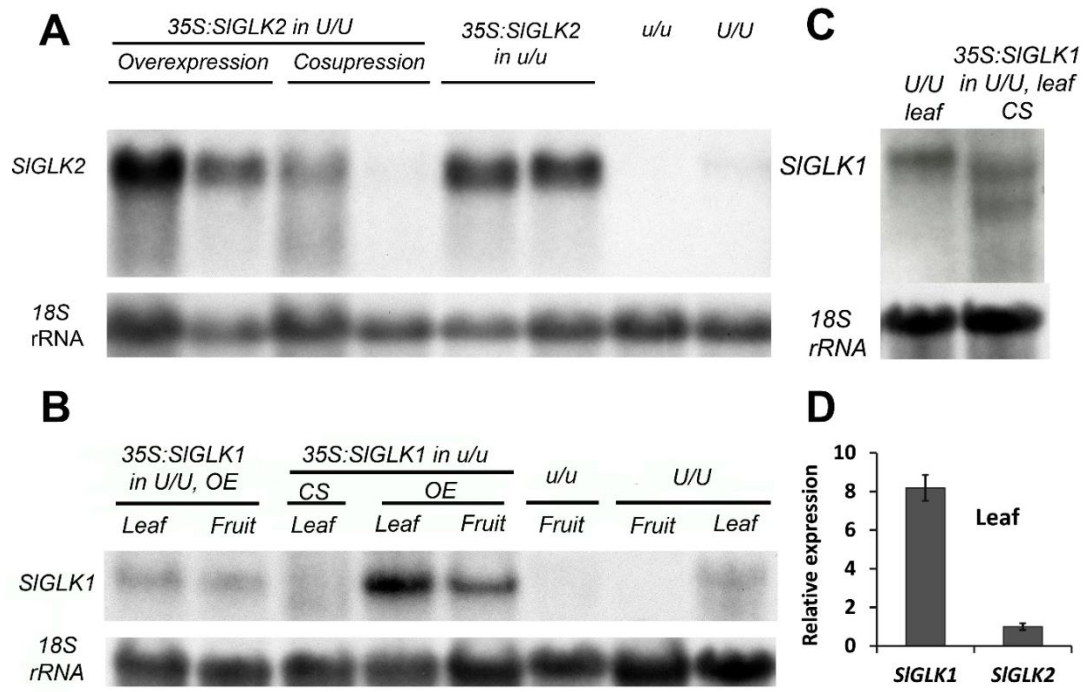


Figure 2.7. Expression of *SlGLK2* and *SlGLK1* in normal and transgenic tomato tissues.

(A) Expression of *SlGLK2* as determined by hybridization of radio-labeled *SlGLK2* probe to gel blot of RNA from immature green fruit tissues showing enhanced *SlGLK2* transcript levels in dark green fruit (over-expression) and reduced or degraded *SlGLK2* transcripts in "mimicked" *u* fruit (co-suppression) from *'Ailsa Craig'* (*U/U*) or *''M82''* (*u/u*) transformed with *35S:SlGLK2*. Each lane has RNA derived from tissue from an independent transgenic line.

(B) Expression of *SlGLK1* as determined by hybridization of radio-labeled *SlGLK1* probe to gel blot of RNA from immature green fruit and leaf tissues showing enhanced transcript levels in dark green fruited plants (over-expression, OE) or reduced and degraded transcripts in pale-leaved plants (co-suppression, CS) from *'Ailsa Craig'* (*U/U*) or *''M82''* (*u/u*) transformed with *35S:SlGLK1*.

(C) Expression of *SlGLK1* as determined by hybridization of radio-labeled *SlGLK1* probe to gel blot of RNA from leaf tissues showing reduced or degraded transcripts in pale leaf plants (co-suppression, CS) from *'Ailsa Craig'* (*U/U*) transformed with *35S:SlGLK1*.

(D) Relative expression levels of *SlGLK1* and *SlGLK2* in wild type *'Ailsa Craig'* (*U/U*) leaves as determined by qRT-PCR. Expression levels were normalized for internal control (18S) and amplification efficiency. Error bars represent standard error and data from a minimum of three biological replicates were used to determine each sample and standard error value presented.

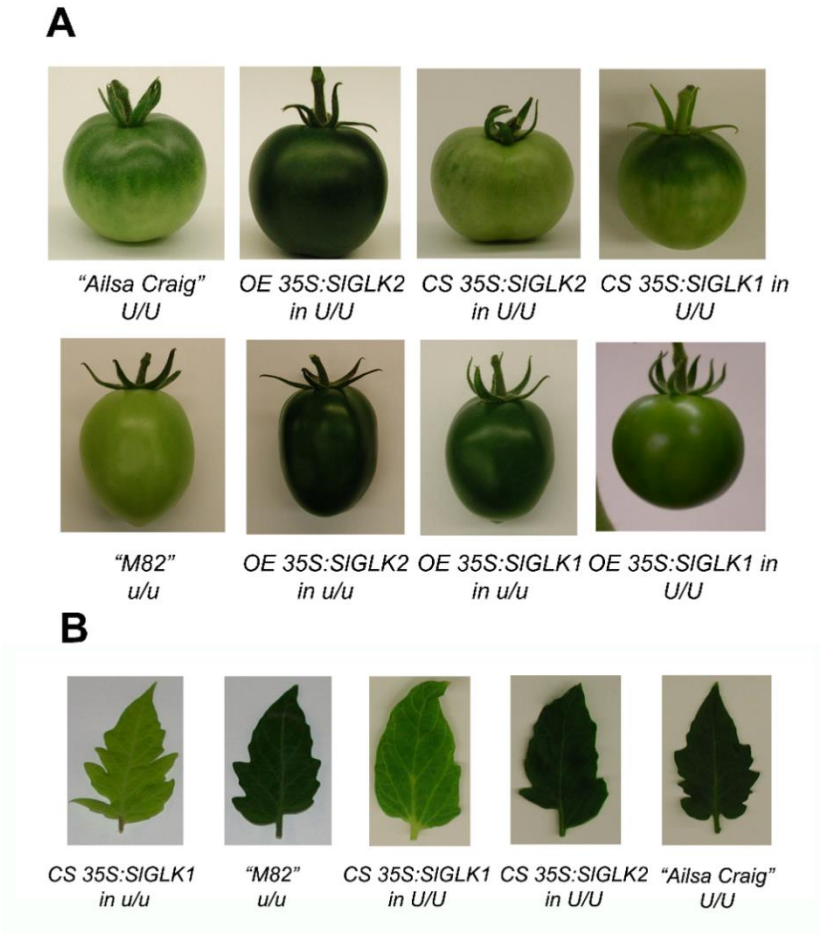


Figure 2.8. Fruit and leaf phenotypes of transgenic plants over-expressing or co-suppressing *SlGLK1* or *SlGLK2*.

(A) Immature green fruit (20 DPA) from *'Ailsa Craig'* (*U/U*) and *''M82''* (*u/u*) containing *35S:SlGLK1* or *35S:SlGLK2* with over-expression (OE) or co-suppression (CS) of *SlGLK1* or *GLK2*.

(B) Leaves from *'Ailsa Craig'* (*U/U*) and *''M82''* (*u/u*) co-suppressing *SlGLK1* and *SlGLK2*. From left to right: *''M82''* (*u/u*) co-suppression *SlGLK1* leaf, *''M82''* (*u/u*) leaf, *'Ailsa Craig'* (*U/U*) co-suppression *SlGLK1* leaf, *'Ailsa Craig'* (*U/U*) co-suppression *SlGLK2* leaf, *'Ailsa Craig'* (*U/U*) leaf.

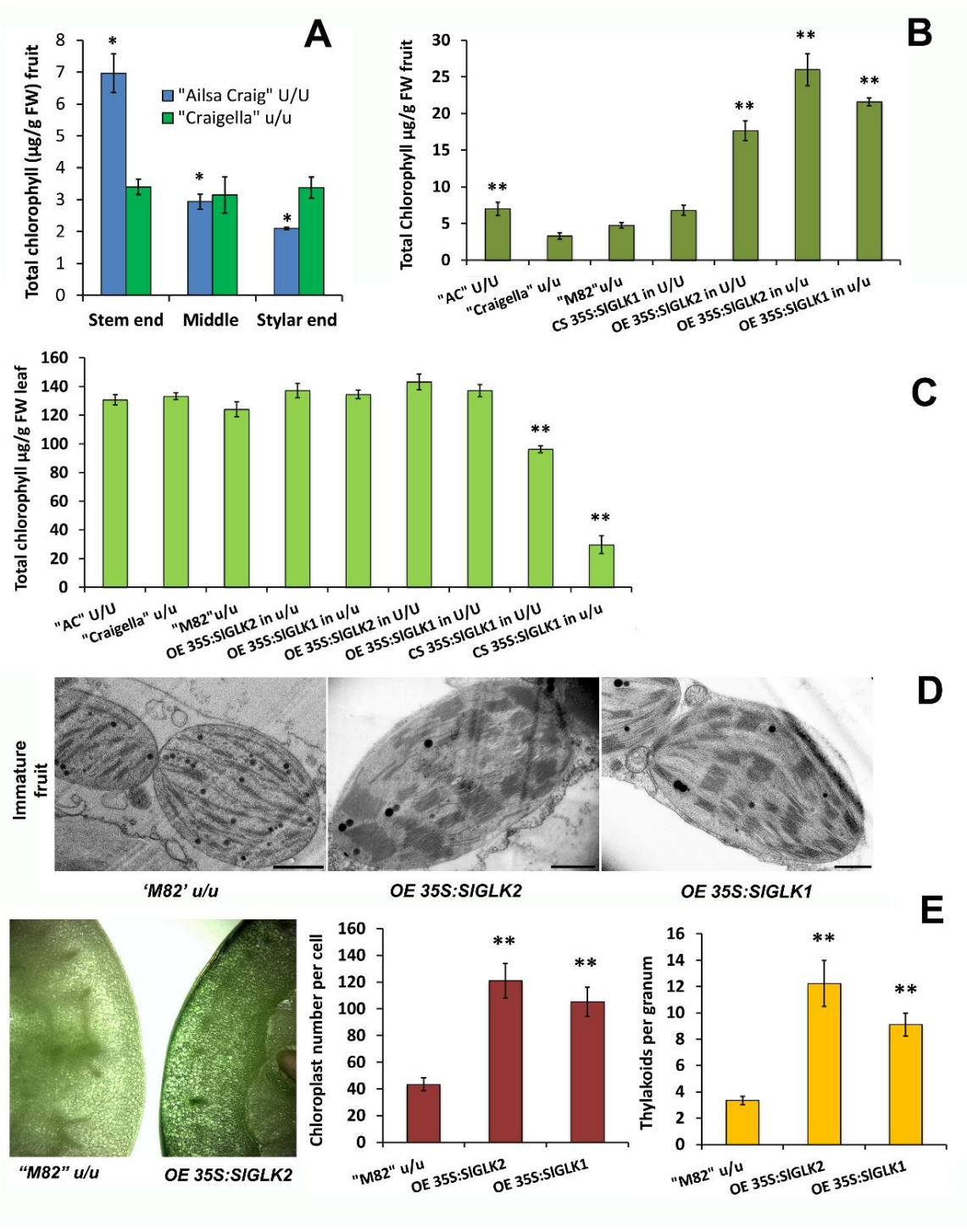

A
Total chlorophyll (µg/g FW) fruit
"Ailsa Craig" U/U
"Craigella" u/u
Stem end
Middle
Stylar end
B
Total Chlorophyll µg/g FW fruit
"AC" U/U
"Craigella" u/u
"M82"u/u
CS 35S:SlGLK1 in U/U
OE 35S:SlGLK2 in U/U
OE 35S:SlGLK2 in u/u
OE 35S:SlGLK1 in u/u
C
Total chlorophyll µg/g FW leaf
"AC" U/U
"Craigella" u/u
"M82"u/u
OE 35S:SlGLK2 in u/u
OE 35S:SlGLK1 in u/u
OE 35S:SlGLK2 in U/U
OE 35S:SlGLK1 in U/U
CS 35S:SlGLK1 in U/U
CS 35S:SlGLK1 in u/u
D
Immature fruit
'M82' u/u
OE 35S:SlGLK2
OE 35S:SlGLK1
E
"M82" u/u
OE 35S:SlGLK2
Chloroplast number per cell
"M82" u/u
OE 35S:SlGLK2
OE 35S:SlGLK1
Thylakoids per granum
"M82" u/u
OE 35S:SlGLK2
OE 35S:SlGLK1

Figure 2.9. Chlorophyll and chloroplast phenotypes of immature green fruit and leaves expressing *35S:SlGLK1* or *35S:SlGLK2*.

(A) Total chlorophyll levels from three sectioned segments along the latitudinal axis of immature green fruit (15DPA) (n=5) of *'Ailsa Craig'* (*U/U*) and *'Craigella'* (*u/u*). *'Craigella'* is a near isogenic line of *'Ailsa Craig'*, except for the *u* locus. *: p value t-test < 0.05. Error bars are standard error (S.E.)

(B) and (C) Total chlorophyll levels from leaves (B) and immature green fruit (C) (n=8) of different genotypes and transgenic plants expressing *35S:SlGLK1* or *35S:SlGLK2*. *'Ailsa Craig'* , AC; OE, over-expression; CS, cosupression. **: p value t-test < 0.001. Error bars are S.E.

(D) Transmission electron microscopy images of immature green fruit chloroplasts from *''M82''* (*u/u*) over-expressing *SlGLK1* or *SlGLK2*. Scale bar = 1 μm.

(E) Left: Horizontal cross-sectioned optical microscopy images of immature green fruit from *''M82''* (*u/u*) and *''M82''* over-expressing *SlGLK2*. Middle: Number of thylakoids per granum (sample size n=20). Right: Number of chloroplasts per cell (sample size n >15) from immature fruit tissues of *''M82''* (*u/u*) over-expressing *SlGLK1* or *SlGLK2*. **: p value t-test < 0.001. Error bars are S.E.

In leaves, both genes are expressed, though *SlGLK1* is expressed at higher levels (Figure 2.7). In the case of over-expression constructs transformed into wild-type *'Ailsa Craig'* (*U/U),* possessing functional alleles for both *SlGLK1* and *SlGLK2*, we observed co-suppression (Napoli et al., 1990) in several transgenic lines (Figure 2.7). Co-suppression of *SlGLK2* resulted in *u*-like fruit in that they are absent the green shoulder, but no discernible leaf phenotypes, whereas co-suppression of *SlGLK1* resulted in pale leaves with reduced chlorophyll, but no notable differences in the fruit (Figure 2.8). Co-suppression of *SlGLK1* in cultivar *''M82''*, a processing variety with the *u/SlGLK2* mutation, also yielded a pale leaf phenotype and low chlorophyll levels as in *'Ailsa Craig'* (*U/U*) wild type. Together these data suggest that both *SlGLK1* and *SlGLK2* are functionally similar peptides, but their expression patterns effectively restrict *Sl*GLK1 largely to leaf and *Sl*GLK2 largely to fruit functions, respectively, *in vivo*.

Over-expression of tomato *GLKs* enhances chloroplast development and nutritional quality

We initially focused on characterizing the dark green fruit of *SlGLK1* and *SlGLK2* over-expression lines. Consistent with the function of *GLK* genes in other plant species, over-expression of *GLKs* increased chlorophyll content, chloroplast number and thylakoid grana stacks in green fruit. Specifically, chlorophyll content increased 5-9 fold, chloroplast numbers 2.5-3 fold and the observed numbers of grana stacks 3 - 4 fold (Figure 2.9). It is interesting to note that the changes in chloroplast number and grana stacks have roughly multiplicative effects on chlorophyll levels, which is expected as most chlorophyll resides in the grana thylakoid membrane. The starch levels in green fruit of over-expressing transgenic plants were also increased (Figure 2.10C) as was observed with over-expression of *Arabidopsis GLK2* in tomato fruit (Powell et al., 2012). Because the chlorophyll content in leaves of *SlGLK1* and *SlGLK2* over-expressing lines was not substantially altered when compared with control leaves (Figure

2.9B), the increase in transgenic fruit starch levels likely directly reflects increased fruit photosynthesis. In red tomato fruit, many desirable nutritional compounds are elevated, including not only carbohydrates/sugars, but also carotenoids and ascorbate (vitamin C) (Figure 2.10). Total soluble solids (^{o}Brix - percent w/w of soluble solids in solution) and sugar levels were significantly increased (20 - 40% and 40 – 60%, respectively) and total carotenoid levels were elevated 25-40% in both *SlGLK1* and *SlGLK2* over-expressing lines as compared to the control *''M82''* (*u/u*). Interestingly, in the red ripe pericarp tissue, β-carotene and lutein, the protective carotenoids usually found at high levels associated with the light harvesting complexes of leaves, were also substantially increased (Figure 2.10A). We also note that the elevation of the predominant carotenoid, lycopene, is consistent with elevated plastid numbers and thus storage capacity, while the larger changes in β-carotene and lutein are more difficult to interpret as they reflect a much smaller proportion of the total carotenoid pool. One possibility is that this increase includes residual thylakoid-associated carotenoids from dark green unripe stages of GLK overexpressing fruit.

GLK2 acts independently of *hp1/DDB1*

Prior studies of tomato mutations reveal additional genes impacting fruit pigmentation and we explored whether or not said genes might act in the same pathway as GLKs. For example, high pigment (*hp*) mutations including *hp1, hp2, hp3* display elevated green fruit chlorophyll and ripe fruit carotenoids as a result of mutations in light signal transduction genes *DDB1(hp1)*, *DET1(hp2)* (Mustilli et al., 1999; Liu et al., 2004), or an abscisic acid biosynthetic gene *zeaxanthin epoxidase* (Zep) in the case of *hp3* (Galpaz et al., 2008). To test for epistasis when combining different genes or mutations that have positive effects on tomato chloroplasts and associated nutritional value, I focused initially on the relationship of the well characterized

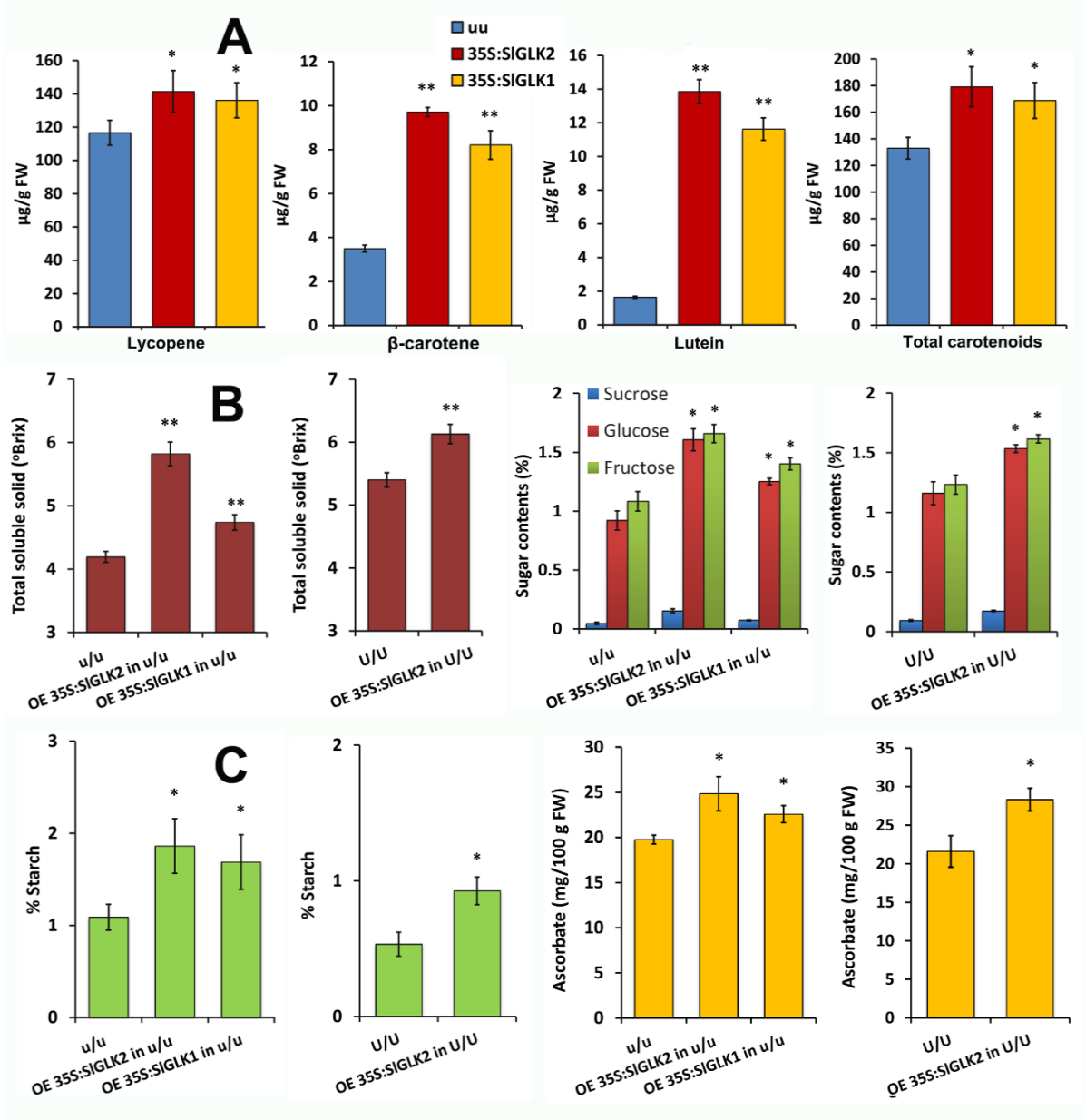


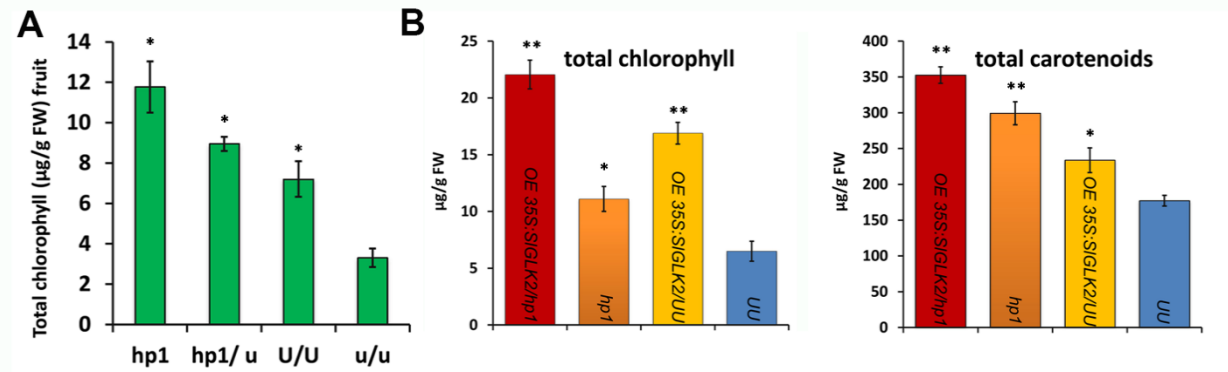
Figure 2.10. Fruit phenotypes resulting from tomato GLK expression. Carotenoids (A), °Brix, sugars (B), starch and ascorbate (C) levels in *''M82''* (*u/u*) or *''Ailsa Craig''* (*U/U*) fruit over-expressing (OE) *SlGLK1* or *SlGLK2*. °Brix (n=30), sugar, carotenoid and ascorbate levels were measured in red fruit (10 days after the breaker stage) (n=8). Starch was measured in immature green fruit (15 DPA). Five biological replicates (n=5) were used for each analysis. *: p value t-test < 0.05; **: p value t-test < 0.001. Error bars are S.E.

HP1/DDB1 gene (Liu et al., 2004) and *U*/*SlGLK2*. *hp1/hp1 u/u* double mutant fruit display darker green color and higher chlorophyll content than *u/u* fruit (Figure 2.11A), suggesting independent and additive activities. The additive actions of *DDB1* and *SlGLK2* were supported when I characterized F2 plants of a cross between *hp1/hp1* and wild type (*U/U)* over-expressing *35S:SlGLK2*, where both lines are in the same *'Ailsa Craig'* genetic background. Homozygous *hp1/hp1* plants with dark green color throughout the fruit due to *35S:SlGLK2* activity had more chlorophyll in green fruit and carotenoid in red fruit than the other genotypic classes in the F2 population (Figure 2.11B). These observations indicate the additive effects of both genes on fruit quality traits and suggest the feasibility of combining elevated expression of *GLKs* with other genes (including *DDB1/HP1* reduction) in improving fruit characteristics.

An additional transcription factor gene that has been characterized in tomato and resulted in dark green immature fruit color and elevated chlorophyll levels when repressed is *TAGL1* (Itkin et al., 2009; Vrebalov et al., 2009; Gimenez et al., 2010). We measured the expression level of *U/GLK2* in immature green fruit of *TAGL1* suppressed (via RNAi) lines and saw no changes in the expression level of *GLK2* as compared to the wild type (Figure 2.12), suggesting that the dark green color of fruit in these lines is not a direct result of changes in *SlGLK2* expression level.

GLKs alters expression of fruit photosynthesis and chloroplast development genes

To test the molecular consequences of over-expressing *SlGLK1* or *SlGLK2* in transgenic fruit, we compared the gene expression levels in the pericarp of immature green transgenic fruit with that in untransformed isogenic control pericarp using strand-specific mRNA sequencing (Zhong et al., 2011). Using cut-off criteria with an expression ratio of ≥ 2 and $p < 0.05$ between transgenic and control tissues, over-expressing *35S:SlGLK1* or *35S:GLK2* in *u/u* up-regulated



	Phytoene	Lutein	β-carotene	γ-carotene	cis-Lycopene	trans-Lycopene
OE 35S:SlGLK2/hp1	9.19 ± 1.32	7.76 ± 0.23	21.68 ± 1.24	4.09 ± 0.14	3.29 ± 0.27	301.10 ± 10.1
hp1	9.42 ± 1.19	3.27 ± 0.17	12.73 ± 1.69	3.64 ± 0.28	2.91 ± 0.34	263.45 ± 12.76
OE 35S:SlGLK2/UU	4.69 ± 0.19	6.32 ± 0.65	16.58 ± 1.24	2.80 ± 0.15	2.59 ± 0.23	198.34 ± 14.20
UU	3.76 ± 0.13	1.98 ± 0.25	12.74 ± 0.99	2.37 ± 0.11	1.42 ± 0.13	152.78 ± 7.79

Figure 2.11. Fruit chlorophyll and carotenoid levels from *hp1/hp1 u/u* double mutant and *hp1/hp1* mutant over-expressing *GLK2*. **(A)** Total chlorophyll levels from immature green fruit (15DPA) (n=5) of the *hp1* mutant, *hp1/hp1 u/u* double mutant, *U/U* wild type *('Ailsa Craig')*, and *u/u* mutant *('Craigella')*. **(B)** Chlorophyll (from immature green fruit 15 DPA) and carotenoid levels (from red ripe fruit Breaker +10) of fruit from the F2 generation of *hp1/hp1* x *'Ailsa Craig' (U/U)* over-expressing (OE) *GLK2*. Eight biological replicates (n = 8) were used for each analysis and all values in the figure are expressed as ug/g fresh weight (FW). *: p value t-test < 0.05; **: p value t-test < 0.001.

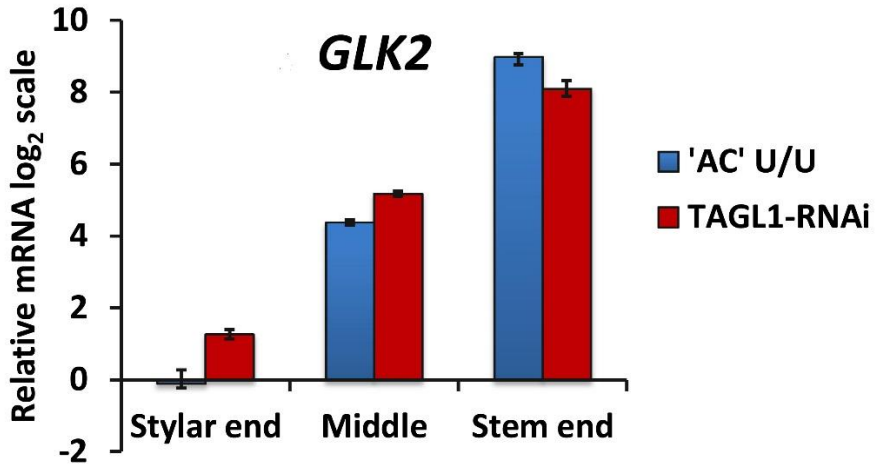
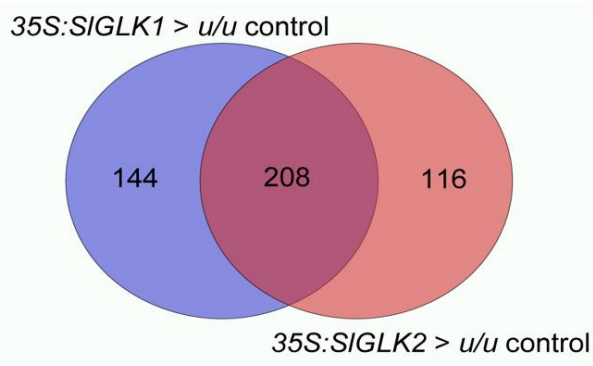


Figure 2.12. *GLK2* expression level in wild type and *TAGL1* RNAi repressed green fruit tissues. Gradient expression levels of *SlGLK2* as determined by qRT-PCR through 3 latitudinal sections of wild type *'Ailsa Craig' (AC)* and *TAGL1-RNAi* immature green (15 DPA) fruit tissues.

324 and 352 genes, respectively. Consistent with microarray studies using *Arabidopsis GLK* genes (Waters et al., 2009), a large portion (>60%) of the up-regulated genes are common to both transgenes (Figure 2.13). It is important to note that we observe higher expression of *35S:SlGLK2* than *35S:GLK1* in our transgenic lines, based on higher RPKM values and supported by higher chlorophyll levels in *35S:SlGLK2* over-expression lines. The additional expression may partially explain the differences in up-regulated genes between *SlGLK1* and *SlGLK2* over-expression in immature green pericarp tissues. Moreover, when using more stringent criteria for defining up-regulated genes (ratio ≥ 5, p value ≤ 0.05), the percentage of shared up-regulated genes increases to ~70%. Assessing GO term enrichment in *35S:SlGLK1* and *35S:SlGLK2* over-expression fruit compared to the control (*'M82" u/u*), yielded similar outcomes for both transgenes (Figure 2.13; Table 2 and 3). A large portion of the common enriched terms relate to photosynthesis and chloroplast structure, and include genes involved in the chlorophyll biosynthesis pathway, thylakoid membrane biogenesis, light harvesting complexes and starch metabolism (Figure 2.13). These changes in gene expression agree well with the phenotype of the fruit, where the chlorophyll content, chloroplast numbers, thylakoid grana stacks and starch levels are significantly elevated. Deploying the same methodology in ripe fruit tissues, however, revealed just a small number of genes that are up-regulated despite the fact that both *SlGLK1* and *SlGLK2* are expressed at high levels (Nguyen et al., 2014). Together these results suggest that tomato GLKs exert their primary influence on green fruit tissues even in cases of phenotypes that manifest in later stages of fruit development. Examples include the higher levels of ripe fruit carotenoids and sugars resulting from increased green fruit plastids and photosynthetic activity, respectively.

A



B

GO term enrichment of overlapping *SlGLK1* and *SlGLK2* up-regulated genes

Gene Ontology term	GO Term	Adjusted p value
GO: Cellular Component		
Thylakoid	GO:0009579	0
Photosystem	GO:0009521	0
Photosystem II	GO:0009523	0
Photosystem I	GO:0009522	0
Thylakoid membrane	GO:0042651	0
Plastoglobule	GO:0010287	0
Chloroplast	GO:0009507	0
Chloroplast part	GO:0044434	0
Organelle subcompartment	GO:0031984	0
GO: Biological Process		
Photosynthesis	GO:0015979	0
Photosynthesis, light reaction	GO:0019684	0
Photosynthesis, light harvesting	GO:0009765	0
Plastid organization	GO:0009657	0
Photosystem II assembly	GO:0010207	0
Chlorophyll biosynthetic process	GO:0015995	0
Tetrapyrrole biosynthetic process	GO:0033014	0
Starch biosynthetic process	GO:0019252	0
Pigment biosynthetic process	GO:0046148	0
Carbohydrate catabolic process	GO:0016051	0
Photosynthetic electron transport chain	GO:0009767	0.0013
GO: Molecular Function		
Chlorophyll binding	GO:0016168	0
Fructose-bisphosphate aldolase activity	GO:0004332	0
Tetrapyrrole binding	GO:0046906	0
Pectinesterase activity	GO:0030599	0
Protochlorophyllide reductase activity	GO:0016630	0.0215
Oxidoreductase activity	GO:0016491	0.0282

Figure 2.13. Over-expression of *SlGLK1* or *SlGLK2* influences similar photosynthesis-related genes in green fruit.

(A) Venn diagram of up-regulated genes unique to and shared between *35S:SlGLK1* and *35S:SlGLK2* immature green fruit (15DPA) tissues.

(B) GO term enrichment analysis of the overlapping up-regulated genes. Terms are ranked by adjusted p values.

Table 2. GO term enrichment of *SlGLK1*- up-regulated genes

Gene Ontology term	GO Term	Raw p-value	Adjusted p value
GO: Cellular Component			
Thylakoid	GO:0009579	1.27E-41	0
Organelle subcompartment	GO:0031984	1.88E-38	0
Thylakoid membrane	GO:0042651	1.39E-34	0
Chloroplast part	GO:0044434	2.59E-34	0
Chloroplast	GO:0009507	6.44E-28	0
Photosystem	GO:0009521	5.26E-24	0
Photosystem II	GO:0009523	1.21E-21	0
Photosystem I	GO:0009522	9.97E-20	0
Thylakoid light-harvesting complex	GO:0009503	9.58E-06	0
GO: Biological Process			
Photosynthesis	GO:0015979	9.79E-49	0
Photosynthesis, light reaction	GO:0019684	1.93E-44	0
Chlorophyll biosynthetic process	GO:0015995	2.06E-15	0
Photosynthesis, light harvesting	GO:0009765	7.39E-12	0
Starch biosynthetic process	GO:0019252	5.00E-11	0
Carbohydrate catabolic process	GO:0016051	2.77E-11	0
Pigment biosynthetic process	GO:0046148	5.97E-09	0
Tetrapyrrole biosynthetic process	GO:0033014	1.29E-09	0
Photosynthetic electron transport chain	GO:0009767	2.15E-09	0
Photosynthesis, dark reaction	GO:0019685	1.20E-09	0
Carotenoid metabolic process	GO:0016116	9.41E-08	0
GO: Molecular Function			
Chlorophyll binding	GO:0016168	7.83E-13	0
Protochlorophyllide reductase activity	GO:0016630	4.67E-06	0
Tetrapyrrole binding	GO:0046906	1.6E-05	0.009
Metal cluster binding	GO:0051540	1.6E-04	0.013
Pectinesterase activity	GO:0030599	3.4E-04	0.02
Oxidoreductase activity	GO:0016491	5.0E-04	0.029

Table 3. GO term enrichment of *SlGLK2*- up-regulated genes

Gene Ontology term	GO Term	Raw p-value	Adjusted p value
GO: Cellular Component			
Photosystem	GO:0009521	2.00E-16	0
Photosystem II	GO:0009523	1.20E-15	0
Thylakoid	GO:0009579	3.23E-13	0
Light-harvesting complex	GO:0030076	2.03E-13	0
Photosystem I	GO:0009522	1.81E-12	0
Organelle subcompartment	GO:0031984	9.58E-11	0
Thylakoid membrane	GO:0042651	2.30E-10	0
Chloroplast part	GO:0044434	5.72E-07	0
Chloroplast	GO:0009507	1.14E-06	0
GO: Biological Process			
Photosynthesis	GO:0015979	5.82E-17	0
Photosynthesis, light reaction	GO:0019684	7.20E-16	0
Photosynthesis, light harvesting	GO:0009765	7.39E-12	0
Chlorophyll biosynthetic process	GO:0015995	7.73E-08	0
Tetrapyrrole biosynthetic process	GO:0033014	3.46E-06	0
Starch biosynthetic process	GO:0019252	2.48E-05	0.0016
Pigment biosynthetic process	GO:0046148	6.92E-05	0.0026
Carbohydrate catabolic process	GO:0016051	9.55E-05	0.0038
Reactive oxygen species metabolic process	GO:0072593	2.7E-03	0.0083
Carotenoid metabolic process	GO:0016116	2.9E-03	0.044
GO: Molecular Function			
Chlorophyll binding	GO:0016168	1.96E-08	0
Metal cluster binding	GO:0051540	2.17E-07	0
Copper ion binding	GO:0005507	3.15E-07	0
Electron carrier activity	GO:0009055	1.14E-05	0.001
Tetrapyrrole binding	GO:0046906	5.82E-05	0.01
Oxidoreductase activity	GO:0016491	3.7E-04	0.03

Over-expression of *SlGLK1* and *SlGLK2* does not affect overall ripening regulation

I questioned whether the early fruit development phenotypes of immature green fruit over-expressing *35S:SlGLKs*, i.e., dark green color and high chlorophyll content, would impact the overall ripening process. I initially examined four GLK over-expression lines through three generations and observed no differences from the time of fruit set to the initiation of ripening (Figure 2.14), although there were changes in carotenoid accumulation and coloration as the fruit matured (Figure 2.14; Figure 2.10). Ethylene evolution rates were similar between the transgenic fruit and controls (Figure 2.14). In the RNA sequencing data comparing gene expression levels in the red fruit of these transgenic lines with those of the control (see Supplemental Data on Nguyen et al 2014), there were no differences in the expression of genes known to coordinate ripening (*RIPENING-INHIBITOR, RIN; NON-RIPENING, NOR; COLOURLESS NON-RIPENING, CNR; TAGL1*) or ethylene biosynthesis and signaling. Together these results suggest that *SlGLK1* and *SlGLK2* do not affect general ripening control systems, including those involving ethylene and critical ripening transcription factors. Rather, they exert their influence on fruit development through phenotypes manifested from altered green fruit plastid activity and plastid numbers, including sugar and later carotenoid accumulation.

III. Discussion

Fruit but not leaf GLK activities are limiting

It has been suggested that *GLK* duplication predated and facilitated the evolution of C4 photosynthesis, as C4 species such as maize and sorghum demonstrate differential mesophyll (*GLK1*) versus bundle sheath (*G2*) *GLK* expression (Wang et al., 2013). Our results extend this

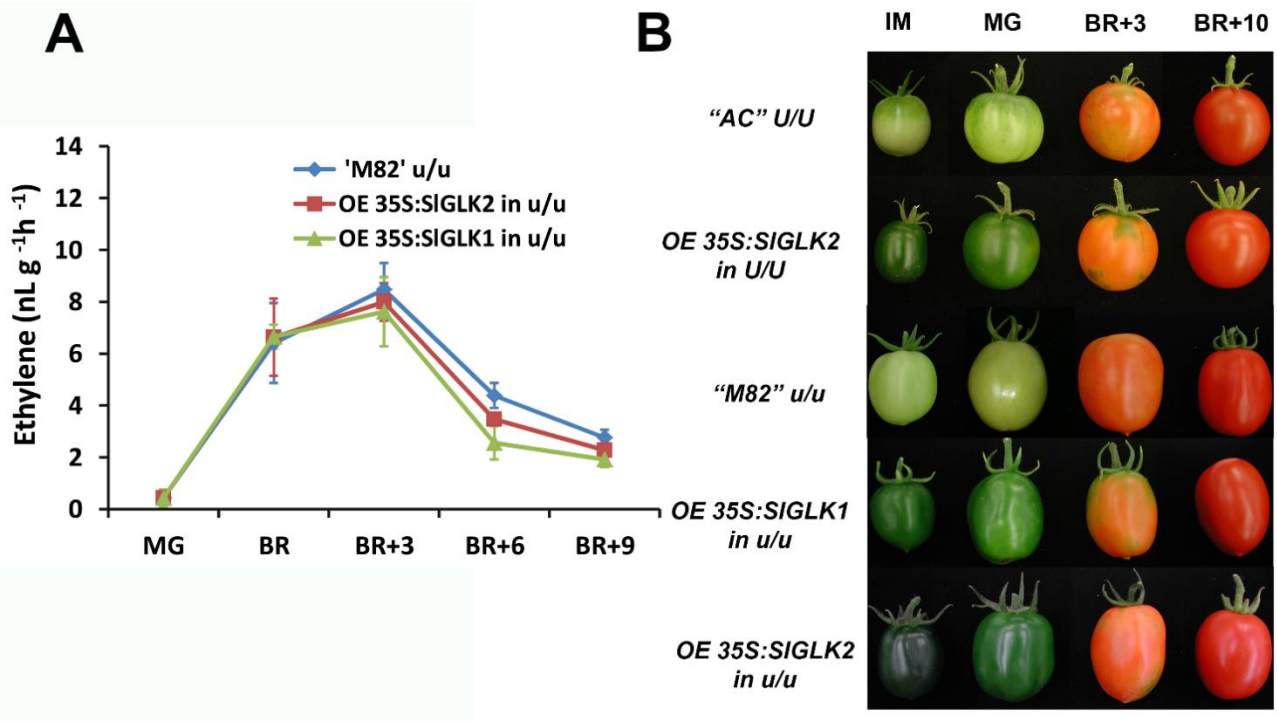


Figure 2.14. Ethylene evolution and external color of *''M82''* (*u/u*) fruit over-expressing *SlGLK1* and *SlGLK2* during fruit ripening.

(A) Ethylene levels as determined by gas chromatography in *''M82''* (*u/u*) control fruit and transgenic fruit over-expressing *SlGLK1* or *SlGLK2* during development (n=10).

(B) Fruit of *'Ailsa Craig'* (*U/U*), *''M82''* (*u/u*), and transgenic plants over-expressing *SlGLK1* or *SlGLK2*. IM: immature green (15 DPA), MG: mature green, BR: Breaker.

hypothesis further to suggest that the duplication facilitated additional regulatory fidelity of chloroplast development for C3 species bearing fleshy fruits.

Fruit are net sink tissues with photosynthate from source tissues (leaves) transported to and largely responsible for their carbohydrate accumulation. With this transgenic tomato GLK data it show that green tomato fruit but not leaves can be elevated in terms of photosynthetic capacity through GLK manipulation. Specifically, over-expression of either *SlGLK1* or *SlGLK2* resulted in dark green tomato fruit with high chlorophyll and chloroplast levels in addition to more stacked thylakoid grana and elevated starch in the fruit. However, no significant changes were observed in leaves. *SlGLK1* repression resulted in chlorotic leaves, demonstrating its importance in leaf plastid development while *SlGLK2* repression had no leaf phenotype. Elevated GLK activity also impacted mature fruit via increased sugar and carotenoid content, but did not influence the rate or timing of ripening. These results indicate that leaf GLK activity is saturated while the potential to increase fruit photosynthetic activity via GLK induction remains. Elevating GLK activity via transgenes or through natural diversity thus represents a strategy to manipulate fruit photosynthesis to enhance fruit crop quality without affecting non-fruit tissues.

Transgenic over-expression of either *SlGLK1* or *SlGLK2* leads to elevated quality and nutrient compound levels in red ripe tomato fruit, e.g., increases in sugars, multiple carotenoids and ascorbate (Vitamin C). High sugar levels in transgenic fruit over-expressing *GLKs* correlate with higher starch in green fruit that is metabolized to glucose and fructose in red fruit. Although carotenoid biosynthetic genes were not differentially expressed (Nguyen et al 2013) in transgenic red fruit tissues as compared to controls, the higher chloroplast numbers and carotenoids in green fruit presents a source of more carotenoids, as chloroplasts are converted to chromoplasts during fruit ripening, increasing the sink capacity for carotenoid accumulation (Lu and Li, 2008).

Elevated fruit ascorbate levels in transgenic GLK fruit likely reflect a photoprotective response to quench hydroperoxides, byproducts of photosynthesis, resulting from higher chlorophyll and chloroplasts in photosynthetic green fruit. From a practical perspective, fruit GLKs can be over-expressed without affecting non-fruit tissues such as leaves, meaning that natural or transgenic variation resulting in elevated GLK expression can be employed as a means of positively impacting fruit quality through increased sugars, carotenoids and organic acids such as ascorbate, which are positively correlated with favorable flavor attributes (Baldwin et al., 2008; Klee and Tieman, 2013).

GLK and DDB1 fruit quality effects are additive, suggesting a route to amplified fruit quality improvement

Several tomato high pigment (*hp*) mutants conferring elevated chlorophyll and carotenoid levels have been described (e.g., *hp1, hp2,* and *hp3*) and result from mutations in light signaling processes [*DET1* (*hp2*) *DDB1* (*hp1*)] or abscicic acid biosynthesis (*hp3*). The data presented here shows that *SlGLK2* acts independently of *HP1/DDB1,* suggesting that different routes regulate plastid/chlorophyll and carotenoid accumulation in tomato fruit. While GLKs are transcription factors that positively regulate chlorophyll biosynthesis and accumulation of light harvesting complexes in addition to chloroplast development in general, DDB1 and DET1 contribute to multiple developmental and response processes, including light signal transduction, chromatin modeling and plastid biogenesis (Benvenuto et al., 2002; Liu et al., 2004). Comparing the numbers of genes differentially regulated in response to *SlGLK2* over-expression (Nguyen et al 2013) as compared to perturbation of DET1 suggests much broader effects resulting from DET1 activity (Enfissi et al., 2010). Because *DET1* and *DDB1* are components of the same light signaling process and interact with each other (Schroeder et al., 2002), we anticipate a similar

result if we were to cross the *hp2* mutant with *SlGLK2* over-expressing lines. As such, breeding combinations deploying alternate high activity alleles of *GLK*s with either high activity *HP1* or *HP2* alleles could be pursued toward optimized fruit quality. *TAGL1* suppression results in both early (high chlorophyll content) and late (reduced carotenoid levels) ripening effects (Vrebalov et al., 2009), suggesting *GLKs* and *TAGL1* as another combination of genes that could influence fruit quality. However, in this case, *TAGL1* alleles with reduced activity during early fruit development but not during ripening would be required to positively impact fruit quality. Alternatively, targeted repression via an early fruit development promoter would be predicted to yield similar results.

IV. Materials and Methods

Plant materials

Tomato cultivars *'Ailsa Craig' (U/U) (LA2838A), 'Craigella' (u/u) (LA3247), ''M82'' (u/u) (LA3475), hp1 (LA0279),* and *hp1/u (LA3371)* were provided by the Tomato Genetics Resource Center, Davis, CA. Craigella harbors the *u* mutation and is nearly isogenic (5 back-crosses) with *'Ailsa Craig'*. Plants were grown in greenhouses at the Boyce Thompson Institute for Plant Research, Ithaca, NY with a 16 h light (27^{o}C) and 8 h dark (19^{o}C) cycle. Tomato fruit were tagged at 7-8 days post anthesis (dpa) when they were 1 cm in diameter and reached breaker stage at 37 to 38 dpa.

Expression constructs and transformation

The full-length sequences of *SlGLK1* and *SlGLK2* cDNA were cloned into pBI121 (Clontech, CA) and pBTEX (Lin et al., 2006) vectors digested with SmaI/SacI and SmaI/SalI, respectively, under the control of the CaMV35S promoter. The fidelity of the construct was confirmed by

DNA sequencing and transgenic *'Ailsa Craig' (U/U)* and *''M82'' (u/u)* tomato plants were generated by *Agrobacterium tumefaciens* (strain LBA4404) mediated transformation using previously described methods (Fillatti et al., 1987). Plants with transgene insertions were screened by PCR and DNA gel-blot using CAMV 35S-specific primers/probe. *TAGL1* RNAi lines were described previously (Vrebalov et al., 2009).

RNA gel-blot analysis

Total RNA was extracted from frozen tissues and gel-blot analysis were performed as described (Vrebalov et al., 2009) using 25 µg total RNA per lane. pGEM-T Easy vectors containing *SlGLK1* and *SlGLK2* full-length cDNAs were used as template to generate *SlGLK1*- and *SlGLK2*-specific probes by PCR using primers GLK1F/R for *SlGLK1* and GLK2F/R for *SlGLK2* (Table 4).

Quantitative RT-PCR analysis

Quantitative real-time PCR was performed using the ABI PRISM 7900HT Sequence Detection System and the Power SYBR Green RNA-to-C_T 1-Step Kit (Applied Biosystems) following the manufacturer's instructions. Quantitation of gene expression was calculated using the relative standard curve method (Applied Biosystems). Gene-specific primers used in the qRT-PCR assay are listed in Table 2.3. Standard errors (SEs) were calculated based on a minimum of three biological replicates.

Soluble solid, sugar and starch analysis

Soluble solids (^{o}Brix) from fruit juice were measured with a digital refractometer (HR200, APT Co, LLC). For sugar analysis, 200 mg frozen red ripe tissue samples were ground in a mortar

with a pestle in 2.5 mL methanol/chloroform/water (12/5/3; vol/vol/vol) and 2 mL chloroform. Then, 0.6X volume water was added to create an aqueous phase to separate from the chloroform phase. After centrifugation, the water layer was saved and purified by anion and cation exchange resins to remove charged components. The neutral fractions were vacuum dried and rehydrated with water and analyzed by HPLC using a Waters SugarPak column as described (Turgeon and Medville, 1998). Starch quantification was determined using a Starch Assay Kit (STA20, Sigma, St. Louis, MO) following the manufacturer's protocol.

Ethylene measurement

Tomato fruit were placed in open 250 ml airtight jars at room temperature for 2 h to reduce harvest stress. The jars were then sealed and after 2 h, 1 ml air samples were taken and analyzed by gas chromatography using a Hewlett-Packard 5890 series II with a flame ionization detector and an activated alumina column. Ethylene concentrations were calculated by comparison with a standard of known concentration and normalizing for fruit mass.

Carotenoid and chlorophyll analysis

Carotenoid and chlorophyll extraction and carotenoid quantification by High Pressure Liquid Chromatography (HPLC) were performed as previously described (Vrebalov et al., 2009). For chlorophyll, the dried down extracts were resuspended in 100% acetone and the amount of chlorophyll was determined spectrophotometrically using published formulas (Lichtenthaler, 1987).

Ascorbate Measurement

Ascorbate levels were determined spectrophotometrically based on the ΔA_{265} following ascorbate oxidation as previously described (Alba et al., 2005).

Table 2.3. DNA primer sequences.

	Forward Primer (5' to 3')	Reverse Primer (5' to 3')
qRT-GLK1	TCATCATCTTCATCCTCCACCTC	TGGTGCTGGAAAGTAAGGAGTTC
qRT-GLK2	GACACAGTTGCCTCTGATTTGC	TGAGGAGGTTTGAGTCCTATGG
GLK1 probe	TTCGCCGTCTCCACAACAT	AATGGCTGCATCTATGCTCTCC
GLK2 probe	ATGCTTGCTCTATCTTCATCATTGA	TTGAAGATGACTAGCAATGTTATGTCT
35S	AAACCTCCTCGGATTCCATTG	AGGAAGGGTCTTGCGAAGGA
qRT-08g083400	AGTACCCAGCTGCTGCTTGG	CTGAGACATGCTGCCAGGTG
qRT-01g100510	TCCGCTGCTTCCATCTCCT	AGCTCAGTTCCCTATGCAAAG
qRT-Sucr	AGGTGGTACCATCCCTTGTGA	TATGCGTGGTCCAATGGCTATG
qRT-18S	CGGAGAGGGAGCCTGAGAA	CCCGTGTTAGGATTGGGTAATTT
HP1 marker	ATCATAATGGGCTTCTTCACCTACT	GACCTTGCCTTTCCAGATCAAC

Chloroplast counting and transmission electron microscopy

For plastid counting, immature outer fruit pericarp tissues were fixed in 3.5% glutaraldehyde solution for 1 h in the dark and then heated in 0.1 M Na_2-EDTA at 60°C for 10–15 min as previously described (Cookson et al., 2003). Numbers of chloroplasts were determined from separated cells using a Leica DM5500 epifluorescence microscope. For transmission electron microscopy, outer pericarp tissues were fixed in 2% glutaraldehyde in 0.05 M Cacodylate buffer with picric acid for 2 h at 4°C, and then rinsed three times for 10 min each in 0.05 M Cacodylate buffer. Tissues were post fixed in 1% OsO_4 for 1 h at 4°C, rinsed again with Cacodylate buffer and dehydrated in an ethanol series. Tissues were infiltrated with EmBed-812 resin, embedded in molds and polymerized at 60°C for 48 h. Ultramicrotomed sections were viewed with a FEI Tecnai T12 Twin transmission electron microscope.

Transcriptome sequencing

Fruit were harvested at the stages indicated in the text from three biological replicates, each replicate contained 4-5 combined fruit of the same developmental stage and from the same plant. For gradient expression, fruit were equally cross-sectioned in five parts and pericarp tissues from the top (stem end), middle and bottom (stylar end) parts were kept for RNA extraction. The two intervening sections were discarded. Total RNA from pericarp tissues was used to construct strand-specific RNA libraries as described (Zhong et al., 2011). Two or three biological replicates for each sample were sequenced using a HiSeq2000 Sequencing System (Illumina, CA) according to the manufacturer's instructions.

RNA seq data processing and GO enrichment analysis

RNA-Seq reads were first processed by removing barcode and adaptor sequences. The resulting reads were aligned to ribosomal RNA (rRNA) and transfer RNA (tRNA) sequences using Bowtie allowing up to three mismatches (Langmead et al., 2009) to remove potential contaminating reads. The resulting cleaned reads were then aligned to the tomato genome using Tophat allowing one segment mismatch (Trapnell et al., 2009). Following alignments, raw counts of mapped reads for each tomato gene model were derived and then normalized to reads per kilobase of exon model per million mapped reads (RPKM). Differentially expressed genes between transgenic and control tissue were identified using the DESeq package (Anders and Huber, 2010). Raw p values were adjusted for multiple testing using the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995). Differentially expressed genes were then fed to the Plant MetGenMAP system to identify enriched GO terms (Joung et al., 2009). Dr. Fei helped me analyze the RNA-seq data.

Accession numbers

Full-length cDNA sequences of tomato *SlGLK1* and *SlGLK2* can be found at Genbank with accession numbers JQ316460 and JQ316459, respectively. Transcriptome sequencing reads were submitted to NCBI Sequence Read Archive (accession number SRA079879 for the *GLK* overexpression fruit tissues and accession number SRA080322 for the cross-sectioned fruit tissues).

REFERENCES

Alba, R., Payton, P., Fei, Z.J., McQuinn, R., Debbie, P., Martin, G.B., Tanksley, S.D., and Giovannoni, J.J. (2005). Transcriptome and selected metabolite analyses reveal multiple points of ethylene control during tomato fruit development. Plant Cell 17: 2954-2965.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biology 11: R106.

Baldwin, E.A., Goodner, K., and Plotto, A. (2008). Interaction of volatiles, sugars, and acids on perception of tomato aroma and flavor descriptors. Journal of Food Science 73: S294-307.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological) 57: 289-300.

Benvenuto, G., Formiggini, F., Laflamme, P., Malakhov, M., and Bowler, C. (2002). The photomorphogenesis regulator DET1 binds the amino-terminal tail of histone H2B in a nucleosome context. Current Biology 12: 1529-1534.

Blanke, M.M., and Lenz, F. (1989). Fruit photosynthesis. Plant, Cell & Environment 12: 31-46.

Bohn, G., and Scott, D. (1945). A second gene for uniform unripe fruit color in the tomato. Journal of Heredity 36: 169-172.

Carrara, S., Pardossi, A., Soldatini, G.F., Tognoni, F., and Guidi, L. (2001). Photosynthetic Activity of Ripening Tomato Fruit. Photosynthetica 39: 75-78.

Cookson, P.J., Kiano, J.W., Shipton, C.A., Fraser, P.D., Romer, S., Schuch, W., Bramley, P.M., and Pyke, K.A. (2003). Increases in cell elongation, plastid compartment size and phytoene synthase activity underlie the phenotype of the high pigment-1 mutant of tomato. Planta 217: 896-903.

Enfissi, E.M., Barneche, F., Ahmed, I., Lichtle, C., Gerrish, C., McQuinn, R.P., Giovannoni, J.J., Lopez-Juez, E., Bowler, C., Bramley, P.M., and Fraser, P.D. (2010). Integrative transcript and metabolite analysis of nutritionally enhanced DE-ETIOLATED1 downregulated tomato fruit. Plant Cell 22: 1190-1215.

Eshed, Y., and Zamir, D. (1995). An introgression line population of Lycopersicon pennellii in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. Genetics 141: 1147-1162.

Fitter, D.W., Martin, D.J., Copley, M.J., Scotland, R.W., and Langdale, J.A. (2002). GLK gene pairs regulate chloroplast development in diverse plant species. Plant Journal 31: 713-727.

Galpaz, N., Wang, Q., Menda, N., Zamir, D., and Hirschberg, J. (2008). Abscisic acid deficiency in the tomato mutant high-pigment 3 leading to increased plastid number and higher fruit lycopene content. Plant Journal 53: 717-730.

Gapper, N.E., McQuinn, R.P., and Giovannoni, J.J. (2013). Molecular and genetic regulation of fruit ripening. Plant Molecular Biology 82: 575-591.

Gillaspy, G., Ben-David, H., and Gruissem, W. (1993). Fruits: A Developmental Perspective. Plant Cell 5: 1439-1451.

Gimenez, E., Pineda, B., Capel, J., Anton, M.T., Atares, A., Perez-Martin, F., Garcia-Sogo, B., Angosto, T., Moreno, V., and Lozano, R. (2010). Functional analysis of the Arlequin mutant corroborates the essential role of the Arlequin/TAGL1 gene during reproductive development of tomato. PLoS One 5: e14427.

Giovannoni, J.J. (2004). Genetic regulation of fruit development and ripening. Plant Cell 16: S170-180.

Hetherington, S.E., Smillie, R.M., and Davies, W.J. (1998). Photosynthetic activities of vegetative and fruiting tissues of tomato. Journal of Experimental Botany 49: 1173-1181.

Itkin, M., Seybold, H., Breitel, D., Rogachev, I., Meir, S., and Aharoni, A. (2009). TOMATO AGAMOUS-LIKE 1 is a component of the fruit ripening regulatory network. Plant Journal 60: 1081-1095.

Joung, J.G., Corbett, A.M., Fellman, S.M., Tieman, D.M., Klee, H.J., Giovannoni, J.J., and Fei, Z. (2009). Plant MetGenMAP: an integrative analysis system for plant systems biology. Plant Physiology 151: 1758-1768.

Klee, H.J., and Giovannoni, J.J. (2011). Genetics and control of tomato fruit ripening and quality attributes. Annual Reviews of Genetics 45: 41-59.

Klee, H.J., and Tieman, D.M. (2013). Genetic challenges of flavor improvement in tomato. Trends in Genetics.

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

Lichtenthaler, H.K. (1987). Chlorophylls and Carotenoids - Pigments of Photosynthetic Biomembranes. Methods in Enzymology 148: 350-382.

Lin, N.C., Abramovitch, R.B., Mm, Y.J., and Martin, G.B. (2006). Diverse AvrPtoB homologs from several Pseudomonas syringae pathovars elicit Pto-dependent resistance and have similar virulence activities. Applied and Environmental Microbiology 72: 702-712.

Liu, Y., Roof, S., Ye, Z., Barry, C., van Tuinen, A., Vrebalov, J., Bowler, C., and Giovannoni, J. (2004). Manipulation of light signal transduction as a means of modifying fruit nutritional quality in tomato. Proceedings of the National Academy of Sciences 101: 9897-9902.

MacArthur, J.W. (1928). Linkage studies with the tomato. II. Three linkage groups. Genetics 13: 410.

Mustilli, A.C., Fenzi, F., Ciliento, R., Alfano, F., and Bowler, C. (1999). Phenotype of the tomato high pigment-2 mutant is caused by a mutation in the tomato homolog of DEETIOLATED1. Plant Cell 11: 145-157.

Nakamura, H., Muramatsu, M., Hakata, M., Ueno, O., Nagamura, Y., Hirochika, H., Takano, M., and Ichikawa, H. (2009). Ectopic overexpression of the transcription factor OsGLK1 induces chloroplast development in non-green rice cells. Plant and Cell Physiology 50: 1933-1949.

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

Powell, A.L., Nguyen, C.V., Hill, T., Cheng, K.L., Figueroa-Balderas, R., Aktas, H., Ashrafi, H., Pons, C., Fernandez-Munoz, R., Vicente, A., Lopez-Baltazar, J., Barry, C.S., Liu, Y., Chetelat, R., Granell, A., Van Deynze, A., Giovannoni, J.J., and Bennett, A.B. (2012). Uniform ripening encodes a Golden 2-like transcription factor regulating tomato fruit chloroplast development. Science 336: 1711-1715.

Rossini, L., Cribb, L., Martin, D.J., and Langdale, J.A. (2001). The maize golden2 gene defines a novel class of transcriptional regulators in plants. Plant Cell 13: 1231-1244.

Schroeder, D.F., Gahrtz, M., Maxwell, B.B., Cook, R.K., Kan, J.M., Alonso, J.M., Ecker, J.R., and Chory, J. (2002). De-etiolated 1 and damaged DNA binding protein 1 interact to regulate Arabidopsis photomorphogenesis. Current Biology 12: 1462-1472.

Tanksley, S., and Hewitt, J. (1988). Use of molecular markers in breeding for soluble solids content in tomato—a re-examination. Theoretical and Applied Genetics 75: 811-823.

Tanksley, S., Mutschler, M., and Rick, C. (1990). Linkage map of the tomato (Lycopersicon esculentum)(2N= 24), pp. 6.3-6.15.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25: 1105-1111.

Turgeon, R., and Medville, R. (1998). The absence of phloem loading in willow leaves. Proceedings of the National Academy of Sciences 95: 12055-12060.

Vrebalov, J., Pan, I.L., Arroyo, A.J., McQuinn, R., Chung, M., Poole, M., Rose, J., Seymour, G., Grandillo, S., Giovannoni, J., and Irish, V.F. (2009). Fleshy fruit expansion and ripening are regulated by the Tomato SHATTERPROOF gene TAGL1. Plant Cell 21: 3041-3062.

Wang, P., Fouracre, J., Kelly, S., Karki, S., Gowik, U., Aubry, S., Shaw, M.K., Westhoff, P., Slamet-Loedin, I.H., and Quick, W.P. (2013). Evolution of GOLDEN2-LIKE gene function in C3 and C4 plants. Planta 237: 481-495.

Waters, M.T., Wang, P., Korkaric, M., Capper, R.G., Saunders, N.J., and Langdale, J.A. (2009). GLK transcription factors coordinate expression of the photosynthetic apparatus in Arabidopsis. Plant Cell 21: 1109-1128.

Yeager, T. (1935). The uniform fruit color gene in the tomato. Proceedings of the American Society for Horticultural Science 33: 512.

Zhong, S., Joung, J.G., Zheng, Y., Chen, Y.R., Liu, B., Shao, Y., Xiang, J.Z., Fei, Z., and Giovannoni, J.J. (2011). High-throughput illumina strand-specific RNA sequencing library preparation. Cold Spring Harbor Protocols 2011: 940-949.

CHAPTER III.

FRUIT QUALITY ANALYSIS OF TOMATO NEARLY ISOGENIC LINES FOR THE GENE CONFERRING THE GREEN SHOULDER PHENOTYPE

Abstract

Tomato is among the most important vegetable crops in the world providing sugars, fibers, vitamins and antioxidants to the human diet. Consumers demand high nutritional and flavorful tomatoes and often find modern fruit deficient in organoleptic properties. The *uniform ripening* (*u*) mutation has been extensively deployed in almost all modern processing cultivars, in addition to fresh consumption tomato cultivars, as a means to better (more uniform) appearance. However, as shown in the previous chapter, the gene coding for *u* is a *Golden 2-like* transcription factor effecting chlorophyll content and chloroplast development, thus influencing sugar, carotenoid and ascorbic acid levels in red ripe fruit. The knockout mutation of *SlGLK2* represented by the *u* allele thus confers an unintended consequence in compromising the quality and nutritional value of tomatoes. While the use of GMOs such as *GLK* overexpression for commercial tomato production is currently controversial, it remains possible to elevate nutrient quality of tomato fruit via use of the wild-type *U* allele in modern varieties. As a first step in demonstrating the lost value of the *U* allele, and for confirming the transgenic results presented in the prior chapter using natural genetic variation, nearly isogenic lines that differ for the *u* locus in an elite variety *(M82),* and which contain tightly linked DNA sequence variation useful for molecular selection, were developed. While there is some return of off-color phenotypes in ripe fruit, *U/U* fruit elevate soluble solids as well as carotenoids by as much as 20% compared to *u/u*

fruit, and the line additionally contains DNA polymorphisms surrounding the *u* locus facilitating molecular assisted selection.

I. Introduction

The *uniform ripening (u)* mutation has been extensively used in tomato breeding to create tomato varieties with uniformly ripened fruit without the yellow discoloration characteristic of some green shoulder cultivars. *u* fruit also help in the selection of fruit stage for optimal timing of destructive harvest (as they are more even in appearance), which is common for tomato processing varieties in the US and elsewhere. As described in the previous chapter, *U* encodes a *Golden 2-like (GLK)* transcription factor and plays an important role in chlorophyll accumulation and chloroplast development. Increased *GLK* expression positively effects fruit nutritional quality. Hence, using the *u* allele in tomato breeding adds uniform fruit color but also leads to unintended consequences in that the nutritional value of the fruit may be reduced. *U*, though its contributions to distribution and content of chlorophyll and chloroplasts in green fruit, has related effects on fruit quality traits mediated by the chloroplasts and photosynthesis such as soluble solids, carotenoids and ascorbic acid. Indeed, our data showed that overexpression of *U/SlGLK* significantly increases chlorophyll and chloroplast levels in green fruit in addition to sugars, carotenoids and ascorbate in red fruit (Powell et al., 2012; Nguyen et al., 2014).

Consumers have long been concerned with the poor organoleptic properties of tomato fruit available through large suppliers such as supermarket chains where many individuals secure much of their fresh fruit and vegetables (Klee and Tieman, 2013). Horticultural practices designed to extend shelf-life and maintain firmness for long distance shipping, including

harvesting unripe fruit as well as use of genetic remedies including delayed ripening mutants like *rin* and the *u* mutation for even coloration, are at least partial contributors to this general problem. In the case of *uniform ripening* (*u*) (he focus of this thesis), the mutant has been extensively used in tomato breeding such that nearly all modern processing varieties and a large portion of fresh consumption varieties carry the homozygous mutation. Chloroplast numbers and chlorophyll content are effected *by* u and chloroplasts are the site of synthesis of many primary and secondary metabolites (e.g., sugars, carotenoids, ascorbic acid, lipids and almost all amino acids), including precursors of flavor-related volatiles. During tomato fruit ripening, chlorophylls are broken down and chloroplasts change their structures to become chromoplasts where the carotenoids (mostly lycopene and β-carotene) accumulate. Perturbations in chlorophyll and chloroplast development in the green fruit as mediated by *u*, negatively effects fruit nutrition and flavor quality.

Overexpression of *GLKs* increases these same nutritional quality traits as noted in the prior chapter. When using the largely constitutive 35S promoter to drive the expression of *SlGLKs*, the dark green color of unripe fruit normally restricted to the top of wild-type fruit now spans the entirety of the unripe transgenic *SlGLKs* overexpressing fruit (Figure 2.6, 2.8). These fruit ripen normally and typically mature to full ripe absent the development of off colors. As the fruit are rightly considered GMO products they would be difficult and expensive to clear for regulatory approval and marketing. At present, the most practical way to exploit the potential benefits of altered *SlGLKs* is to deploy available genetic variation and use useful alleles of the gene through traditional breeding, which may result in reintroduction of the green shoulder phenotype. It is worth noting that this phenotype is currently associated with heirloom varieties that are perceived as having higher quality by some consumers. In addition, not all green

shoulder fruit develop off colors and the potential of improving fruit quality by using green shoulder varieties may compensate for this problem.

My goal in this portion of my thesis project was to develop a non-transgenic system to study the effect of the uniform/green shoulder phenotype on fruit quality traits. *Craigella* is a line derived from *'Ailsa Craig'* through five back-crosses and homozygous for the *u* allele, but records on the development and purity of this line are limited and likely included selection for additional traits (Tomato Genetics Resource Center, http://tgrc.ucdavis.edu). The objective here was to create more nearly isogenic lines differing at just the *u* locus and in a modern elite cultivar so that the effect of the mutation on fruit quality could be measured accurately. The development of nearly isogenic lines for the *u* locus and quality assessment of their fruit is presented here.

II. Results and discussion

Development of nearly isogenic lines for the u locus

To assess the specific effects of the *uniform*/green shoulder (*u*) allele, isogenic lines that differ for as little of the genomic region surrounding the gene/allele are needed. Ideally the lines also should have a relevant modern genetic background as close to utilize commercial varieties as possible. *S. pennellii* introgression lines (ILs) (Eshed and Zamir, 1995) that contain introgressed segments from chromosome 10 (spanning where *U/SlGLK2* resides) represent starting material that meet these criteria. *'M82'* was an elite commercial processing cultivar and germplasm developed from pyramiding several of these ILs have been widely used in the central valley of California for large-scale production of processing tomatoes. The drawback with these ILs is that the introgressed segments can be large (up to the entire chromosome), containing hundreds to thousands of genes from the wild species. This "linkage drag" issue is common

when bringing traits from wild species, where the locus of interest may be linked to undesirable genes that can negatively affect important agronomical traits (Young and Tanksley, 1989).

Two *s. pennellii* ILs, 10-1 and 10-1-1, present the green shoulder phenotype: IL10-1 has a very large segment from chromosome 10 of *S. pennellii* (Eshed and Zamir, 1995) spanning more than 55 Mb, while IL 10-1-1 has a relatively small 2.6 Mb introgressed segment (still containing hundreds of genes from *S. pennellii*) (Chitwood et al., 2013). Linkage drag was addressed by additional back-crossing and use of molecular markers to select for recombination events near *U*.

Genetic marker mapping using IL 10-1 and *'M82'* showed that chromosome 10 of *S. pennellii* and S. *lycopersicum* have typical (non-suppressed) recombination rates, with recombination events breaking linkages surrounding the *u* locus. From the mapping population characterized from this cross of IL 10-1 and *'M82'*, a line with a recombination event close to the *u* locus was chosen for back-crossing with *'M82'* to develop a new segregating population with reduced *S. pennellii* DNA in this region (Figure 3.1). The next goal was to select for a second recombination event at the other side of the *u* locus. Based on a genetic map of *S. pennellii* (Fulton et al., 2002), CAPS markers which are ~1, ~3 and ~5 cM away from the *u* locus were developed and a total of ~ 700 plants were scored with these markers. Three recombinant plants with cross-over events closest to *U* were chosen for further analysis. These plants were self-pollinated and homozygous progeny were selected using the T787 marker, which is ~50kb to the distal side of the *U* locus from the prior flanking recombination event noted above.

To determine more precisely the position of the recombination events and the size of the *S. pennellii* introgression distinguishing the NILs, RNA from immature green fruit tissues were sequenced by strand specific Illumina sequencing (Zhong et al., 2011) and SNPs were determined from the resulting sequence data (see Figure 3.2). The positions of boundary SNPs on both sides of the *U* locus were used to approximate corresponding *S. pennellii* introgressed segments sizes. The smallest introgressed segment was estimated at 180 kb in line NIL530, and the largest at ~770kb in line NIL414. All lines analyzed had introgressions significantly smaller compared to the introgressed sizes of IL10-1 and IL 10-1-1, which are ~55Mb and 2.5 Mb, respectively (Chitwood et al., 2013). The smallest (180 kb) NIL has 22 predicted genes from *S. pennellii*, none with predicted functions in chlorophyll and chloroplast biology other than *U/SlGLK2* (Table 3.1)

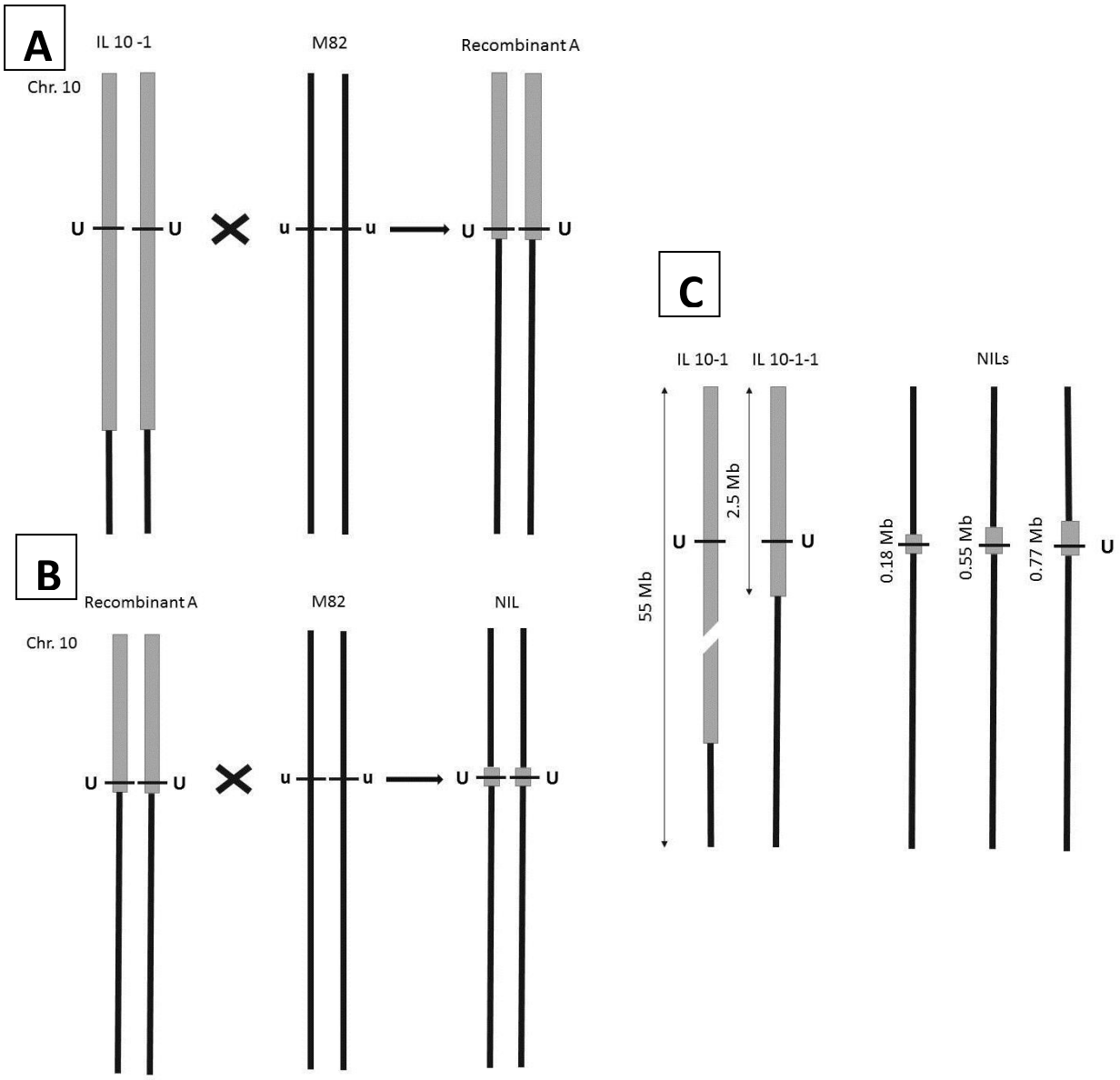


Figure 3.1. Illustration of development of near isogenic lines (NILs) for the *u* locus.

A. F2 population of a cross between IL10-1 and *'M82'*, selected for recombinants as close to *u* as possible through the use of molecular markers.

B. Starting with the recombinant defined in A, selection was performed for recombination events on the distal side of the *u* locus.

C. Representation of the small introgression *U* NILs developed following backcrossing to *'M82'* and selection with molecular markers. The left most line with the calculated 0.18 Mb introgression harbors only 22 *S. pennellii* genes.

Table 3.1. List of genes in introgressed segments NIL 530

Gene	Annotation version 2.40
Solyc10g007980	50S ribosomal protein L5 (AHRD V1 ***- RL5_XANP2); contains Interpro domain(s) IPR002132 Ribosomal protein L5
Solyc10g007990	Cupin RmlC-type (AHRD V1 **-- B6T568_MAIZE); contains Interpro domain(s) IPR012864 Protein of unknown function DUF1637
Solyc10g008000	Light-dependent short hypocotyls 1 (AHRD V1 ***- D7M6V0_ARALY); contains Interpro domain(s) IPR006936 Protein of unknown function DUF640
Solyc10g008010	Proteasome subunit alpha type (AHRD V1 ***- Q38M52_SOLTU); contains Interpro domain(s) IPR001353 Proteasome, subunit alpha/beta
Solyc10g008020	Methyltransferase (AHRD V1 *-*- B8H501_CAUCN); contains Interpro domain(s) IPR013216 Methyltransferase type 11
Solyc10g008030	Os06g0207500 protein (Fragment) (AHRD V1 ***- Q0DDQ9_ORYSJ); contains Interpro domain(s) IPR004253 Protein of unknown function DUF231, plant
Solyc10g008040	Seed biotin-containing protein SBP65 (AHRD V1 *-*- SBP65_PEA); contains Interpro domain(s) IPR006777 Microvirus H protein (pilot protein)
Solyc10g008050	Unknown Protein (AHRD V1)
Solyc10g008060	LOC445834 protein (Fragment) (AHRD V1 **-- Q6DF75_XENLA); contains Interpro domain(s) IPR007402 Protein of unknown function DUF455
Solyc10g008070	Bromodomain-containing RNA-binding protein 2 (AHRD V1 *-*- Q5K541_TOBAC)
Solyc10g008080	Ring H2 finger protein (AHRD V1 ***- D9ZHD8_HYPPE); contains Interpro domain(s) IPR018957 Zinc finger, C3HC4 RING-type
Solyc10g008090	Oligosaccharide transporter (AHRD V1 **-* B4FSV8_MAIZE); contains Interpro domain(s) IPR006844 OST3/OST6
Solyc10g008100	Glucosidase 2 subunit beta (AHRD V1 *--- B5X0Z0_SALSA)
Solyc10g008110	Acyl-CoA oxidase 6 (AHRD V1 **** D7KG21_ARALY); contains Interpro domain(s) IPR012258 Acyl-CoA oxidase
Solyc10g008120	O-methyltransferase 3 (AHRD V1 ***- B0ZB57_HUMLU); contains Interpro domain(s) IPR016461 O-methyltransferase, COMT, eukaryota
Solyc10g008130	Pre-rRNA-processing protein ESF2 (AHRD V1 ***- D0MVG4_PHYIN); contains Interpro domain(s) IPR012677 Nucleotide-binding, alpha-beta plait
Solyc10g008140	Prohibitin 1-like protein (AHRD V1 ***- Q9AXM0_BRANA); contains Interpro domain(s) IPR000163 Prohibitin
Solyc10g008150	Glutaredoxin (AHRD V1 ***- B9MYC1_POPTR); contains Interpro domain(s) IPR011905 Glutaredoxin-like, plant II
Solyc10g008160	SlGLK2
Solyc10g008170	26S proteasome regulatory subunit (AHRD V1 ***- C6HL17_AJECH)
Solyc10g008180	26S proteasome regulatory subunit (AHRD V1 ***- A8J3A4_CHLRE); contains Interpro domain(s) IPR016643 26S proteasome regulatory complex, non-ATPase subcomplex, Rpn1 subunit
Solyc10g008190	OB-fold nucleic acid binding domain containing protein (AHRD V1 ***- B6SHT0_MAIZE); contains Interpro domain(s) IPR012340 Nucleic acid-binding, OB-fold

Fruit nutritional quality comparisons of u NIL lines

U NIL plants developed with molecular screening for recombination events minimizing *S. pennellii* introgression described above were used to assess differences in fruit nutritional quality resulting from the *u* mutation. All three *U/U* NILs (NIL530, NIL559, and NIL514) and *u/u 'M82'* control plants were grown in the greenhouse and in the field. 20-30 plants for each line were grown in randomized blocks and red fruit were harvested for analysis of off coloring (white or yellow patches) and quality traits (total soluble solids [TSS] and carotenoids).

One of the main reasons for selection against the green shoulder phenotype is the uneven coloration developing during ripening of some *U/U* fruit (Figure 2.1, Figure 3.2). These yellow spots reflect reduced lycopene accumulation. Indeed, when comparing the carotenoid profiles of the yellow and red sectors of *U/U* fruit, lycopene levels are significantly reduced (Figure 3.6). In this experiment, fruit from greenhouse plants showed minimal off-coloring with just 5-7 % of the fruit presenting white or yellow patches. In the field, off-colored fruit frequency was substantially higher at 15-25%. Interestingly, 8-10% of fruit from the *u/u 'M82'* control also presented yellow spots suggesting additional factors effect shoulder color patterning. The higher percentage of patchy coloration in the field compared to the greenhouse may result from higher light intensity. The degree of dark green coloration on the shoulder correlated with the amount of light received (Figure 3.2) in that fruit that were exposed to direct sun light had darker green coloration as compared to fruit shaded by leaves. During ripening, chlorophylls are broken down and chloroplasts convert to chromoplasts (Camara et al., 1995). High chlorophyll and chloroplast levels coupled with high light intensity may interfere with chromoplast formation and lycopene accumulation, resulting in off coloring of fruit though the nature of plastids in the light colored patches of *U/U* or *u/u* fruit remain to be thoroughly examined.

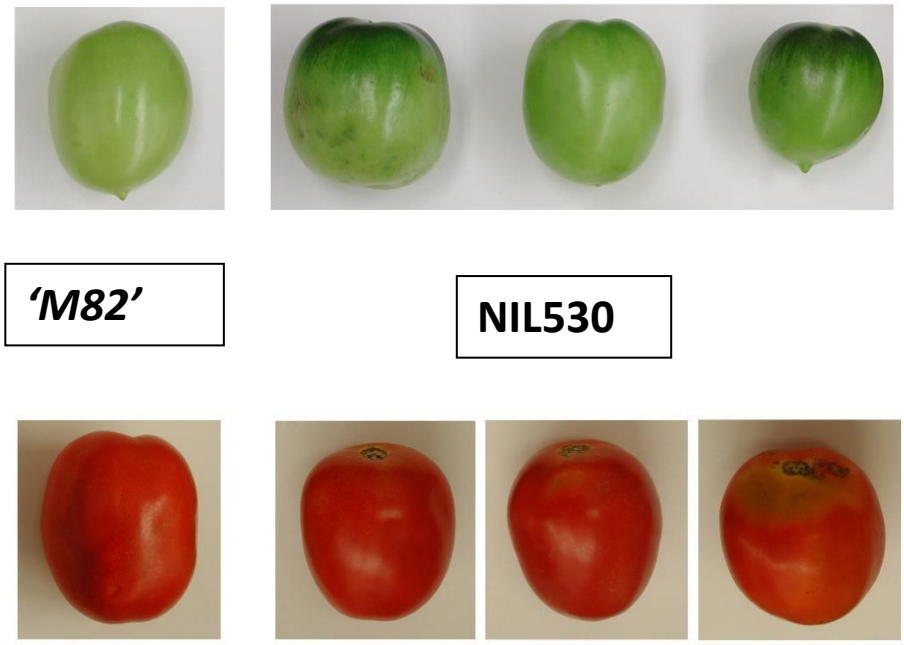


Figure 3.2. Mature green and ripe fruit of *'M82'* and NIL530 cultivars grown in the Freeville, NY Cornell University field showing variable dark green shoulder and off-colored regions in ripe fruit.

Table 3.2. Percent of fruit with off-color segment development (white to yellow regions).

	'M82'	NIL530	NIL 559	NIL514	average
Field	10.80%	23.30%	18.75%	23.60%	22.20%
Greenhouse	<1%	6.30%	5.10%	7.30%	6.20%

Resulting data showed that green shoulder *U/U* fruit develop off-coloration on the shoulder 5 – 15 % more often than their *u/u* counterparts. This negative visual effect was compensated by the development of higher levels of nutrient/quality metabolites. Indeed, red fruit harvested from both the greenhouse and field accumulated higher soluble solids and total carotenoids as compared to the *'M82' u/u* control (Figures 3.3and 3.4). ^{o}Brix, a measure of total soluble solids, a primary quality trait for processing tomatoes, increased 8-12% under greenhouse conditions and 10-15% in the field (Figure 3.4). Similarly, lycopene levels were also 10-20% higher in the *u/u* NILs compared to *'M82'*. β-carotene in *'M82'* was low (~1-2% of total carotenoids) as compared to other varieties (e.g. *'Ailsa Craig'*, see Nguyen et al., 2013) but was still elevated in all three NILs (Figure 3.4). When all other carotenoids are included, total carotenoid levels are 10-20% higher in the NILs as compared to *'M82'*, suggesting the entire carotenoid pathway is elevated, as opposed to a subset of specific carotenoids These data are consistent with those resulting from transgenic plants overexpressing *SlGLKs*, although fruit overexpressing *SlGLKs* have higher ^{o}Brix (20-70% compared to control) in addition to larger increases in β-carotene and lutein levels (Powell et al., 2012; Nguyen et al., 2014 and prior chapter). The mutant data presented here confirm the transgenic effect of over-expressing *U/SlGLK2* gene on fruit quality and point directly to the unintended fruit metabolite quality consequence of deploying the *u* mutation in modern tomato breeding. These results also suggest that greenhouse production and/or reduced lighting may mitigate some of the appearance issues resulting from the wild-type *U* allele, implying that elevated quality and appearance are not mutually exclusive.

Two heirloom varieties differing at the *u* locus, *'Ailsa Craig'* (*U/U*) and *Craigella* (*u/u*) were also grown under greenhouse conditions at the same time as the *'M82'*/NILs experiment

and measured for °Brix and carotenoid composition. Consistent with the *'M82'*/NIL results, *U/U 'Ailsa Craig'* fruit have higher soluble solids and carotenoid compared to *u/u Craigella* (Figure 3.6), suggesting that the effects of *U/SlGLK2* are broad and not variety specific.

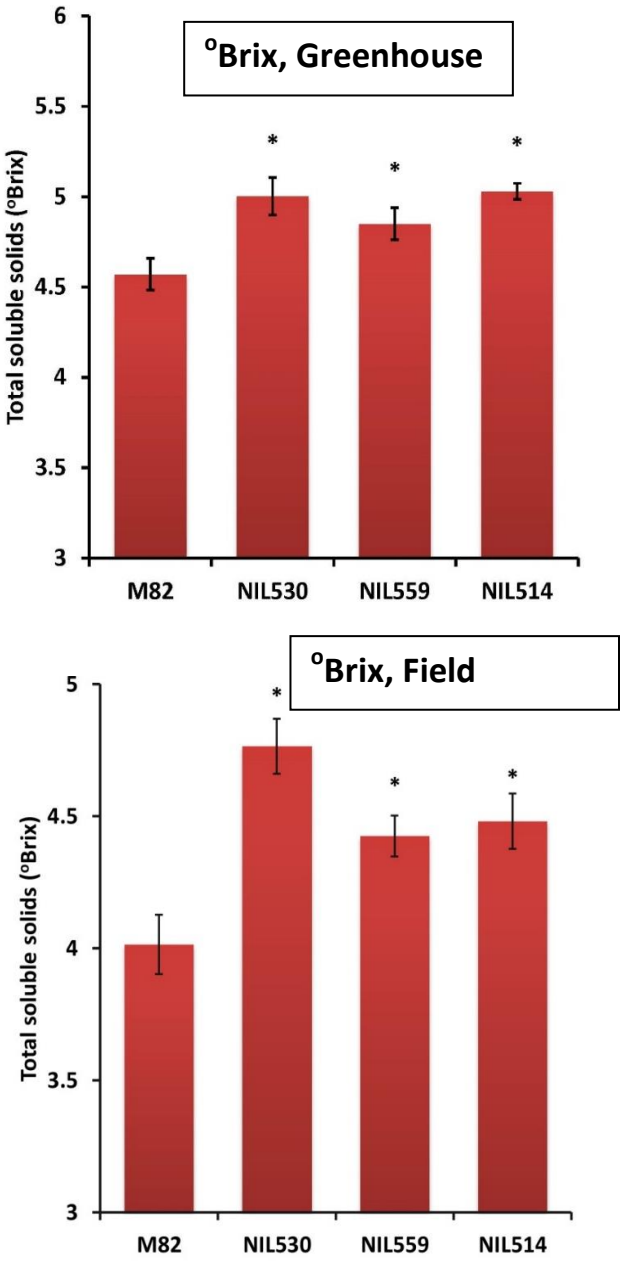


Figure 3.3. °Brix levels in *''M82''* (*u/u*) and NIL (*U/U*) fruit harvested from field and greenhouse trials. n=30, *: p value t-test (NILs vs. *'M82''*) ≤ 0.05. Error bars are S.E.

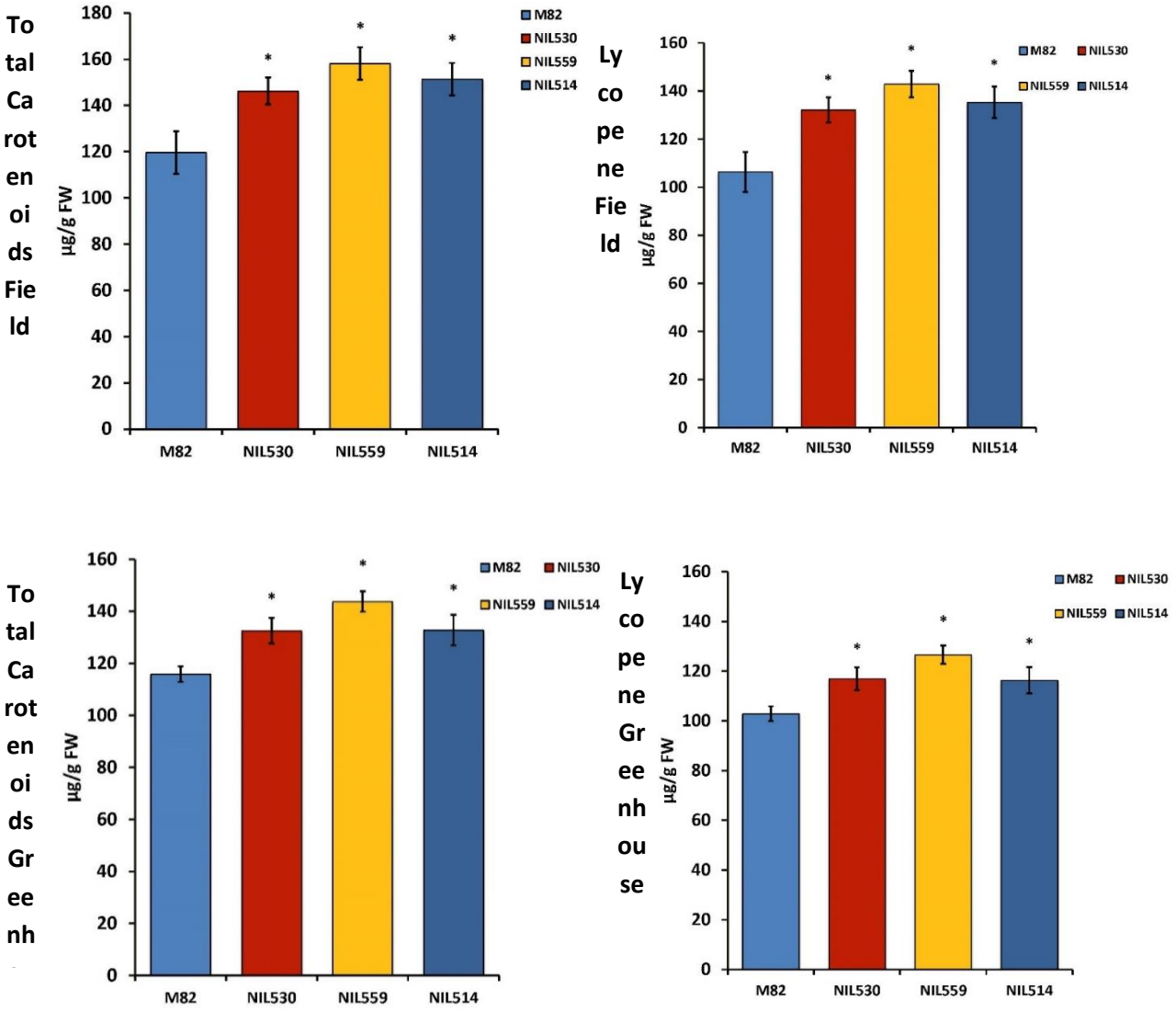


Figure 3.4. Carotenoid levels of *''M82''* (*u/u*) and NIL (*U/U*) fruit from the field and greenhouse. n=10. *: p value t-test (NIL vs. *''M82''*) ≤ 0.05. Error bars are S.E.

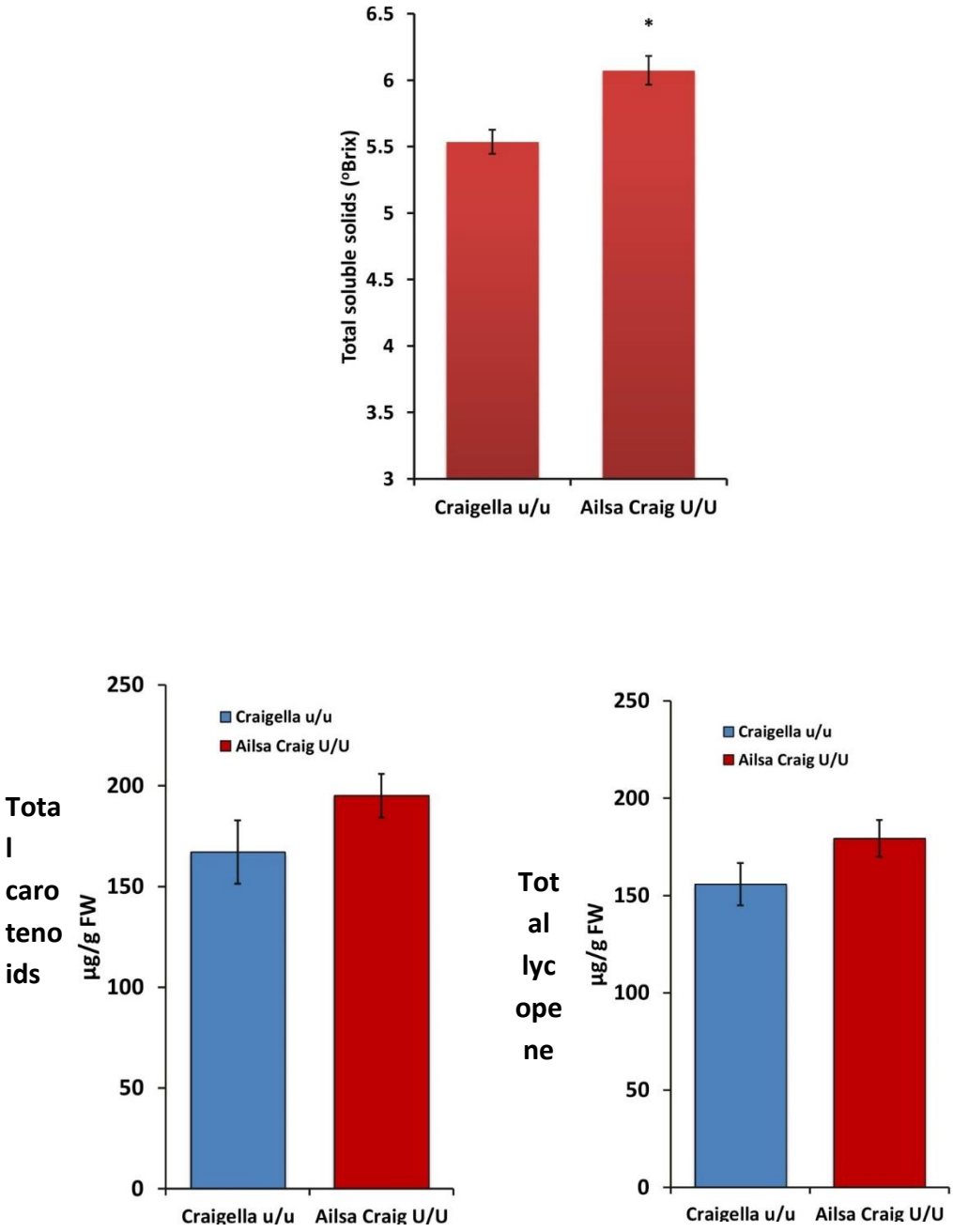
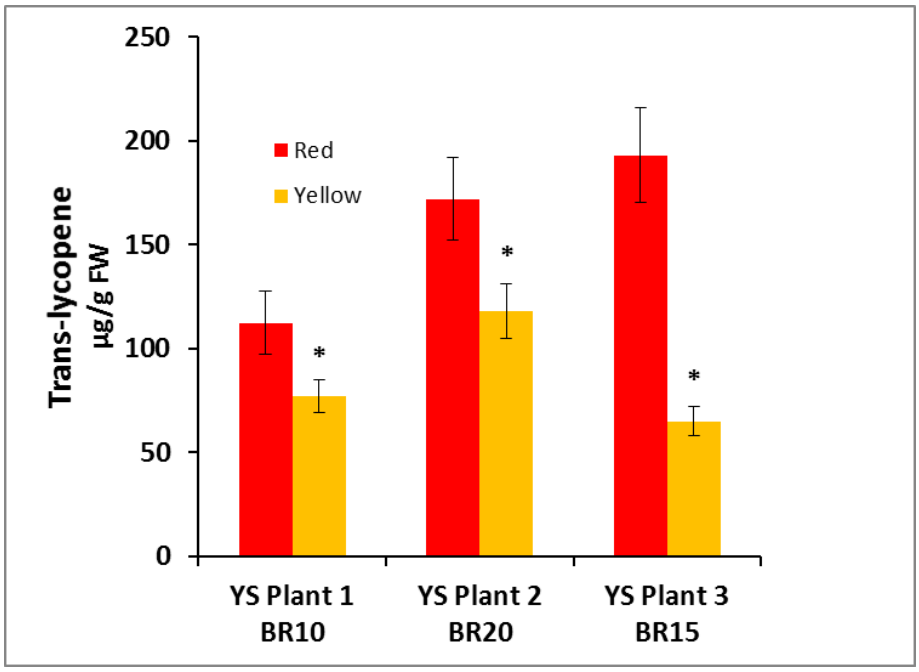


Figure 3.5. °Brix and carotenoid levels in fruit from *''Craigella"* (*u/u*) and *'Ailsa Craig'* U/U grown under greenhouse conditions. n=30 for Brix, n=12 for carotenoid measurements,. *: p value t-test (NILs vs. '*'M82*'') < 0.05. Error bars are S.E.



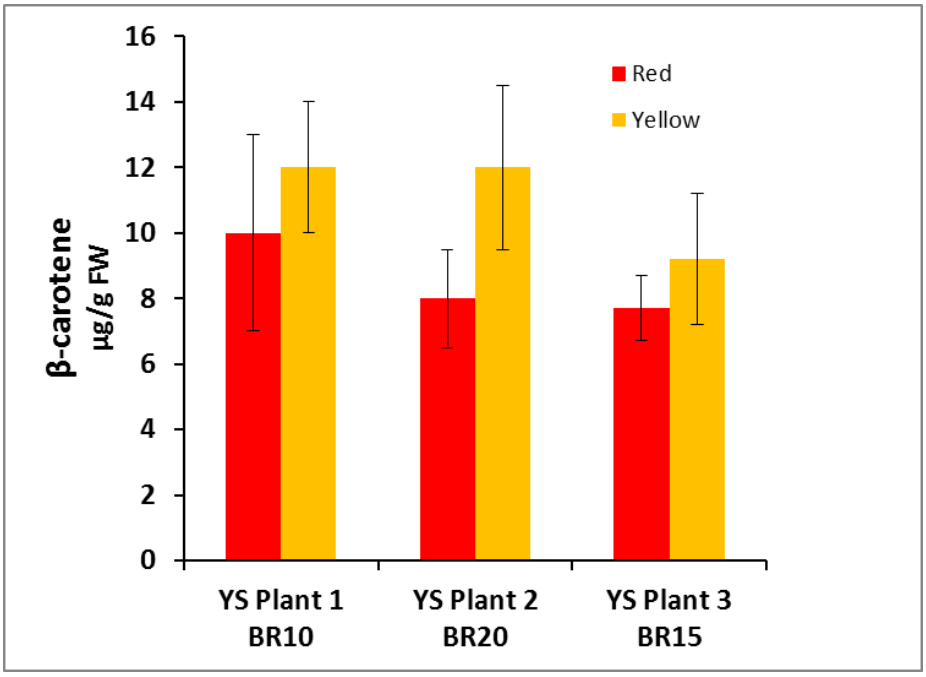


Figure 3.6. Carotenoids (trans-lycopene and beta-carotene) extracted from off-color (YS) and red segments of fruit developing shoulder pattern coloration. YS: Yellow shoulder fruit of NILs. BR10, BR15, BR20: Breaker +10, 15, 20 days, respectively. *: p value t-test (NILs vs. '*'M82*") < 0.05.

For processing tomatoes, elevated total soluble solids and carotenoids are desirable as they impact paste quality and are among the major attributes sought by growers and processors. Higher soluble solids translate to more paste and product at a lower energy/production cost while more lycopene (richer red coloration) make the products more visually appealing. A single gene with significant effects on both represents substantial value potential in the $2 billion US tomato market, and even greater potential in the worldwide tomato economy.

NILs as the source for U molecular markers

To select more efficiently for the green shoulder phenotype, seedling screens with molecular markers can now be readily deployed. The basis of the *u* allele mutation is an insertion of an adenine (A) to a six base stretch of adenine sequence (Figure 2.2). The nature of this mutation and the mutation position make it difficult to identify useful restriction enzymes and/or develop PCR based markers (which are preferable for tomato breeders with limited resources). Even with sequencing of PCR products, it can be difficult to differentiate heterozygous/homozygous genotypes. In the case of the NILs developed here, as the introgressed segments are from *S. pennellii*, there are many polymorphisms within and linked to the *U* gene itself facilitating design of PCR or other molecular markers. For example, in developing these NILs, a CAPS marker which is 30 kb from *U* was developed to select for homozygous *U/U* genotypes (Figure 3.1). This marker can be easily scored with *AluI* restriction enzyme and gel electrophoresis. With the genome sequence of *S. pennellii* now complete (Bolger et al., 2014), additional polymorphisms can be chosen at the gene itself (for example, introns) or in promoter sequences, such that the probability of inaccurate scoring due to recombination approaches zero.

Transcriptome of immature green NIL fruit

Transcriptome profiles resulting from RNA-seq of NIL immature green fruit tissue was compared to corresponding *'M82' u* control tissue to identify differential gene expression resulting from *U* activity. Differentially expressed genes were subject to GO term analysis as outlined in chapter 2. As expected, most of the differentially expressed genes have functions related to chlorophyll synthesis, photosynthesis, and chloroplast development (Table 3.3). These results are similar to those observed with the overexpression experiments described in the prior chapter and again confirm the function of GLK in chlorophyll and chloroplast development. It is noteworthy that these data also support the "substantial equivalence" concept regarding the differences between a GMO and non-GMO product (Miller, 1999). While the choice of using *u* or *U* alleles remains at the discretion of the breeder, the results described here suggest there are substantial benefits for both growers and consumers in using the *U* allele, especially under growth conditions that limit the off-color shoulder phenotype. A determining factor may be the extent to which consumers are receptive of the green shoulder or off-color patterns of ripe fruit. It is worth noting that some consumers are beginning to associate differential shoulder coloration with heirloom varieties commanding premium prices.

III. Material and methods

Plant materials.

Tomato plants (*Solanum* sps.*)* were grown in fields and greenhouses in Ithaca, NY, USA. The *S. lycopersicum* varieties *'Ailsa Craig'* (LA3736), the monogenic *u* mutant of *'Ailsa Craig'* called *'Craigella'* (LA3247) (*36*), IL 10-1 and IL10-1-1 were germinated from stock seed collections

(Tomato Genetics Resource Center, UC Davis), transplanted and grown in two gallon pots in greenhouses or in furrowed irrigated fields.

Marker assisted selection

From a F2 mapping population derived from the cross IL10-1 x *'M82'*, a line with green shoulder fruit and a recombination event close to the *u* locus was identified. This line was subsequently crossed with *'M82'* to make a F2 segregating population. CAPS markers which are ~1, ~3 and ~5 cM away from the *u* locus (based on *S. pennellii* IL genetic map, http://solgenomics.net) were screened in a population of ~ 700 plants. Three recombinant plants with cross-over events closest to *u* were chosen for further analysis. These plants were self-pollinated and homozygous progeny were selected using marker T787, which is ~50kb away on the distal side of the *U* locus relative to the recombination noted above.

Soluble solids analysis

Soluble solids (^{o}Brix) from fruit juice were measured with a digital refractometer (HR200, APT Co, LLC). Fruit were cut in half and soluble solids were obtained by squeezing a few drops of juice directly on to the refractometer.

Carotenoid and chlorophyll analysis

Carotenoid and chlorophyll extraction and carotenoid quantification by High Pressure Liquid Chromatography (HPLC) were performed as described (Vrebalov et al., 2009).

Transcriptome sequencing

Fruit were harvested at the stages indicated in the text from three biological replicas with each containing 4-5 combined fruit of the same developmental stage and from the same plant.

Pericarp tissues (without locule, septa and columella tissues) were frozen in liquid nitrogen. Total RNA from frozen pericarp tissue was used to construct strand-specific RNA libraries as described (Zhong et al., 2011). Two or three biological replicates for each sample were sequenced using a HiSeq2000 Sequencing System (Illumina, CA) according to the manufacturer's instructions.

RNA-seq data processing and GO enrichment analysis

RNA-Seq reads were first processed by removing barcode and adaptor sequences. The resulting reads were aligned to ribosomal RNA (rRNA) and transfer RNA (tRNA) sequences using Bowtie allowing up to three mismatches (Langmead et al., 2009) to remove potential contaminating reads. The resulting cleaned reads were then aligned to the tomato genome using Tophat allowing one segment mismatch (Trapnell et al., 2009). Following alignments, raw counts of mapped reads for each tomato gene model were derived and then normalized to reads per kilobase of exon model per million mapped reads (RPKM). Differentially expressed genes between transgenic and control tissue were identified using the DESeq package (Anders and Huber, 2010). Raw p values were adjusted for multiple testing using the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995). Differentially expressed genes were then fed to the Plant MetGenMAP system to identify enriched GO terms (Joung et al., 2009). Samtools (Li et al., 2009) and SnpEff (Cingolani et al., 2012) were used to call SNP from RNA seq data. Dr. Zhangjun Fei helped with this analysis.

REFERENCES

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biology 11: R106.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological) 57: 289-300.

Bolger, A., Scossa, F., Bolger, M.E., Lanz, C., Maumus, F., Tohge, T., Quesneville, H., Alseekh, S., Sorensen, I., Lichtenstein, G., Fich, E.A., Conte, M., Keller, H., Schneeberger, K., Schwacke, R., Ofner, I., Vrebalov, J., Xu, Y., Osorio, S., Aflitos, S.A., Schijlen, E., Jimenez-Gomez, J.M., Ryngajllo, M., Kimura, S., Kumar, R., Koenig, D., Headland, L.R., Maloof, J.N., Sinha, N., van Ham, R.C., Lankhorst, R.K., Mao, L., Vogel, A., Arsova, B., Panstruga, R., Fei, Z., Rose, J.K., Zamir, D., Carrari, F., Giovannoni, J.J., Weigel, D., Usadel, B., and Fernie, A.R. (2014). The genome of the stress-tolerant wild tomato species Solanum pennellii. Nature Genetics 46: 1034-1038.

Camara, B., Hugueney, P., Bouvier, F., Kuntz, M., and Monéger, R. (1995). Biochemistry and molecular biology of chromoplast development. International Review of Cytology 163: 175-247.

Chitwood, D.H., Kumar, R., Headland, L.R., Ranjan, A., Covington, M.F., Ichihashi, Y., Fulop, D., Jimenez-Gomez, J.M., Peng, J., Maloof, J.N., and Sinha, N.R. (2013). A quantitative genetic basis for leaf morphology in a set of precisely defined tomato introgression lines. Plant Cell 25: 2465-2481.

Cingolani, P., Platts, A., Wang le, L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) 6: 80-92.

Eshed, Y., and Zamir, D. (1995). An introgression line population of Lycopersicon pennellii in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. Genetics 141: 1147-1162.

Fulton, T.M., Van der Hoeven, R., Eannetta, N.T., and Tanksley, S.D. (2002). Identification, analysis, and utilization of conserved ortholog set markers for comparative genomics in higher plants. Plant Cell 14: 1457-1467.

Joung, J.G., Corbett, A.M., Fellman, S.M., Tieman, D.M., Klee, H.J., Giovannoni, J.J., and Fei, Z. (2009). Plant MetGenMAP: an integrative analysis system for plant systems biology. Plant Physiology 151: 1758-1768.

Klee, H.J., and Tieman, D.M. (2013). Genetic challenges of flavor improvement in tomato. Trends in Genetics 29: 257-262.

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

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

Miller, H.I. (1999). Substantial equivalence: Its uses and abuses. Nature Biotechology 17: 1042-1043.

Nguyen, C.V., Vrebalov, J.T., Gapper, N.E., Zheng, Y., Zhong, S., Fei, Z., and Giovannoni, J.J. (2014). Tomato GOLDEN2-LIKE transcription factors reveal molecular gradients that function during fruit development and ripening. Plant Cell 26: 585-601.

Powell, A.L., Nguyen, C.V., Hill, T., Cheng, K.L., Figueroa-Balderas, R., Aktas, H., Ashrafi, H., Pons, C., Fernandez-Munoz, R., Vicente, A., Lopez-Baltazar, J., Barry, C.S., Liu, Y., Chetelat, R., Granell, A., Van Deynze, A., Giovannoni, J.J., and Bennett, A.B. (2012). Uniform ripening encodes a Golden 2-like transcription factor regulating tomato fruit chloroplast development. Science 336: 1711-1715.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25: 1105-1111.

Vrebalov, J., Pan, I.L., Arroyo, A.J., McQuinn, R., Chung, M., Poole, M., Rose, J., Seymour, G., Grandillo, S., Giovannoni, J., and Irish, V.F. (2009). Fleshy fruit expansion and ripening are regulated by the Tomato SHATTERPROOF gene TAGL1. Plant Cell 21: 3041-3062.

Young, N.D., and Tanksley, S.D. (1989). RFLP analysis of the size of chromosomal segments retained around the Tm-2 locus of tomato during backcross breeding. Theor Appl Genet 77: 353-359.

Zhong, S., Joung, J.G., Zheng, Y., Chen, Y.R., Liu, B., Shao, Y., Xiang, J.Z., Fei, Z., and Giovannoni, J.J. (2011). High-throughput illumina strand-specific RNA sequencing library preparation. Cold Spring Harbor Protocols 2011: 940-949

CHAPTER IV.

CHARACTERIZATION OF MOLECULAR GRADIENTS IN FRUIT DEVELOPMENT

Abstract

A latitudinal gradient of *SlGLK2* expression influences the typical uneven coloration of green and ripe wild-type tomato fruit. Transcriptome profiling revealed a broader fruit gene expression gradient throughout development suggesting additional molecular complexity in ripening control. The gradient included genes associated with general ripening activities beyond plastid development and was consistent with the easily observed, yet poorly studied, visual ripening gradient present in tomato and many fleshy fruits. In green fruit, many genes that have differential expression along the latitudinal axis are under the influence of *SlGLK2*, however several transcription factors including *KNOX*, *MADS*, *NGATHA* (*ARF*-related) and *WRKY* families showed gradient expression throughout green and red stages. In red fruit, the sucrose hydrolysis gene invertase had gradient expression from the stylar end to the stem end and many ethylene-related genes were observed to be expressed more at the top (stem end) compared to the stylar end (bottom) of the fruit. Interestingly, no previously described ripening control genes such as *RIN*, *NOR*, *CNR*, *TAGL1* nor *SlAP2a* display such patterns, suggesting that other factors regulate the formation of the gradient during the ripening process in tomato. The *KNOX* gene has been reported by others to act upstream of *SlGLK2*. RNAi suppression of a *NGATHA* gene led to elongated and pointed fruit phenotypes consistent with the established *Arabidopsis* model of an auxin gradient along the apical basal axis of the gynoecium. Overexpression and suppression of a fruit top-bottom differentially expressed *WRKY* gene, *SlWRKY55*, did not yield any obvious

phenotype. Preliminary data from *SlWRKY51* overexpression showed that some fruit had elongated morphology. In summary, transcriptome and functional analysis suggest that there are gradients of gene expression along the vertical axis of tomato fruit that influence ripening polarity, chlorophyll distribution and lycopene accumulation. Several candidates were targeted for functional analyses and transgenic plants altered in expression of several of these genes have been created as a first step toward better understanding of the ripening gradient.

I. Introduction

Fleshy fruit are important components of human and animal diets with a central role in the reproductive biology of plants. Fruit biology research has shed light on basic fruit development and ripening in addition to avenues to improve fruit and produce quality characteristics. Many fruit develop from carpels (true fruit) or other floral-associated tissues (false or accessory fruit). Accordingly, *MADS* box genes have been shown to play important roles in carpel and fruit development. As examples, *AGAMOUS (AG)* clade genes *AG*, *SHATTERPROOF1* (*SHP1*) and *SHATTERPROOF2* (*SHP2*) from *Arabidopsis* and their putative orthologs TOMATO AGAMOUS (*TAG1*) and *TOMATO AGAMOUS-LIKE1* (*TAGL1*) in tomato are required for carpel identity and fruit development, dehiscence or ripening (Pnueli et al., 1994; Liljegren et al., 2000; Vrebalov et al., 2009). A-class *MADS*-box genes such as *FRUITFUL* (*FUL*) in *Arabidopsis* and *TDR4* in tomato act together with *SHP/TAGL1* to regulate fruit dehiscence and fruit ripening, respectively (Ferrandiz et al., 2000; Bemer et al., 2012). *SEPALLATA* genes *SHP1, SHP2, SHP3* in *Arabidopsis* show functional redundancy in normal petal, stamen, and carpel development, whereas the *RIN-MADS* gene in tomato plays an important role as a master regulator of fruit ripening (Vrebalov et al., 2002; Zhong et al., 2013).

Several additional classes of transcription factors also have defined functions in *Arabidopsis* and tomato carpel and fruit tissues, including *KNOX, BHLH, RAVB3, HB, HD-Zip, YABBY, SBP,* and *AP2/ERF* (reviewed in (Giovannoni, 2004; Ostergaard, 2009; Klee and Giovannoni, 2011), suggesting complex and layered regulation during carpel/fruit development. Screens for such regulators have by no means been saturated and additional players remain to be identified.

. Wild-type tomatoes display a "green shoulder" when unripe, often resulting in orange, green or white ripe regions at the top (stem end) of the ripe fruit. This uneven color phenotype is absent in the loss of function *uniform ripening (u)* mutant. *U* encodes a *Golden 2-like* transcription factor, *SlGLK2*, belonging to the *GARP* sub-family of the *myb* transcription factor superfamily in plants (see prior chapters and Powell et al., 2012; Nguyen et al., 2014). The normal *GLK2* gene is expressed in the fruit in a latitudinal gradient and with higher expression in the stem end (top) of the fruit conferring the green shoulder phenotype (Figure 2.5).

The green shoulder phenotype suggests a gradient of developmental regulation along the latitudinal (top – bottom) axis of the fruit. Such developmental gradients are described in both animal and plant systems. Examples include anterior-posterior axis patterning during *Drosophila* embryogenesis (Berleth et al., 1988; Lehmann and Nussleinvolhard, 1991), while in plants, the spatial pattern of auxin specifies apical-basal axis formation during embryogenesis (Friml et al., 2003). The development of the apical-basal axis of the *Arabidopsis* gynoecium has also been hypothesized to involve auxin gradients (Nemhauser et al., 2000; Balanza et al., 2006). Transcription factors also influence patterning of the *Arabidopsis* gynoecium (Ostergaard, 2009; Trigueros et al., 2009), though the impact of regulatory gene expression gradients has not been described in tomato or other fleshy fruits beyond the prior characterization of *GLK2*.

In this chapter I demonstrate transcriptome gradients throughout the latitudinal axis during tomato fruit development, consistent with the observed uneven gradient of ripening displayed by tomato and many fleshy ripening fruits, and revealing additional molecular complexity in ripening control. *The fact that the ripening gradient occurs in the absence of GLK2 activity (i.e., in the u/u uniform mutant) indicates that while GLK2 activity effects ripe fruit phenotype and quality characteristics, it operates separately from the bottom-top ripening gradient.* Preliminary data on functional analysis of gradient expressed genes confirm the existence of the gradient.

II. Results and discussion

1. The *GLK2* green shoulder phenotype results from a gradient of expression through the fruit

Expression of *SlGLK2* reflects a gradient of differential expression from the stem end to the blossom end of the fruit with expression varying 4- to 16-fold between the two ends (Figure 2.5, 4.1). Ectopic expression of either *GLK* gene minimized the repertoire of genes comprising this gradient and resulted in the immature fruit having a more uniform dark green color (Figure 2.6, 2.8), supporting the role of the *GLK2* expression gradient as cause of the green shoulder phenotype in wild-type fruit. Interestingly, tomato cultivars with elevated anthocyanin content also demonstrate a higher concentration of metabolites at the shoulder/stem end of the fruit (Gonzali et al., 2009), possibly reflecting along with the intensity of chlorophyll accumulation in light versius shaded tomato fruit, a role of light or other environmental factors. Moreover, the ripening process, at least as defined by pigment accumulation, usually does not occur uniformly throughout the fruit but initiates at the stylar end (bottom) of the fruit (Figure 4.1). These

phenotypes suggest a gradient of gene expression along the latitudinal axis of the fruit that may extend beyond *SlGLK2* and which has not been extensively characterized. Indeed, the fact that this ripening gradient occurs in the absence of functional SlGLK2 (*u/u* mutation; Figure 4.1) confirms that the *SlGLK2* expression and ripening gradients act largely independently.

To specifically explore this phenomenon of gradient ripening, RNA-seq was deployed to characterize the transcriptome activity of genes in three latitudinal sections along the vertical axis of wild type *'Ailsa Craig'* (*U/U*) fruit pericarp: the stem end or the top, the middle and the stylar end (bottom) of the fruit, and through six stages of fruit development (10 DPA, 20 DPA, Mature green, Breaker, Breaker+5 and Breaker+10 days). Fruit were cut into five equal sections and the intervening two sections discarded to minimize developmental and gene expression overlap (Figure 4.1C). I focused my analysis on two specific gene expression profile classes: high at the stem end, low at the stylar end or vice versa, and with 2 fold or greater ratio differences between the two ends and adjusted p value ≤ 0.05. The result was tens to hundreds of genes exhibiting differential expression along the latitudinal axis in all examined fruit stages (Figure 4.1C and 4.1D; Figure 4.2; see Supplemental Data in Nguyen et al., 2014).

At 10 DPA, many differentially expressed genes were related to photosynthesis and chloroplast structure, consistent with the high expression level of *GLK2* in the top of the fruit. At Breaker+5 and Breaker+10 (both red ripe fruit stages), several genes involved in ethylene biosynthesis and additional ripening pathways showed differential expression (high at the top and low at the bottom; Figure 4.3), including *ACO1*, *ACO3*, *ACS2*, *ACS4*, *E4*, and *E8*. Interestingly, at the breaker stage there is no significant differential expression of these genes between the sections. This may be due to the fact that all are regulated by the gaseous and readily diffusible hormone ethylene.

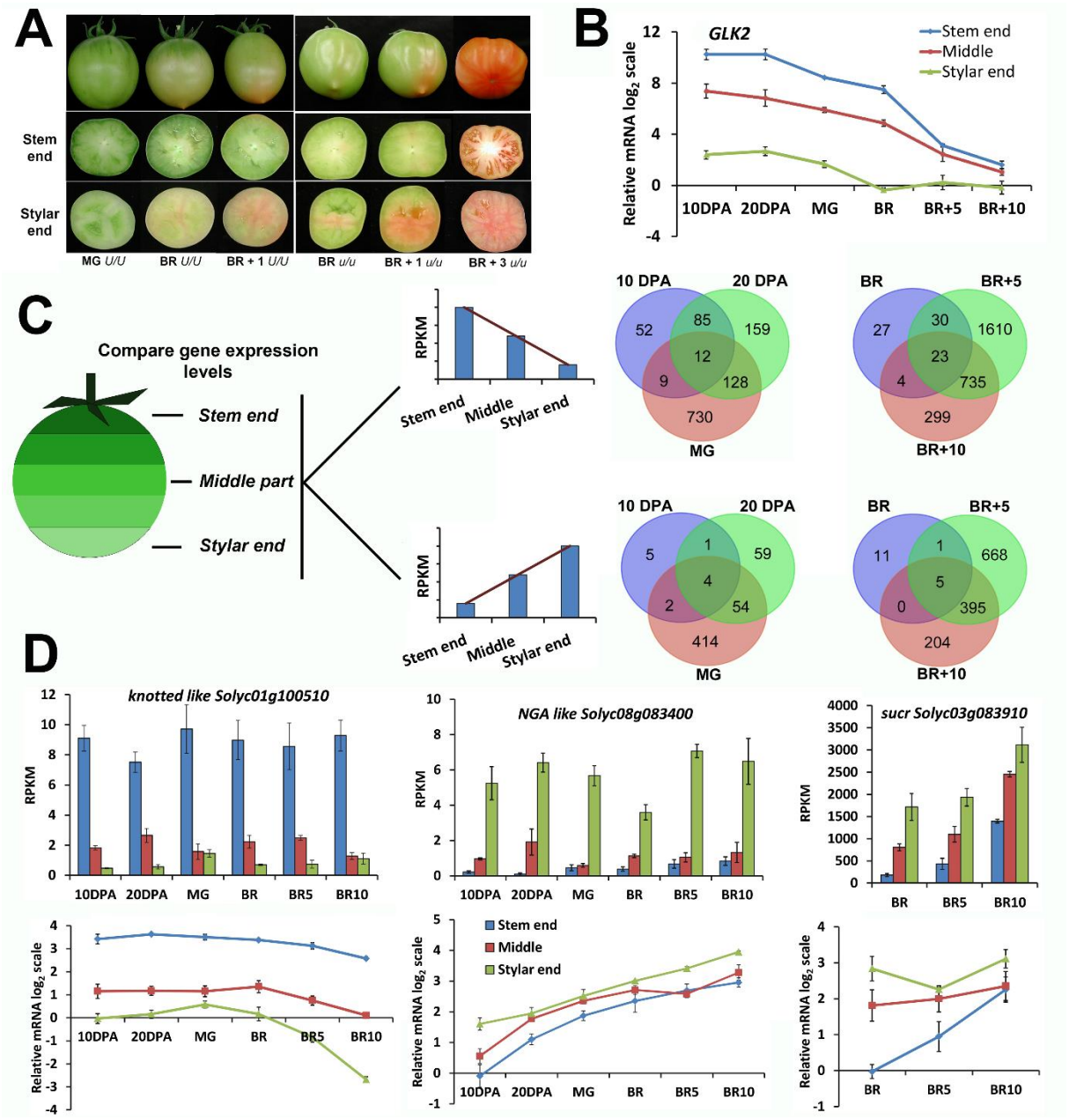


Figure 4.1. Gradient of gene expression along the latitudinal axis of tomato fruit

(A) Top row: wild-type *U/U* and mutant *u/u* tomato fruit at mature green (MG), Breaker (BR) and Breaker+1 day (BR+1) and Breaker+3 day (BR+3) stages. In BR and BR+1 color can be visually observed to "break" at the stylar end of the fruit. Second and third rows: Cross-sections .

(B) Gradient expression levels of *GLK2* as determined by qRT-PCR through three latitudinal sections of wild type *'Ailsa Craig'* (*U/U*) fruit during six stages of tomato fruit development. DPA=days post anthesis.

(C) Transcriptome analysis for differential gene expression through three latitudinal sections of wild type *'Ailsa Craig' (U/U)* fruit. Two expression profiles used to define differential expression: Expression levels (RPKM value) highest at stem end, lower at the middle part, and lowest at stylar end, or vice versa. Three green fruit stages (10 DPA, 20 DPA, Mature green) and three red fruit stages (Breaker, Breaker + 5, Breaker + 10) were represented in Venn diagrams to determine the shared differentially expressed genes among the different stages.

(D) Example of differential expression pattern of *Solyc01g100510* (a *knotted* transcription factor), *Soly08g083400* (*ARF/NGATHA* homolog gene) and *Solyc03g083910* (*sucr* - invertase gene) as determined by RNA seq (top) and confirmed by qRT-PCR (bottom).

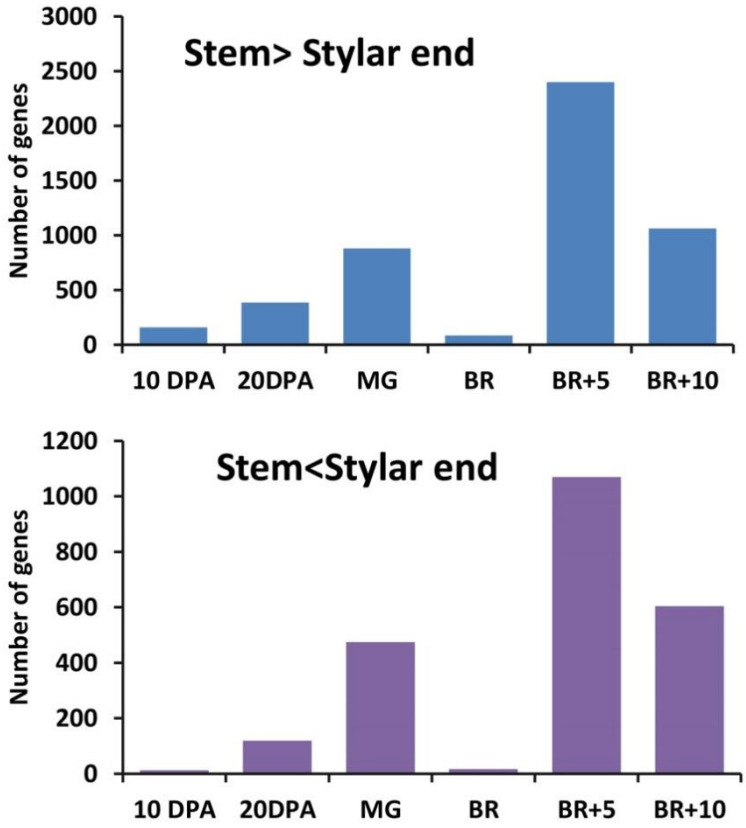


Figure 4.2. Numbers of genes showing differential expression along three latitudinal sections of tomato fruit tissues during fruit development.

Top: Genes with expression levels high at the stem end, low at the stylar end.

Bottom: Genes with expression levels low at the stem end, high at the stylar end.

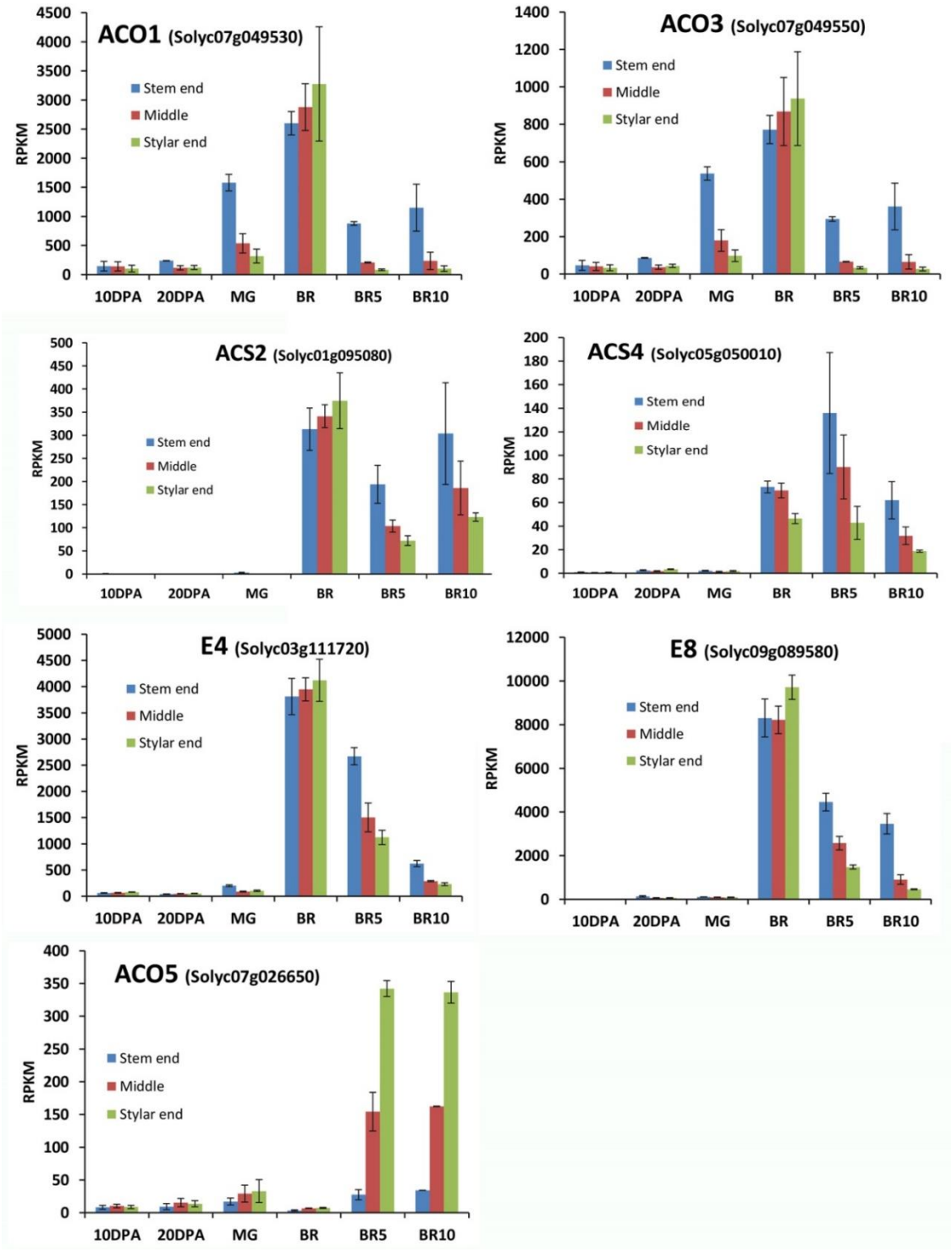


Figure 4.3. **Gradient expression of ethylene-related genes.** Expression patterns of ethylene biosynthetic and ethylene responsive genes that have high expression levels (RPKM >100) during fruit development and show differential expression along the 3 latitudinal sections of tomato fruit tissues as determined by RPKM value from RNA-seq.

Using Venn diagrams (Figure 4.1C), it is possible to describe genes that showed differential expression in all three green stages (Table 4.1), all red fruit stages (Table 4.2) and both green and red fruit stages (Table 4.3). To confirm the RNA-seq results, I also measured expression levels of selected genes by qRT-PCR that showed differential expression through multiple fruit stages and the results agree well with the RNA-seq data (Figure 4.1B and 4.1D). As examples, in green fruit tissues, in addition to *SlGLK2* showing differential expression in all green fruit stages as expected, there are also two genes of the light harvesting complex (*Solyc02g070990* and *Solyc02g070950*, both chlorophyll binding proteins) and one in chlorophyll biosynthesis (*Solyc10g006900, protochlorophyllide reductase*) that presumably reflect *SlGLK2* responses given the *u* phenotype. Indeed, most of the genes that show differential expression from the top to bottom of fruit in the green stages are also upregulated in response to *SlGLK2* overexpression (Table 4.1).

Interestingly, the macrocalyx *MC-MADS* gene and the tomato *AP1* ortholog functioning in floral determinacy and sepal development (Vrebalov et al., 2002), also showed differential expression in these green stages. Fruit of the *mc* mutant display the normal green shoulder as in the wild type and its repression yielded no discernible fruit phenotypes (Vrebalov et al., 2002) so its role in fruit development may be redundant with other of the many MADS-box genes expressed in fruit. In all three stages of red fruit, the *sucr* gene that hydrolyzes sucrose to glucose and fructose in tomato red fruit (Klann et al., 1996) showed high expression levels in the stylar end with low expression in the stem end consistent with the movement of the ripening gradient. It is interesting to note that known ripening regulators like *RIN*, *NOR*, *CNR*, *TAGL1*, *AP2a* in tomato do not display these gradient patterns, that is, there are no significant differences in expression levels between the sections during fruit development (Figure 4.4), suggesting that

other factors beyond these well described and necessary ripening regulators regulate the formation of the gradient during the ripening process in tomato.

Table 4.1. Genes showing differential expression along the 3 latitudinal sections of tomato fruit in all 3 green stages.

Gene ID[a]	Annotation
Solyc12g094640[c]	GAPDH B
Solyc03g045070[c]	Ammonium transporter
Solyc02g070990[c]	Chlorophyll a/b binding protein
Solyc02g070950[c]	Chlorophyll a/b binding protein
Solyc07g065900[c]	Fructose-bisphosphate aldolase
Solyc02g081120	Knotted-1-like homeobox protein
Solyc01g100510	Knotted-like homeobox protein
Solyc07g056540[c]	L-lactate dehydrogenase
Solyc05g056620	Macrocalyx MADS box
Solyc10g006900[c]	Protochlorophyllide reductase
Solyc10g008160[c]	SlGLK2
Solyc08g005010[c]	Unknown Protein

[a]: Genes with expression levels high at the stem end, low at the stylar end
[c]: Genes that are also upregulated in 35S:SlGLK2 OE fruit

Gene ID[b]	Annotation
Solyc07g053360	Seed biotin-containing protein
Solyc08g013690	ARF24 Transcriptional factor B3
Solyc08g083400	ARF24 Transcriptional factor B3
Solyc08g013700	ARF24 Transcriptional factor B3

[b]: Genes with expression levels low at the stem end, high at the stylar end
[c]: Genes that are also upregulated in 35S:SlGLK2 OE fruit

Table 4.2. Genes showing differential expression along the 3 latitudinal sections of tomato fruit in all 3 red stages

Gene ID[a]	Annotation
Solyc08g007460	Non-specific lipid-transfer protein
Solyc03g033840	26S protease regulatory Subunit 6B homolog
Solyc03g044900	Calmodulin-like protein 41 EF-Hand type
Solyc04g048900	Calreticulin 2 calcium-binding protein
Solyc10g017980	Chitinase Chitin-binding, type 1
Solyc07g009530	Chitinase Chitin-binding, type 1
Solyc10g017970	Chitinase Chitin-binding, type 1
Solyc04g078900	Cytochrome P450
Solyc10g051020	Cytochrome P450
Solyc07g006890	Cytochrome P450
Solyc02g089730	Endo-1 3-beta-glucanase
Solyc01g105370	Enoyl-(Acyl-carrier-protein) reductase II
Solyc04g051490	Essential meiotic endonuclease 1B
Solyc01g100510	Knotted-like homeobox protein
Solyc02g089630	Proline dehydrogenase
Solyc05g024260	RAG1-activating protein 1 homolog
Solyc09g082860	Sulfate adenylyltransferase ATP-sulfurylase
Solyc01g006950	Syntaxin
Solyc01g014840	Tir-nbs-lrr, resistance protein
Solyc09g083060	U-box domain-containing protein
Solyc01g086660	Unknown Protein (AHRD V1)
Solyc04g072070	WRKY transcription factor 16
Solyc04g051690	WRKY transcription factor 16
Gene ID[b]	Annotation
Solyc08g013700	ARF24 Transcriptional factor B3
Solyc08g083400	ARF24 Transcriptional factor B3
Solyc08g013690	ARF24 Transcriptional factor B3
Solyc01g099200	Lipoxygenase
Solyc03g083910	Acid beta-fructofuranosidase (invertase)

[a]: Genes with expression levels high at the stem end, low at the stylar end in all 3 red stages

[b]: Genes with expression levels low at the stem end, high at the stylar end in all 3 red stages

Table 4. 3. Genes showing differential expression along the 3 latitudinal sections of tomato fruit in all 6 stages

Gene	Annotation
Solyc01g100510	Knotted-like homeobox protein
Solyc08g013690	ARF24 Transcriptional factor B3
Solyc08g083400	ARF24 Transcriptional factor B3
Solyc08g013700	ARF24 Transcriptional factor B3

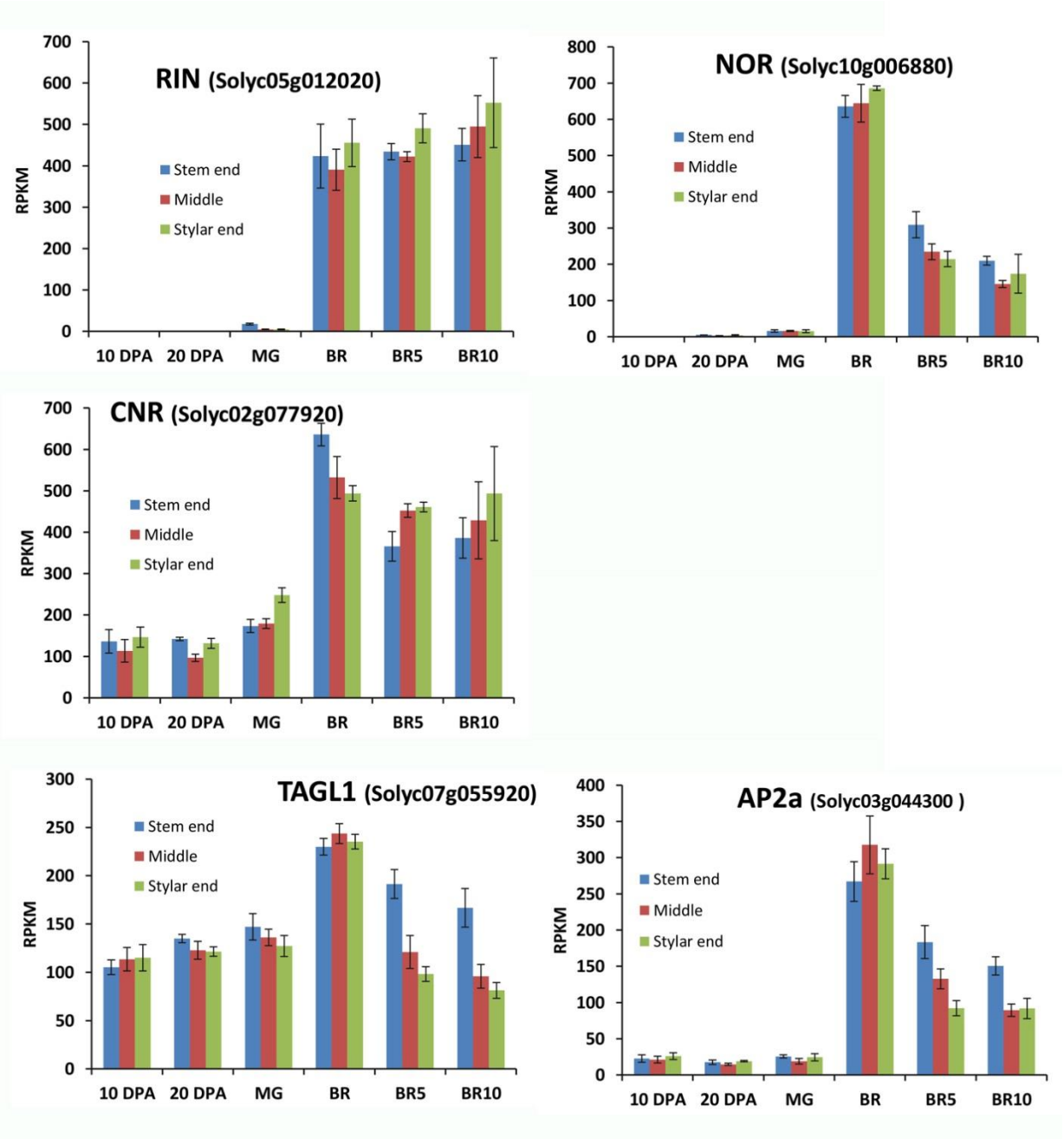


Figure 4.4 Expression levels of known ripening regulators during fruit development along the latitudinal axis of tomato fruit. Expression patterns of 5 known transcription factor that play roles in controlling tomato fruit ripening along the 3 latitudinal sections of tomato fruit tissues as determined by RPKM value from RNA-seq.

A *knotted* related transcription factor and three genes annotated as auxin responsive genes showed differential expression throughout the six stages of fruit development analyzed. These three auxin responsive genes are very similar to the *NGATHA* transcription factor family in *Arabidopsis* (Alvarez et al., 2009; Trigueros et al., 2009) that has been shown to affect gynoecium development. *knotted* transcription factors have been shown to be involved in leaf and shoot meristem development (Pautot et al., 2001; Ragni et al., 2008), but there is still no report on the effect of this class of transcription factor in fleshy fruit, making them interesting candidates for further functional characterization.

To more fully assess the differences in transcriptome activity resulting from the *GLK2* gradient versus the ripening gradient (and to ascertain the degree of any regulatory overlap), we analyzed transcriptome data from the top versus bottom of mature green fruit in addition to comparing the bottom of mature green to the bottom of breaker fruit. The former differ substantially in *GLK2* expression and are not ripe, so should represent GLK mediated differences, while the latter have very low and similar *GLK* expression and thus reflect differences largely due to the ripening process (Figure 4.5). Using high stringency cutoff values (>5-fold differential expression, p-value < 0.01), 371 genes were found to be differentially expressed between the top and bottom of the mature green fruit, while 1872 were differentially expressed between the mature green and breaker fruit bottoms (see Supplemental Data in Nguyen at al., 2014). Only 66 genes were found in common, suggesting that the ripening and GLK2 gradients are distinct (see Supplemental Data in Nguyen et al., 2014)

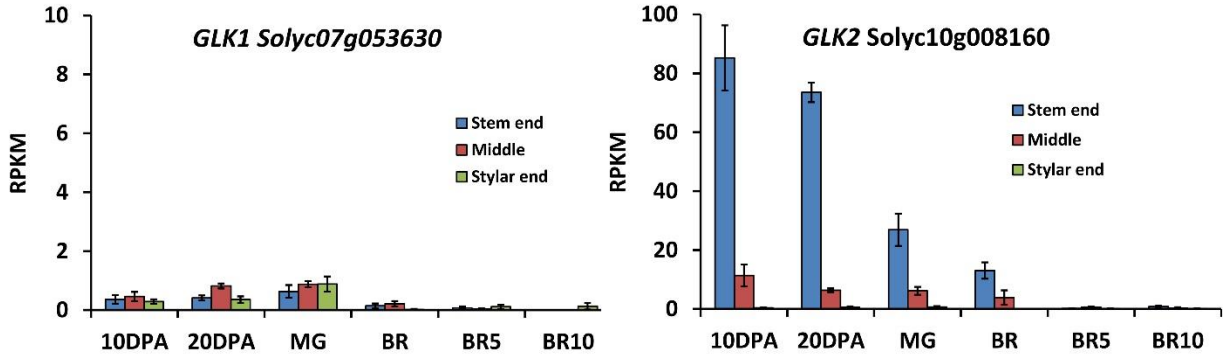


Figure 4.5. Expression levels of *GLK1* and *GLK2* during fruit development along the latitudinal axis of tomato fruit. Expression patterns of *GLK1* and *GLK2* along the 3 latitudinal sections of tomato fruit tissues as determined by RPKM value from RNA-seq.

GLK2 is a component of the latitudinal gradient of ripening

The mature tomato fruit is derived from the ovary/carpels of the gynoecium. In terms of tissue patterning, there are several axes of development and morphological variation in fruit: apical-basal, abaxial-adaxial, and medial-lateral (Ostergaard, 2009). Here we focused on the gradient of expression levels of genes in tomato fruit along the apical-basal or latitudinal axis that displays an obvious gradient of chlorophyll accumulation (defined by *U/SlGLK2*) and later ripening, that has not been characterized at the molecular level. Indeed, most reports on tomato (including our own) typically focus on whole fruit pericarp characterization, without distinguishing between the phenotypic and physiological variation across the latitudinal gradient. To distinguish the transcriptome effects of *SlGLK2* versus ripening gradients, and to establish whether genes in addition to *SlGLK2* present expression gradients through the fruit, we sectioned developing fruit and assayed transcriptome activity by RNA-seq in the top, middle and bottom fruit pericarp tissues. We showed that in each stage of development there are numerous genes that exhibit expression gradients through the fruit, and the number of differentially expressed genes generally increases with fruit development. Analogous studies have been performed on developing maize leaves, consistent with the gradient of blade maturation (Li et al., 2010). Unlike the maize blade that displays a developmental gradient reflecting the synthesis of new cells from meristematic stem cells and a gradient across the blade effectively representing cell age and maturity, the fruit ripening gradient occurs in the context of a tissue that has long ceased cell division. Interestingly, we found several transcription factors that maintained gradient expression patterns during one or more fruit stages. These include members of the *KNOX*, *WRKY* and *RAVB3/NGATHA* gene families (Tables 4.1 - 3). Homologous transcription factors have

been shown to play roles in gynoecium development in *Arabidopsis*. For example, the *knat1/brevipedicellus (bp)* mutants produce downward rather than upward oriented siliques (Venglat et al., 2002), whereas *KNAT2* and *KNAT6* (*Arabidopsis knotted*-like genes) are restricted by *KNAT1*/*BP* and *REPLUMLESS* (*RPL*) to promote correct inflorescence development (Pautot et al., 2001; Ragni et al., 2008). *KNAT2* over-expression in *Arabidopsis* induced homeotic conversion of the nucellus into a carpel-like structure. Here *Solyc01g100510*, a tomato homolog of *Arabidopsis KNAT2* and *KNAT6*, displayed high expression at the stem end of the fruit as compared to the basal end and throughout development. It is noteworthy that *knotted*-like genes are usually active in the shoot apical meristem (Hake et al., 2004). The fact that fruit are developed from carpels and the stem end of carpels originate from floral receptacles containing the shoot apical meristem, may explain the gradient expression of this *knotted*-like gene along the apical-basal axis. *NGATHA* gene family members also play roles in style and stigma development in *Arabidopsis* via a mechanism dependent on relative expression levels (Alvarez et al., 2009; Trigueros et al., 2009). The fact that the tomato homologs have higher expression levels at the stylar end derived from the tips of the carpels may indicate a similar gradient with roles during both gynoecium and later fruit maturation. It is noteworthy that previously described transcriptional regulators that are necessary for ripening, including *RIN*, *CNR*, and *TAGL1*, do not display gradients of expression, suggesting additional regulators are central to the ripening gradient.

The gradient expression of tomato *sucr* is also noteworthy. The stylar end is the location of the first visual signs of ripening initiation and would be anticipated to be the site of initial conversion of starch and sucrose to monosaccharides. Indeed, soluble solid (^{o}Brix) levels are significantly higher at the stylar end compared to the stem end of red ripe tomato fruit (Figure

4.6). Multiple ethylene biosynthesis and ethylene responsive genes displayed gradient expression at the Breaker + 5 and Breaker + 10 stages when the fruit are most ripe, consistent with the fact that the stem end or top of the fruit displays ripening phenotypes later than the bottom.

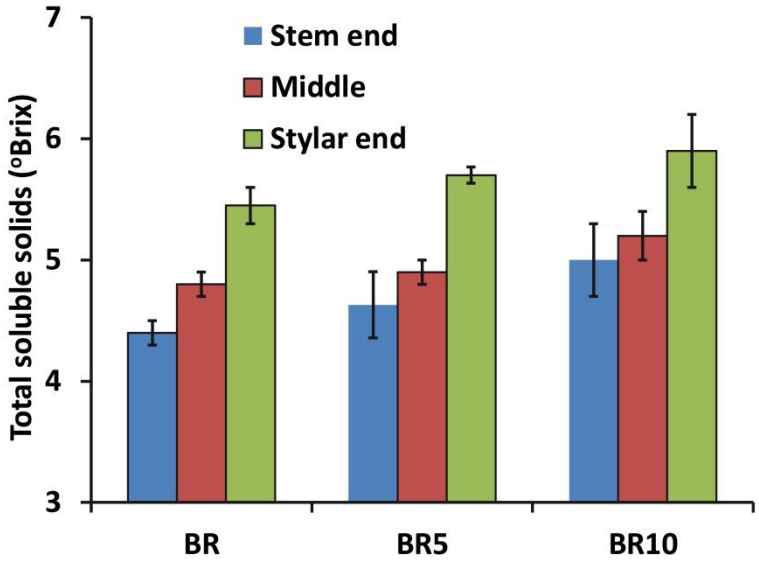


Figure 4.6. °Brix level along the latitudinal axis of tomato red fruit. Total soluble solids (°Brix) levels of 3 latitudinal fruit sections of wild type *'Ailsa Craig' U/U*. BR (Breaker stage) BR5 (Breaker+5 days) BR10 (Breaker+10 days).

Characterization of the two tomato *GLK* genes demonstrates the positive impact on fruit nutritional value stemming from more chlorophyll and chloroplasts in pre-ripe green fruit and suggests a means to improve the nutritional quality of tomato and additional fruit species. The data presented here revealed that while GLK is not limited in leaf tissue, it is limited in fruit, confirming the potential for using transgenes or allelic diversity to increase fruit quality by elevating GLK activity. Previously described ripening regulators, such as the *RIN* MADS-box

gene, did not show gradient expression though a number of putative regulators previously uncharacterized in tomato are candidates for gradient control.

2. Transgenic analysis of candidate genes showing gradient expression patterns

To assess function of genes displaying molecular gradients, transgenic plants were generated targeting several candidate genes that showed differential expression along the latitudinal axis during fruit development. Transcription factors were preferentially chosen as they are more likely to effect broader genetic pathways or processes and could act upstream (possibly of *SlGLK2*) in the molecular gradient(s) operating during fruit development.

KNOX gene TKN4 Solyc01g100510 (characterized recently by Nadakuduti et al., 2014)

One of the strongest candidate genes for further characterization is the *knotted*-like gene, *Solyc01g100510*. This gene maintained high differential expression levels throughout six stages of fruit development analyzed. It belongs to the homeodomain KNOX transcription factor family that plays an important role in shoot apical meristem function and morphogenesis, especially as related to leaf shape (Hay and Tsiantis, 2010).

Independent from this research, (Nadakuduti et al., 2014) recently performed positional cloning of the *ug* (*uniform globe*) mutation (Bohn and Scott, 1945), (which results in loss of the green shoulder phenotype similar to the *u* mutation phenotype) revealing that the gene is *Solyc01g100510,* which has been previously described as *TKN4* (Parnis et al., 1997). This result presents a link between *GLK2* and this particular *knotted*-like gene. Further support for the role of *knotted*-like or KNOX genes in fruit chloroplast development were drawn from characterization of the *Curl* (*Cu*) mutant (Parnis et al., 1997). This mutation is a dominant gain-of-function mutation resulting from ectopic expression of a KNOX gene, *TKN2*. Cu green fruit

display dark green color and elevated chlorophyll levels throughout the entire surface of the fruit. *SlGLK2* expression is upregulated in this mutant, suggesting *SlGLK2* is downstream of the KNOX genes in the regulatory hierarchy. This independent result confirms the transcriptome analysis by RNA-seq presented earlier in this chapter (Figure 4.1) and supports the hypothesis that genes displaying gradient expression patterns indeed have relevant biological consequences.

NGATHA genes

Three similar genes, *Solyc08g013690*, *Solyc08g083400*, *Solyc08g013700* maintain gradient expression patterns throughout the fruit developmental stages analyzed with all three low at the stem end and high at the stylar end (Figure 4.7). Homology analysis showed that these genes are homologs of *NGATHA* genes in *Arabidopsis* (Alvarez et al., 2009; Trigueros et al., 2009). *NGATHA* genes belong to the RAV clade of the B3 transcription factor family and are related to *AUXIN RESPONSE FACTOR (ARF)* transcription factor genes. Based on *ettin (ett)* mutants, altered in *AUXIN RESPONSE FACTOR 3,* Nemhauser et al. (2000) proposed a model for gradient distribution of auxin along the apical-basal axis of the gynoecium effecting the differentiation of style, ovary and gynophore tissues. According to this hypothesis, high auxin levels at the apical end direct the differentiation of style and stigma, intermediate levels of auxin promote the ovary and low auxin levels contribute to the development of the gynophore. Though the detailed mechanisms remain unclear, the model is consistent with a role of auxin through *NGATHA* genes in apical-basal patterning and similar expression in the maturing tomato fruit may reflect regulatory roles possibly including ripening effects.

Mutant and transgenic analyses of *NGATHA (NGA)* genes in *Arabidopsis* demonstrated that the 4 *NGA* genes act redundantly to promote style and stigma development. *nga* quadruple

mutants completely lack style and stigma and develop instead valve-like structures, whereas overexpression of *NGA* genes promotes ectopic style tissues in the flower (Alvarez et al., 2009; Trigueros et al., 2009).

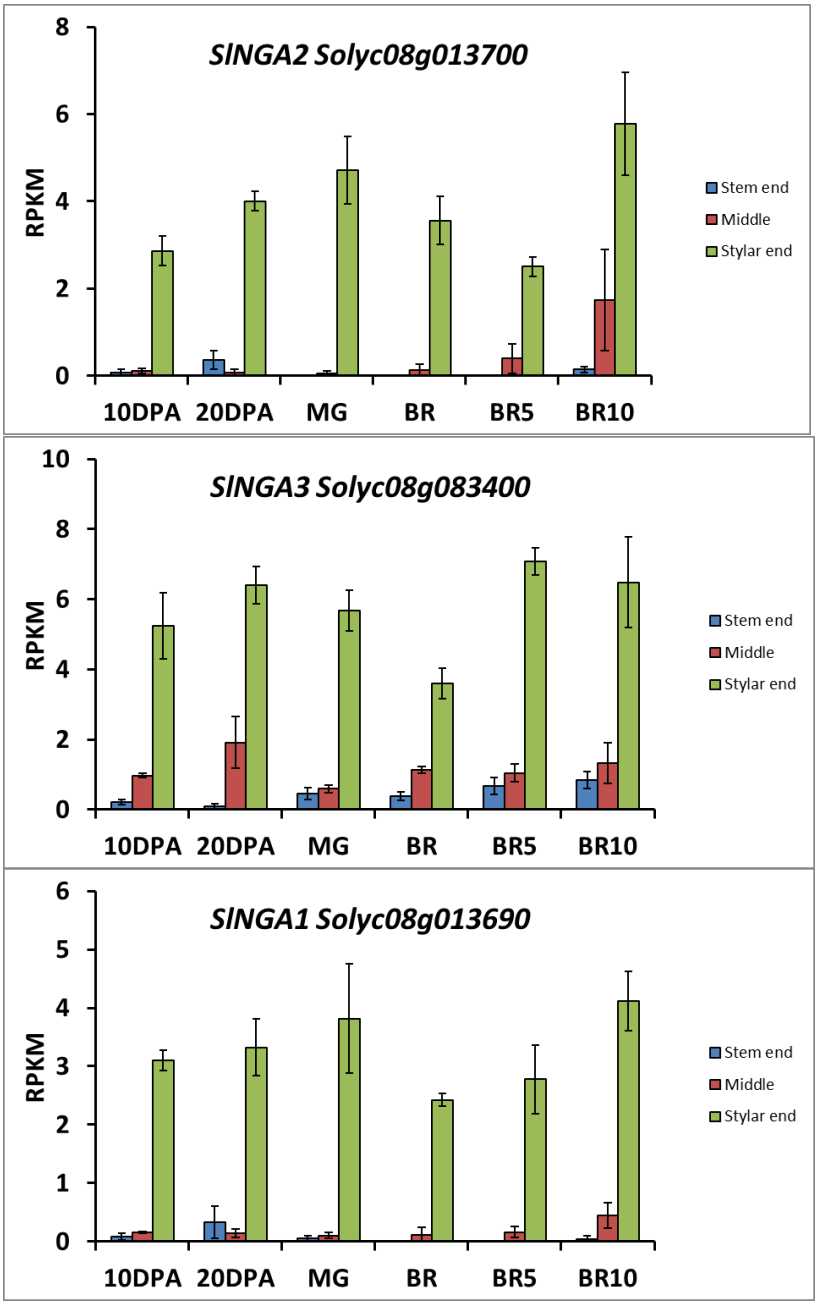


Figure 4.7. Expression levels of 3 *NGATHA* genes during fruit development along the latitudinal axis of tomato fruit as determined by RPKM value from RNA-seq.

NGATHA RNAi constructs for all three genes were created and transformed into wide type *'Ailsa Craig'* plants to create knockdown plants uniquely targeting each gene whose expression is described in Figure 4.7: *SlNGA1 (Solyc08g013690), SlNGA2 (Solyc08g083400), SlNGA3 (Solyc08g013700).*

In the T0 generation a total of 16, 15 and 10 *RNAi SlNGA1, SlNGA2* and *SlNGA3* plants, respectively, were recovered. Morphologically, *SlNGA2* and *SlNGA3* plants appeared normal and there were no notable visual differences as compared to wild type *'Ailsa Craig'*. For *SlNGA1* RNAi lines, 4 lines presented fruit shape phenotypes (Figure 4.9). In general, fruit were significantly smaller and more elongated compared to wild-type controls that did not carry the transgene. Some fruit had a pointed stylar end, similar to the heart-shaped fruit phenotype that occurs in fruit subject to application of auxin (Pandolfini et al., 2002) or down regulation of the auxin negative regulators *SlARF4* (Sagar et al., 2013) or SlARF7 (de Jong et al., 2009). Other fruit presented additional aberrant and bent shapes (Figure 4.8). It is important to note that the fruit phenotype was exaggerated in fruits that developed as the plants aged and were less severe in the first fruits. In some lines, the very first fruit set presented normal fruit phenotypes. In addition to abnormal shapes and reduced fruit size, said fruit were often parthenocarpic or produced non-viable seed. They typically also had thinner pericarp and less jelly in the locules compared to the wild type.

These fruit phenotypes were confirmed in the T1 generation plants although the phenotypes were generally less severe than in the T0 generation (Figure 4.8). Some flowers had exerted stigma (Figure 4.9) though the presentation of this phenotype was not as consistent as the fruit morphology alterations.

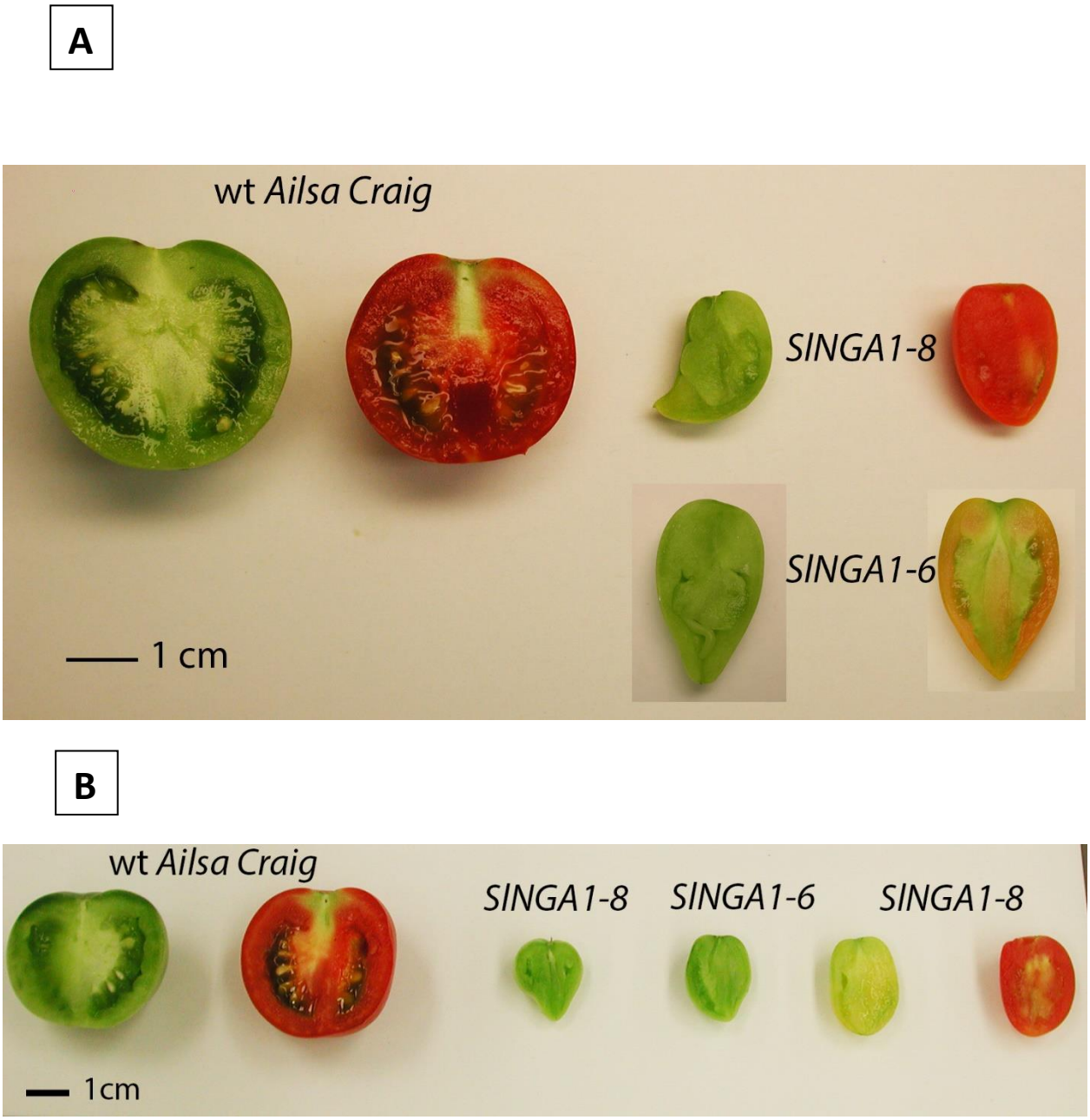


Figure 4.8. *SlNGA1* RNAi fruit phenotype. A. T0 generation fruit. B. T1 generation fruit. wild-type *'Ailsa Craig'* fruit at mature green and red stage.

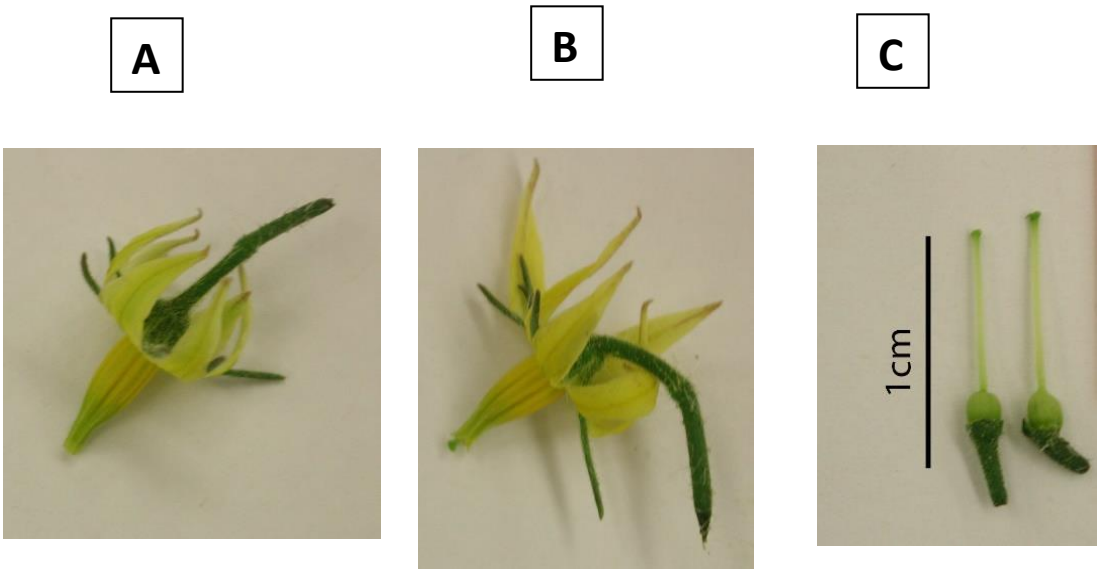


Figure 4.9. **Exerted stigma in control and *SlNga1* RNAi flowers**.

A: wild type *'Ailsa Craig'*. B: *SlNGA1* RNAi. C: Styles of wild-type (left) and SlNGA1 RNAi (right)

To test whether *SlNGA* genes have higher expression in style tissues analogous to *Arabidopsis*, RNA from style and ovary tissues in wild type fruit were extracted and subject to qPCR analysis. The results showed that *SlNGA1* indeed is expressed two-fold higher in styles compared to ovaries (Figure 4.10), suggesting conservation of function of *NGA* genes in style development between both species.

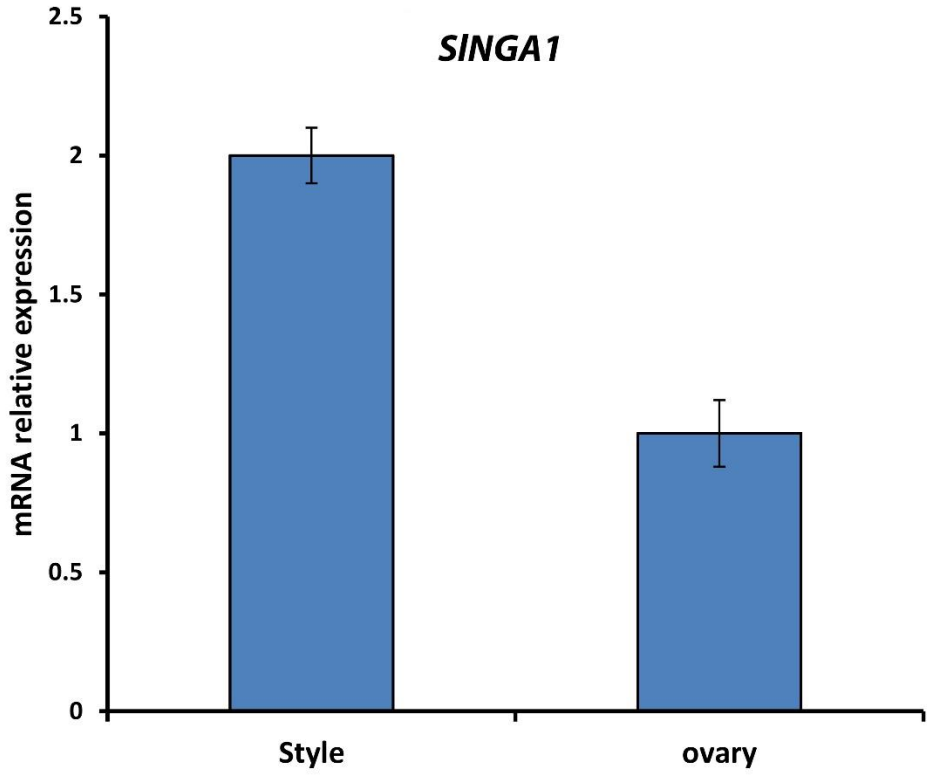


Figure 4.10. Relative *SlNGA1* expression determined by qPCR in tomato style and ovary.

Recently Martinez-Fernandez et al. (2014) performed transcriptome and transgenic analyses in *Arabidopsis* to show that *NGA* factors are essential for auxin synthesis in the apical gynoecium. In addition, *Arabidopsis nga* mutants affect auxin transport through PIN (auxin transporter) subcellular localization, thus disrupting auxin distribution in the gynoecium. The

fruit phenotypes resulting from *SlNGA1* repression in transgenic RNAi lines are reflective of and consistent with, auxin gradient distribution, though auxin levels remain to be quantitated in these lines. In contrast to the negative regulators of auxin such as *SlARF7* (de Jong et al., 2009) , *NGA* genes are positive auxin effect regulators, and thus low auxin levels have been associated with the *nga* phenotypes effecting style and stigma development in *Arabidopsis*. Auxin is known to promote cell expansion during fruit development (Gillaspy et al., 1993) and the small size and thin pericarp phenotypes in *SlNGA1 RNAi* fruit are thus consistent with low levels of auxin restricting cell and tissue expansion. Alterations in fruit shape likely reflect changes in auxin response mediated by altered transport resulting from perturbation of *NGA* expression. The fact that the phenotype appeared later during plant development, as opposed to the first fruits in lines suppressed by RNAi, suggests that *NGA* genes are essential but not sufficient in promoting auxin biosynthesis and transport, at least during the development of the first fruits. In these tissues there may be a requirement for additional factors or heightened accumulation of compensating fruit *SlNGAs*. It is noteworthy that in *Arabidopsis*, the strongest phenotypes were obtained from quadruple mutants (for *nga1, nga2, nga3, nga4*), indicating the potential for redundant functions in tomato as well and implying that the phenotype in *SlNGA1* repressed lines might be enhanced if additional *SlNGAs* genes were repressed. Expression analysis of the additional *SlNGA* family members has not yet been completed for early and late fruit, but could shed light on the stated hypothesis regarding differential family member expression in fruit set at different times in the lifespan of the transgenic *SlNGA1* repressed tomato plants.

WRKY genes

Two tomato WRKY genes *Solyc04g051690* and *Solyc04g072070,* exhibited higher expression at the stem end compared to the stylar end during most fruit stages (Figure 4.11).

WRKY transcription factors belong to a large plant-specific transcription factor family that have been shown to be involved in various plant processes, notably in resistance to both biotic and abiotic stresses (Ülker and Somssich, 2004; Rushton et al., 2010; Chen et al., 2012). In tomato, there is a paucity of data on WRKY that have been characterized functionally, and the knowledge regarding this gene family in tomato is limiting though the tomato genome sequence revealed 81 genes in the WRKY family (TGC, 2012). In order to assess possible functions of the *Solyc04g051690* and *Solyc04g072070* WRKY genes, repression (RNAi) and overexpression constructs were developed and delivered into transgenic tomato plants.

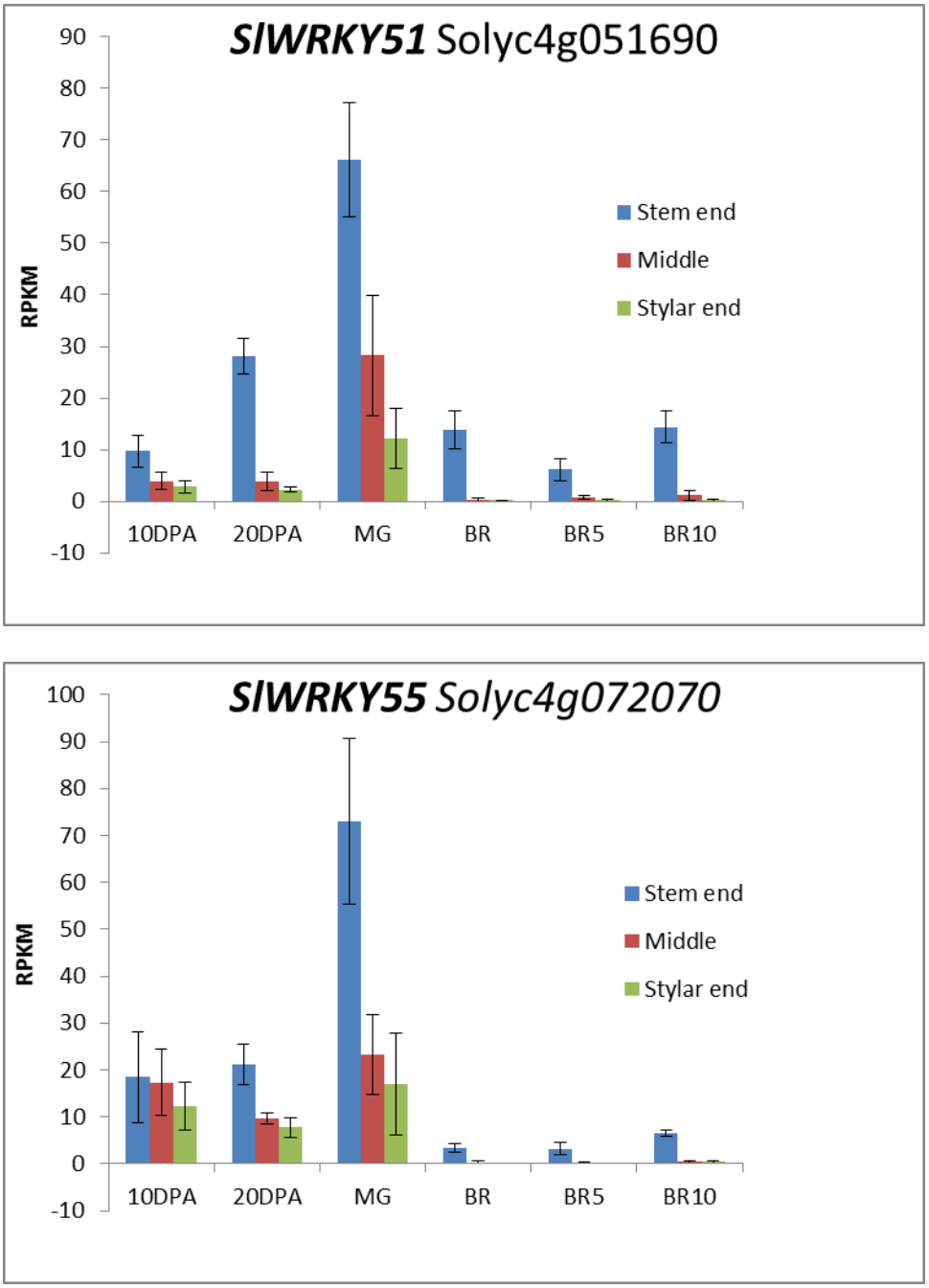


Figure 4.11. Expression levels of 2 WRKY genes during fruit development along the latitudinal axis of tomato fruit as determined by RPKM value via RNA-seq. *Solyc04g051690* and *Solyc04g072070* were named *SlWRKY51* and *SlWRKY55*, respectively.

Solyc04g051690 and *Solyc04g072070* are termed *SlWRKY51* and SlWRKY55, respectively. Phylogenetic analysis showed that both genes are closely related to each other and similar to *AtWRKY51* in *Arabidopsis* (Huang et al., 2012).

For *SlWRKY55*, preliminary data of RNAi and overexpression lines showed that there were no significant changes in the plant as a whole, nor in specific aspects of fruit morphology in the RNAi or over –expression (OE) OE-SlWRKY55 lines. Four SlWRKY55 RNAi lines with reduced transcript levels (compared to wild type *'Ailsa Craig'*) showed normal fruit morphology and ripening times. In the T1 generation, however, there were a number of T1 plants from one T0 line that showed pale leaf and fruit phenotypes (Figure 4.12). The segregation ratio of T1 progeny from this T0 line was 3:1 normal: pale leaf. This phenotype segregated in the progeny of a single T0 plant while T1 offspring of ~15 other T0 plants did not display signs of chlorophyll deficiency, suggesting the phenotype is independent of *SlWRKY55* gene repression, possibly reflecting insertion of the transgene into a chlorophyll or plastid-related locus. Another possible explanation would be a spontaneous mutation during tissue culture or propagation. Identification of the transgene insertion site by sequencing or inverse PCR would help explain this particular transgenic event.

For *SlWRKY51*, there were no *RNAi* plants recovered from two independent transformation experiments. This occurrence might suggest that *SlWRKY51* is an essential gene in early plant development or regeneration and repression is lethal. Among four overexpression plants generated, one plant showed some fruit with an elongated phenotype, including several with dark green color (Figure 4.13). While further investigation is needed (including additional lines with the same phenotype) the fruit phenotype suggests the possibility of disrupted auxin

signaling. Several studies in rice and *Arabidopsis* also showed overexpression of some WRKY genes can affect auxin response in transgenic plants (Zhang et al., 2008; Song et al., 2010).

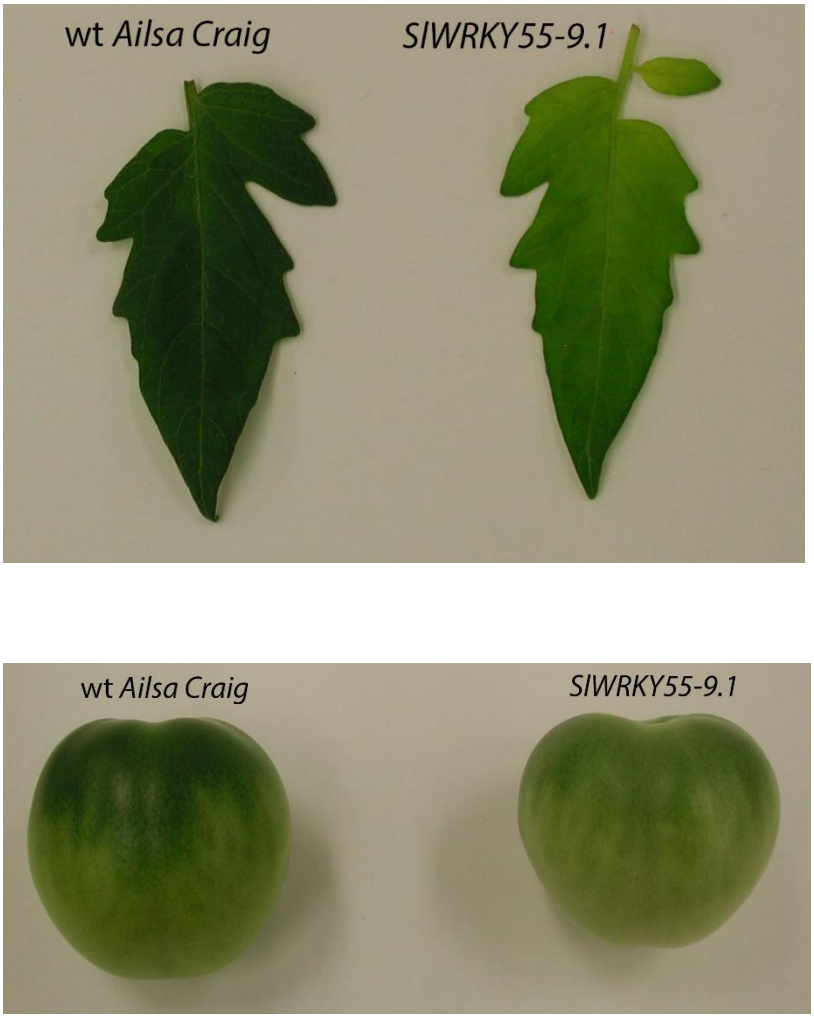


Figure 4.12. Pale leaf and fruit phenotype of a T1 *SlWRKY55 -9* RNAi transgenic.

Figure 4.13. Fruit from a *SlWRKY51* overexpression line.

Left: wild type immature *'Ailsa Craig'*. Right: Elongated *OE SlWRKY51* RNAi

Table 4.4. DNA primer sequences.

	Forward Primer	Reverse Primer
qRT-GLK1	5'-TCATCATCTTCATCCTCCACCTC -3'	5'-TGGTGCTGGAAAGTAAGGAGTTC-3'
qRT-GLK2	5'-GACACAGTTGCCTCTGATTTGC-3'	5'-TGAGGAGGTTTGAGTCCTATGG-3'
GLK1 probe	5'-TTCGCCGTCTCCACAACAT-3'	5'-AATGGCTGCATCTATGCTCTCC-3'
GLK2 probe	5'-ATGCTTGCTCTATCTTCATCATTGA -3'	5'-TTGAAGATGACTAGCAATGTTATGTCT -3'
35S	5'-AAACCTCCTCGGATTCCATTG-3'	5'-AGGAAGGGTCTTGCGAAGGA-3'
qRT-08g083400	5'-AGTACCCAGCTGCTGCTTGG -3'	5'-CTGAGACATGCTGCCAGGTG -3'
qRT-01g100510	5'-TCCGCTGCTTCCATCTCCT -3'	5'-AGCTCAGTTCCCTATGCAAAG -3'
qRT-Sucr	5'-AGGTGGTACCATCCCTTGTGA-3'	5'-TATGCGTGGTCCAATGGCTATG-3'
qRT-18S	5'-CGGAGAGGGAGCCTGAGAA-3'	5'-CCCGTGTTAGGATTGGGTAATTT-3'
HP1 marker	5'-ATCATAATGGGCTTCTTCACCTACT-3'	5'-GACCTTGCCTTTCCAGATCAAC-3'

Summary

Overall, functional analyses of gradient expression genes confirm the existence of a molecular gradient system along the vertical axis in tomato fruit pericarp. In the case of the KNOX gene *TKN4*, RNA-seq revealed a candidate gene acting upstream of *GLK2* as validated by Nadakuduti et al., 2014. This work serves as a platform for further spatial and temporal studies in tomato and other fruit species. Several additional candidates have been identified and preliminary functional studies have been performed, though additional work is needed to validate their functions possibly through assessment of activities of potentially functionally redundant genes or through tissue-specific expression that allows recovery of viable transgenic plants.

III. Materials and Methods

Plant materials

Tomato cultivars *'Ailsa Craig'* was provided by the Tomato Genetics Resource Center, Davis, CA. Plants were grown in greenhouses at the Boyce Thompson Institute for Plant Research, Ithaca, NY with a 16 h light (27^{o}C) and 8 h dark (19^{o}C) cycle. Tomato fruit were tagged at 7-8 days post anthesis (dpa) when they were 1 cm in diameter and reached breaker stage at 37 to 38 dpa.

RNAi constructs

RNAi constructs were made using the pHELLSGATE 2 vector (kindly provided by CSIRO, Plant Industry, Canberra, Australia). cDNA sequences used in the hairpin included both sequences from the C domain and the subsequent 3′ UTR. The target sequences were PCR amplified from total immature fruit cDNA using gene-specific primers with addition of the corresponding recombination sequences as defined in the kit for site-specific recombination used (Gateway BP Clonase enzyme mix; Invitrogen). The resulting PCR product was gel purified and cloned into pHELLSGATE 2 via homologous recombination using the kit above. The resulting construct was sequence confirmed and transformed into tomato cv *'Ailsa Craig'* by *Agrobacterium tumefaciens* (strain ABA4404)

Expression constructs and transformation

The full-length sequences of *SlWRKY51* and *SlWRKY55* cDNA were cloned into pBTEX (Lin et al., 2006) vectors digested with *SmaI/SalI*, and are under expression control via the CaMV35S promoter. The fidelity of the construct was confirmed by DNA sequencing and transgenic *'Ailsa Craig' (U/U)* tomato plants were generated by *Agrobacterium tumefaciens* (strain LBA4404) mediated transformation using previously described methods (Fillatti et al., 1987). Plants with

transgene insertions were screened by PCR and DNA gel-blot using CAMV 35S-specific primers/probe. *TAGL1* RNAi lines were described previously (Vrebalov et al., 2009).

Quantitative RT-PCR analysis

Quantitative real-time PCR was performed using the ABI PRISM 7900HT Sequence Detection System and the Power SYBR Green RNA-to-C_T 1-Step Kit (Applied Biosystems) following the manufacturer's instructions. Quantitation of gene expression was calculated using the relative standard curve method (Applied Biosystems). Gene-specific primers used in the qRT-PCR assay are listed in Table 2.3. Standard errors (SEs) were calculated based on a minimum of three biological replicates.

Transcriptome sequencing

Fruit were harvested at the stages indicated in the text from three biological replicates, each replicate contained 4-5 combined fruit of the same developmental stage and from the same plant. For gradient expression, fruit were equally cross-sectioned in five parts and pericarp tissues from the top (stem end), middle and bottom (stylar end) parts were kept for RNA extraction. The two intervening sections were discarded. Pericarp tissues (without locule, septa and columella tissues) were frozen in liquid nitrogen. Total RNA from frozen pericarp tissues was used to construct strand-specific RNA libraries as described (Zhong et al., 2011). Two or three biological replicates for each sample were sequenced using a HiSeq2000 Sequencing System (Illumina, CA) according to the manufacturer's instructions.

RNA seq data processing

RNA-Seq reads were first processed by removing barcode and adaptor sequences. The resulting reads were aligned to ribosomal RNA (rRNA) and transfer RNA (tRNA) sequences using Bowtie allowing up to three mismatches (Langmead et al., 2009) to remove potential contaminating reads. The resulting cleaned reads were then aligned to the tomato genome using Tophat allowing one segment mismatch (Trapnell et al., 2009). Following alignments, raw counts of mapped reads for each tomato gene model were derived and then normalized to reads per kilobase of exon model per million mapped reads (RPKM). Differentially expressed genes between transgenic and control tissue were identified using the DESeq package (Anders and Huber, 2010). To identify differentially expressed genes across the fruit gradient, the raw counts of RNA-seq expression data were first transformed using the getVarianceStabilizedData function in the DESeq package. The variance-stabilizing transformed expression data were then fed to the LIMMA package (Smyth, 2004), and F tests were performed. Raw p values were adjusted for multiple testing using the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995). Dr. Zhangjun Fei helped me analyze the RNA-seq data.

REFERENCES

Alvarez, J.P., Goldshmidt, A., Efroni, I., Bowman, J.L., and Eshed, Y. (2009). The NGATHA distal organ development genes are essential for style specification in Arabidopsis. Plant Cell 21: 1373-1393.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol 11: R106.

Balanza, V., Navarrete, M., Trigueros, M., and Ferrandiz, C. (2006). Patterning the female side of Arabidopsis: the importance of hormones. Journal of Experimental Botany 57: 3457-3469.

Bemer, M., Karlova, R., Ballester, A.R., Tikunov, Y.M., Bovy, A.G., Wolters-Arts, M., Rossetto Pde, B., Angenent, G.C., and de Maagd, R.A. (2012). The tomato FRUITFULL homologs TDR4/FUL1 and MBP7/FUL2 regulate ethylene-independent aspects of fruit ripening. Plant Cell 24: 4437-4451.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society. Series B (Methodological) 57: 289-300.

Berleth, T., Burri, M., Thoma, G., Bopp, D., Richstein, S., Frigerio, G., Noll, M., and Nusslein-Volhard, C. (1988). The role of localization of bicoid RNA in organizing the anterior pattern of the Drosophila embryo. EMBO Journal 7: 1749-1756.

Bohn, G., and Scott, D. (1945). A second gene for uniform unripe fruit color in the tomato. Journal of Heredity 36: 169-172.

Chen, L., Song, Y., Li, S., Zhang, L., Zou, C., and Yu, D. (2012). The role of WRKY transcription factors in plant abiotic stresses. Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms 1819: 120-128.

de Jong, M., Wolters-Arts, M., Feron, R., Mariani, C., and Vriezen, W.H. (2009). The Solanum lycopersicum auxin response factor 7 (SlARF7) regulates auxin signaling during tomato fruit set and development. Plant Journal 57: 160-170.

Ferrandiz, C., Liljegren, S.J., and Yanofsky, M.F. (2000). Negative regulation of the SHATTERPROOF genes by FRUITFULL during Arabidopsis fruit development. Science 289: 436-438.

Fillatti, J.J., Kiser, J., Rose, R., and Comai, L. (1987). Efficient transfer of a glyphosate tolerance gene into tomato using a binary *Agrobacterium tumefaciens* vector. Nature Biotechnology 5: 726-730.

Friml, J., Vieten, A., Sauer, M., Weijers, D., Schwarz, H., Hamann, T., Offringa, R., and Jurgens, G. (2003). Efflux-dependent auxin gradients establish the apical-basal axis of Arabidopsis. Nature 426: 147-153.

Gillaspy, G., Ben-David, H., and Gruissem, W. (1993). Fruits: A developmental perspective. Plant Cell 5: 1439-1451.

Giovannoni, J.J. (2004). Genetic regulation of fruit development and ripening. Plant Cell 16: S170-180.

Gonzali, S., Mazzucato, A., and Perata, P. (2009). Purple as a tomato: towards high anthocyanin tomatoes. Trends in Plant Science 14: 237-241.

Hake, S., Smith, H.M., Holtan, H., Magnani, E., Mele, G., and Ramirez, J. (2004). The role of knox genes in plant development. Annual Review of Cell and Developmental Biology 20: 125-151.

Hay, A., and Tsiantis, M. (2010). KNOX genes: versatile regulators of plant development and diversity. Development 137: 3153-3165.

Huang, S.X., Gao, Y.F., Liu, J.K., Peng, X.L., Niu, X.L., Fei, Z.J., Cao, S.Q., and Liu, Y.S. (2012). Genome-wide analysis of WRKY transcription factors in Solanum lycopersicum. Molecular Genetics and Genomics 287: 495-513.

Klann, E.M., Hall, B., and Bennett, A.B. (1996). Antisense acid invertase (TIV1) gene alters soluble sugar composition and size in transgenic tomato fruit. Plant Physiology 112: 1321-1330.

Klee, H.J., and Giovannoni, J.J. (2011). Genetics and control of tomato fruit ripening and quality attributes. Annual Review of Genetics 45: 41-59.

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

Lehmann, R., and Nussleinvolhard, C. (1991). The Maternal Gene Nanos Has a Central Role in Posterior Pattern-Formation of the Drosophila Embryo. Development 112: 679-691.

Li, P., Ponnala, L., Gandotra, N., Wang, L., Si, Y., Tausta, S.L., Kebrom, T.H., Provart, N., Patel, R., Myers, C.R., Reidel, E.J., Turgeon, R., Liu, P., Sun, Q., Nelson, T., and Brutnell, T.P. (2010). The developmental dynamics of the maize leaf transcriptome. Nature Genetics 42: 1060-1067.

Liljegren, S.J., Ditta, G.S., Eshed, Y., Savidge, B., Bowman, J.L., and Yanofsky, M.F. (2000). SHATTERPROOF MADS-box genes control seed dispersal in Arabidopsis. Nature 404: 766-770.

Martinez-Fernandez, I., Sanchis, S., Marini, N., Balanza, V., Ballester, P., Navarrete-Gomez, M., Oliveira, A.C., Colombo, L., and Ferrandiz, C. (2014). The effect of NGATHA altered activity on auxin signaling pathways within the Arabidopsis gynoecium. Frontiers in Plant Science 5.

Nadakuduti, S.S., Holdsworth, W.L., Klein, C.L., and Barry, C.S. (2014). KNOX genes influence a gradient of fruit chloroplast development through regulation of GOLDEN2-LIKE expression in tomato. Plant Journal 78: 1022-1033.

Nemhauser, J.L., Feldman, L.J., and Zambryski, P.C. (2000). Auxin and ETTIN in Arabidopsis gynoecium morphogenesis. Development 127: 3877-3888.

Nguyen, C.V., Vrebalov, J.T., Gapper, N.E., Zheng, Y., Zhong, S., Fei, Z., and Giovannoni, J.J. (2014). Tomato GOLDEN2-LIKE transcription factors reveal molecular gradients that function during fruit development and ripening. Plant Cell 26: 585-601.

Ostergaard, L. (2009). Don't 'leaf' now. The making of a fruit. Current Opinion in Plant Biology 12: 36-41.

Pandolfini, T., Rotino, G.L., Camerini, S., Defez, R., and Spena, A. (2002). Optimisation of transgene action at the post-transcriptional level: high quality parthenocarpic fruits in industrial tomatoes. BMC Biotechnology 2: 1.

Parnis, A., Cohen, O., Gutfinger, T., Hareven, D., Zamir, D., and Lifschitz, E. (1997). The dominant developmental mutants of tomato, Mouse-ear and Curl, are associated with distinct modes of abnormal transcriptional regulation of a Knotted gene. Plant Cell 9: 2143-2158.

Pautot, V., Dockx, J., Hamant, O., Kronenberger, J., Grandjean, O., Jublot, D., and Traas, J. (2001). KNAT2: evidence for a link between knotted-like genes and carpel development. Plant Cell 13: 1719-1734.

Pnueli, L., Hareven, D., Rounsley, S.D., Yanofsky, M.F., and Lifschitz, E. (1994). Isolation of the tomato AGAMOUS gene TAG1 and analysis of its homeotic role in transgenic plants. Plant Cell 6: 163-173.

Powell, A.L., Nguyen, C.V., Hill, T., Cheng, K.L., Figueroa-Balderas, R., Aktas, H., Ashrafi, H., Pons, C., Fernandez-Munoz, R., Vicente, A., Lopez-Baltazar, J., Barry, C.S., Liu, Y., Chetelat, R., Granell, A., Van Deynze, A., Giovannoni, J.J., and Bennett, A.B. (2012). Uniform ripening encodes a Golden 2-like transcription factor regulating tomato fruit chloroplast development. Science 336: 1711-1715.

Ragni, L., Belles-Boix, E., Gunl, M., and Pautot, V. (2008). Interaction of KNAT6 and KNAT2 with BREVIPEDICELLUS and PENNYWISE in Arabidopsis inflorescences. Plant Cell 20: 888-900.

Rushton, P.J., Somssich, I.E., Ringler, P., and Shen, Q.J. (2010). WRKY transcription factors. Trends in Plant Science 15: 247-258.

Sagar, M., Chervin, C., Mila, I., Hao, Y., Roustan, J.P., Benichou, M., Gibon, Y., Biais, B., Maury, P., Latche, A., Pech, J.C., Bouzayen, M., and Zouine, M. (2013). SlARF4, an auxin response factor involved in the control of sugar metabolism during tomato fruit development. Plant Physiology 161: 1362-1374.

Smyth, G.K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Statistical Applications in Genetics and Molecular Biology 3: Article3.

Song, Y., Chen, L.G., Zhang, L.P., and Yu, D.Q. (2010). Overexpression of OsWRKY72 gene interferes in the abscisic acid signal and auxin transport pathway of Arabidopsis. Journal of Biosciences 35: 459-471.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25: 1105-1111.

Trigueros, M., Navarrete-Gomez, M., Sato, S., Christensen, S.K., Pelaz, S., Weigel, D., Yanofsky, M.F., and Ferrandiz, C. (2009). The NGATHA genes direct style development in the Arabidopsis gynoecium. Plant Cell 21: 1394-1409.

Ülker, B., and Somssich, I.E. (2004). WRKY transcription factors: from DNA binding towards biological function. Current Opinion in Plant Biology 7: 491-498.

Venglat, S.P., Dumonceaux, T., Rozwadowski, K., Parnell, L., Babic, V., Keller, W., Martienssen, R., Selvaraj, G., and Datla, R. (2002). The homeobox gene BREVIPEDICELLUS is a key regulator of inflorescence architecture in Arabidopsis. Proceedings of the National Academy of Sciences 99: 4730-4735.

Vrebalov, J., Ruezinsky, D., Padmanabhan, V., White, R., Medrano, D., Drake, R., Schuch, W., and Giovannoni, J. (2002). A MADS-box gene necessary for fruit ripening at the tomato ripening-inhibitor (rin) locus. Science 296: 343-346.

Vrebalov, J., Pan, I.L., Arroyo, A.J., McQuinn, R., Chung, M., Poole, M., Rose, J., Seymour, G., Grandillo, S., Giovannoni, J., and Irish, V.F. (2009). Fleshy fruit expansion and ripening are regulated by the Tomato SHATTERPROOF gene TAGL1. Plant Cell 21: 3041-3062.

Zhang, J., Peng, Y.L., and Guo, Z.J. (2008). Constitutive expression of pathogen-inducible OsWRKY31 enhances disease resistance and affects root growth and auxin response in transgenic rice plants. Cell Research 18: 508-521.

Zhong, S., Joung, J.G., Zheng, Y., Chen, Y.R., Liu, B., Shao, Y., Xiang, J.Z., Fei, Z., and Giovannoni, J.J. (2011). High-throughput illumina strand-specific RNA sequencing library preparation. Cold Spring Harbor Protocols 2011: 940-949.

Zhong, S., Fei, Z., Chen, Y.R., Zheng, Y., Huang, M., Vrebalov, J., McQuinn, R., Gapper, N., Liu, B., Xiang, J., Shao, Y., and Giovannoni, J.J. (2013). Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening. Nature Biotechnology 31: 154-159.

CHAPTER V

TRANSCRIPTOME CHARACTERIZATION OF LOCULAR GEL AND PERICARP TISSUES DURING TOMATO FRUIT DEVELOPMENT

Abstract

The fruit is a complex plant organ housing different structures and cell types as well as representing a developmental program composed of multiple physiological and biochemical processes. Fruit ripening is a unique process to plants characterized by changes to multiple characteristics including texture, color, aroma and nutrient quality. The process is especially interesting in that the organ undergoes a change entirely counter to the normal defensive manifestations in dropping its chemical and physical defenses to become attractive to seed dispersing organisms. Although considerable knowledge regarding molecular mechanisms of fruit development and ripening has been acquired in recent years, information concerning the spatial and temporal differences in different tissues of the fruit is lacking. As shown in the previous chapter, within the pericarp tissue there exist molecular gradients along the latitudinal axis operating during fruit development. In tomato, the locular tissue becomes jelly-like prior to ripening suggesting that early ripening phenomena may occur in this tissue, possibly providing key ripening signals to the pericarp. In this chapter, transcriptomes from locular tissue are compared to the pericarp tissue through a time course defined by multiple stages. The transcriptome analysis presented herein indicates that ethylene biosynthetic genes in general are expressed higher in the pericarp compared to the locule. However, the increase of expression of ethylene responsive genes such as E4 and E8, as well as master ripening regulators including

RIN and NOR, occurs earlier in locular tissue compared to the pericarp, suggesting that ripening phenomena initiate in the locule. Pericarp and locular tissues also display differences in expression levels of several cell wall-related genes including Polygalacturonase (PG) and Expansin1 (LeExp1). Finally, this analysis revealed tissue-specific transcription factors from the *STK*, *NTT* and *YABBY* gene families that may prove to be interesting targets for functional analyses as well as sources for tissue-specific promoters.

I. Introduction

Fruit expansion and further development typically occurs after fertilization of the ovary. Developing fruit undergo enlargement through cell division and cell expansion, followed by the later ripening process to facilitate the release of seeds. In some fruits such as tomato, these processes occur in clear sequential order, while in others, such as grape, continued fruit expansion and ripening processes can occur simultaneously. In fleshy fruit like tomato, the ovary wall develops into the mature pericarp that can be separated into three layers: endocarp, mesocarp and exocarp (Gillaspy et al., 1993; Seymour et al., 2013). Matas et al. (2011) classified the tomato pericarp more precisely as 5 major tissues: outer epidermis, collenchyma, parenchyma, vascular tissue and inner epidermis. Inside the fruit, the septa function as the barrier to divide the fruit into locules and the parenchymous placenta is the location of seed attachment to, and is the origin of, locular tissue that fills the locular cavity and becomes jellied during later fruit maturation and ripening.

In dehiscent fruits like the well-studied plant *Arabidopsis,* the fruit, or siliques, achieve seed dispersal by a process of dehiscence or pod shatter where lignification of valve margin cells

adjacent to an abscission cell layer (dehiscent zone) leads to separation of the valves, liberating the enclosed seed (Liljegren et al., 2000). Interestingly, some of the same regulatory genes regulate the seed release process in both Arabidopsis and tomato, albeit in the context of very different fruit structures and developmental processes (Vrebalov et al., 2009).

Classic physiological studies define fruit species into two groups with different ripening characteristics: climacteric fruit are defined by a peak in respiration and a concomitant burst of ethylene hormone biosynthesis during fruit ripening, whereas non-climacteric fruit lack these changes (Alexander and Grierson, 2002; Giovannoni, 2004; Klee and Giovannoni, 2011; Gapper et al., 2013). Ethylene is required for climacteric fruit to ripen normally, as disruptions in either ethylene production and/or ethylene signaling transduction inhibit fruit ripening (Wilkinson et al., 1995). Some non-climacteric fruit have minimal if any response to exogenous ethylene (e.g. strawberry) while others (e.g. some melons, peppers and citrus) display ripening responses upon ethylene exposure, suggesting they are either more sensitive to ethylene or ethylene synthesis or response impaired (Giovannoni, 2004). In tomato, a climacteric fruit, a number of important mutants including *rin*, *nor*, and *Cnr* do not ripen in response to exogenous ethylene, although they do control ethylene biosynthesis and perception (Giovannoni, 2007), suggesting that both ethylene-dependent and ethylene-independent gene regulation pathways coexist to coordinate the ripening process. Furthermore, they suggest that there are likely shared regulators of fruit ripening in climacteric and non-climacteric fruit.

Much published work on analysis of ripening molecular biology, physiology and biochemistry has looked at whole or easily separated fruit tissues, for example, whole fruit pericarp of tomato. This approach has a limitation in that it neglects different tissue types as well as potential spatial differences in genetic control (as shown in the previous chapter). Matas et al.

(2011), using laser capture micro dissection (LCM) to isolate different tissue types in tomato pericarp, were able to show many genes differentially expressed in different tissues, validating the importance of understanding tissue and spatial distinctions during organ development.

The locular tissue is of interest as it shows some of the earliest signs of change prior to or at the onset of ripening. The locular tissue differentiates from the placenta and gradually assumes a jelly-like consistency during fruit maturation and ripening. This tissue is highly parenchymous and surrounds the seed, likely to protect and assist seed development and maturation. Prior evidence suggests that climacteric ethylene production may first occur in locular tissue prior to the pericarp (Brecht, 1987; Atta-Aly et al., 2000) and thus ripening may initiate from this tissue. Using RNA sequencing, the transcriptomes of tomato locular (minus seed) tissue were profiled during different stages of fruit development and we compared said expression data with the transcriptome activity of corresponding pericarp from the same fruits. The goal was to identify ripening marker and novel genes that are differentially or specifically transcribed in each tissue toward better understanding fruit development and ripening, and especially the initiation of this critical plant process which is so important to human food and nutrient security.

II. Results and discussion

In contrast to the molecular gradients along the latitudinal axis where differentially expressed genes occurs in the same tissue (pericarp), the locular gel and the pericarp are distinct tissues, with the locular gel originating from the placenta and the pericarp from the ovary walls. The locular tissue is highly parenchymous and gradually solubilizes during fruit maturation. In

immature fruit, this tissue contains chlorophyll and is green in color while during ripening it generally changes color to white, orange or red, much like the pericarp (Figure 5.1).

To specifically identify unique transcriptomes in the locular gel and pericarp, RNA sequencing was deployed to characterize gene expression activities in each tissue in wild type (cv. *'Ailsa Craig'*) through five stages of fruit development (20 DPA, 25 DPA, Mature green, Breaker and Breaker +8 days). Using criteria of 5-fold or greater ratio difference between the tissues and adjusted p value ≤ 0.05, hundreds to thousands of genes exhibiting differential expression in all examined fruit stages were identified (Table 5.1). The first general observation was that there are more genes that are expressed more highly in pericarp than in locular tissue.

Figure 5.1. Cross section of tomato fruit. Cross-section of *'Ailsa Craig'* showing locule tissue in mature green and red fruit

Table 5.1 Numbers of genes showing differential expression between locular and pericarp tissues at each fruit stage.

Greater than 5 fold (Locule>pericarp)	
BR 8	459
BR	505
MG	391
25 DPA	387
20 DPA	609
Greater than 5 fold (Pericarp>Locule)	
BR8	618
BR	1253
MG	1050
25 DPA	1335
20 DPA	1160

In characterizing these differences more specifically, it was noted a number of genes showed differential expression in all five stages (Table 5.2-3). There are 37 genes that are expressed 5 fold or higher in locule compared to pericarp in all 5 stages and 109 genes that expressed 5 fold or higher in pericarp compared to locule in all 5 stages. Comparing expression patterns of ethylene biosynthesis genes between the two tissues indicated that these genes are expressed more highly in the pericarp as compared to the locule tissue (Figure 5.2). The committed steps of the ethylene synthesis pathway consists of two types of enzymes: ACC synthase (ACS) that converts S-adenosylmethionine to 1-aminocyclopropane-1-carboxylate (ACC) (Klee and Giovannoni, 2011); ACC is subsequently converted to ethylene by ACC oxidase (ACO), the so-called ethylene forming enzyme. Studies have shown that *ACS* transcription is a major point of regulation for ethylene synthesis though in tomato both ACS and ACO genes are upregulated during ripening (Alexander and Grierson, 2002). *ACS2* and *ACS4*, whose transcription rates are significantly increased at the onset of ripening, showed significantly higher expression in pericarp compared to the locular tissue. *ACS1A* is almost exclusively expressed in the pericarp. ACO genes like *ACO1*, *ACO3*, *ACO4 and ACO5* all showed higher expression in the pericarp, especially at the Breaker stage.

Ethylene regulated genes including E4 and E8, however, showed very high expression levels in the locular tissue and seemed to increase their level of expression earlier than in the pericarp (Figure 5.3). For example, at the Mature green stage, the expression level of E8 is ~5 fold higher in the locule compared to the pericarp. At Breaker stage, E4 is expressed higher in the locule. NR and ETR4, the two dominant ethylene receptors that are expressed in fruit, have similar expression level between pericarp and locule, suggesting that level of these ethylene receptors are not the reason for the differences. ETR6 however has higher expression in the

locule compared to pericarp at the Breaker stage. One potential explanation is that since E4 and E8 are strongly regulated and very sensitive to ethylene, ethylene synthesis may be increased first in the locular tissue at the onset of ripening (Brecht, 1987), though the molecular data for ethylene synthesis genes presented here is not consistent with this hypothesis. Also, the possibility remains that the stages examined here bracket (e.g. MG and Breaker) the point where initial ethylene synthesis may occur in the locule. Another explanation is that additional factors besides ethylene regulate these genes. Indeed, it is known that ripening transcription factors and the epigenome may play such roles (Zhong et al., 2013).

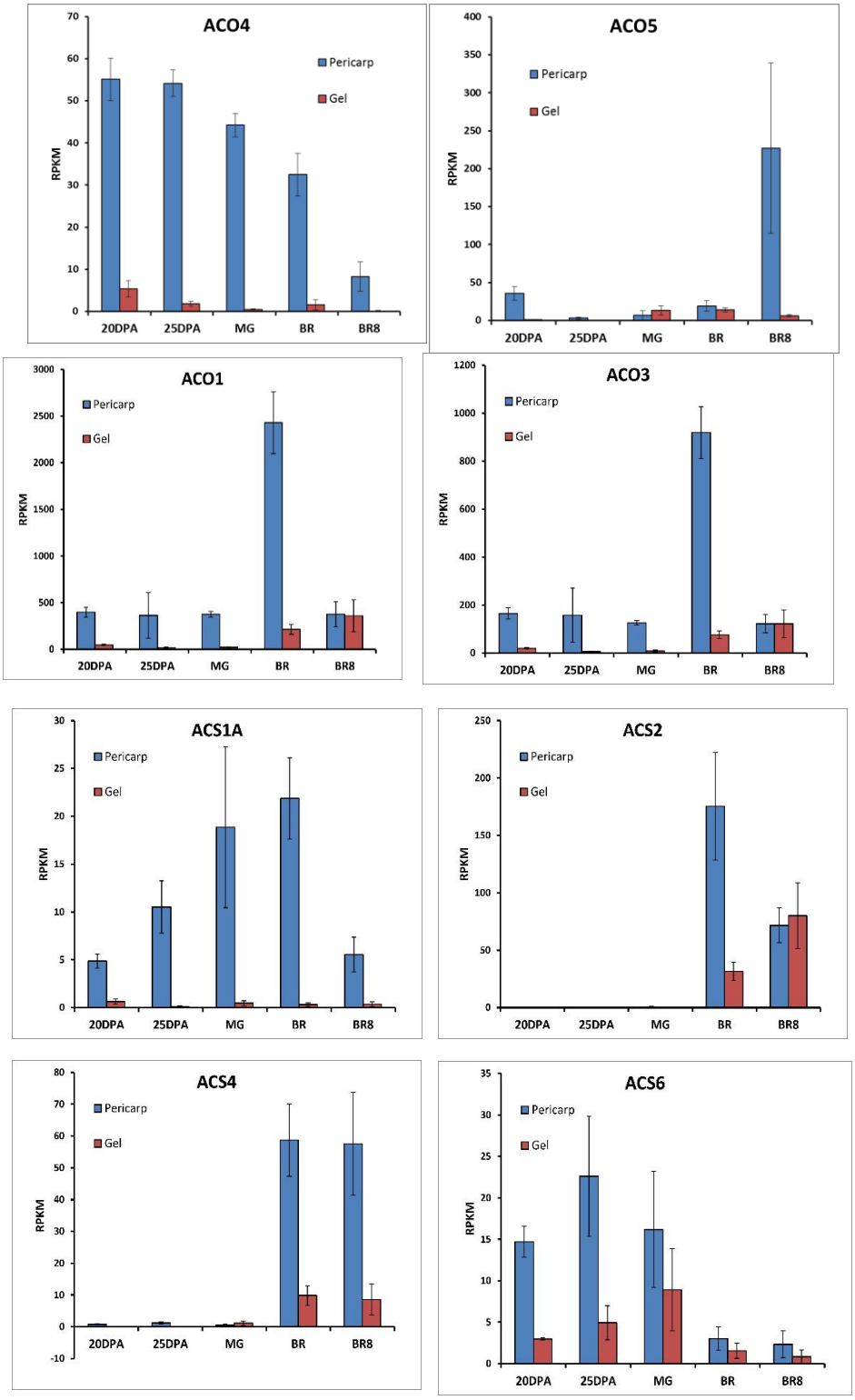


Figure 5.2. Expression levels of ripening-associated ethylene biosynthetic genes in pericarp and locular gel tissues during fruit development as determined by RPKM value from RNA-seq.

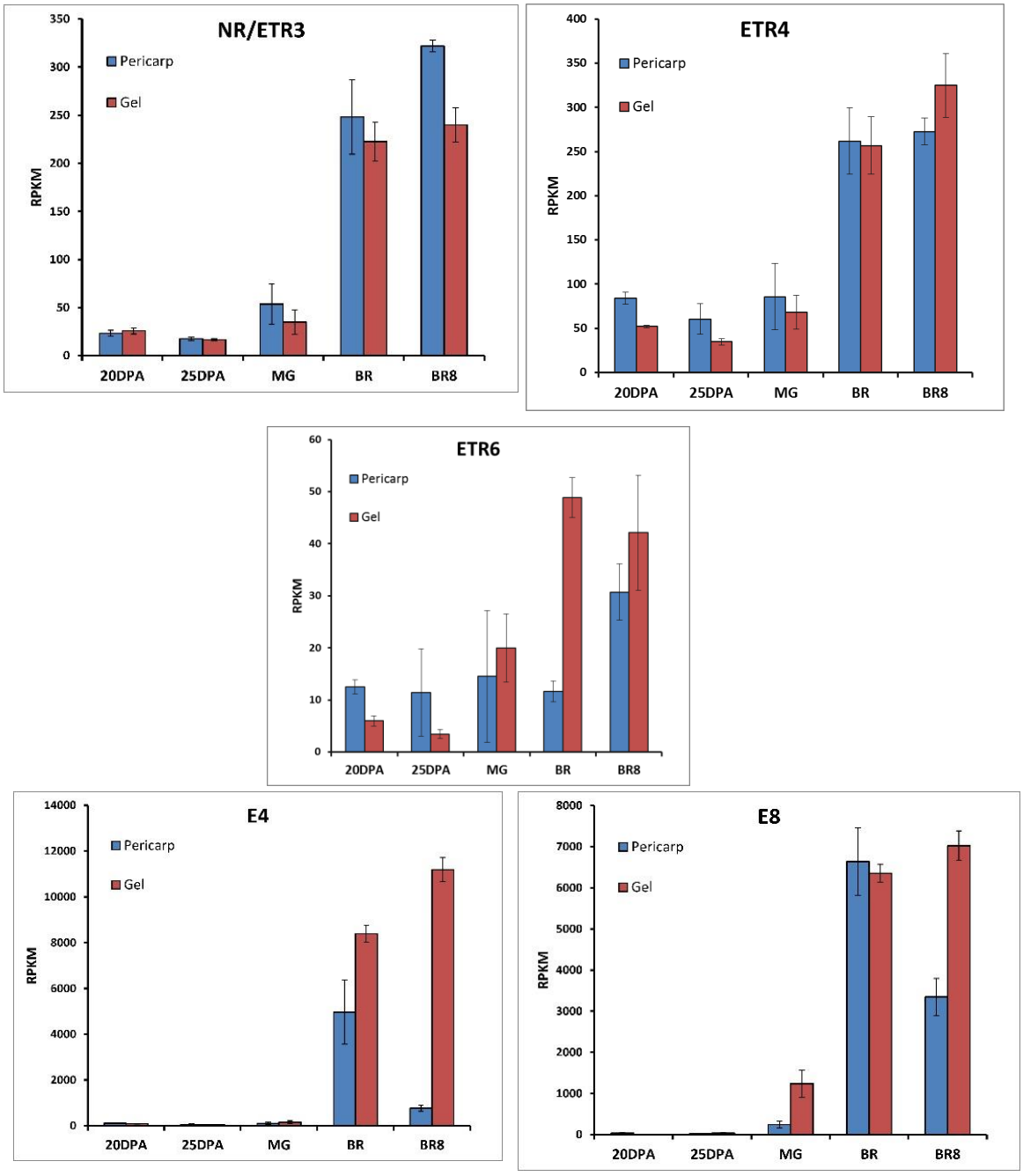
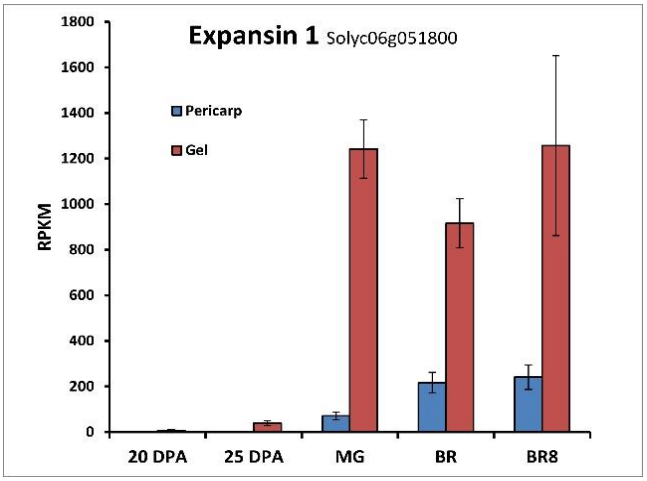
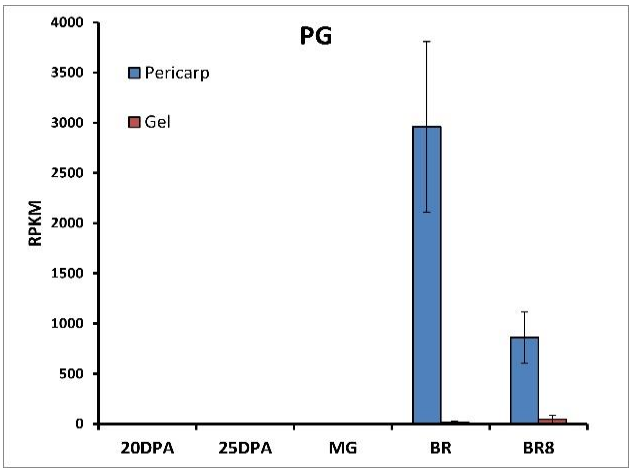


Figure 5.3. Expression levels of ethylene receptors (*NR*, *ETR4* and *ETR6*) and ethylene regulated genes (*E4* and *E8*) in pericarp and locular gel tissues during fruit development as determined by RPKM value from RNA-seq.

Two important fruit cell wall related proteins showed significant tissue expression differences. *Polygalacturonase* (*PG*) (Smith et al., 1988) and some pectinesterases like *Solyc07g064170* are predominantly expressed in pericarp tissue, whereas *Expansin1* (*LeExp1*) (Rose et al., 1997) is expressed in both tissues but is significantly higher in the locule (Figure 5.4). The low activity of PG in locular tissue has been reported previously (Cheng and Huber, 1996). *PG* is a pectinase that degrades polygalacturonan by hydrolysis of the glycosidic bonds that link galacturonic acid residues whereas *Expansin1* is a nonenzymatic protein that functions in cell wall loosening. The structure and function of each tissue may explain these expression differences. The locular tissue is very parenchymous, it needs to be enlarged to fill the locules during the liquefaction process occurring at ripening. Cheng and Huber (1996) showed that the pectic polymers of this tissue exhibit minimal depolymerization and contain unusually high quantities of xylose. The locule environment is also acidic. Overall, structures and environmental conditions in the locule favor expansin activity that is known to prefer acidic conditions (Sampedro and Cosgrove, 2005), likely maintained in part by auxin synthesis from the seeds.

NOR and *RIN*, two important ripening transcriptional regulators showed interesting expression pattern between these two tissues. The expression levels of *NOR* and *RIN* are both increased at the onset of ripening from Mature green to Breaker stage. In case of *NOR*, expression increased in the locular tissue first, and maintained higher expression levels compared to the pericarp until Mature Green (Figure 5.5). In locular tissue, expression levels of *RIN* also increased first at the Mature green stage. The master ripening regulation functions of NOR and RIN showing earlier accumulation in pre-ripening locule tissues may be a sign of early ripening events preceding ethylene and possibly promoting its synthesis and perception.





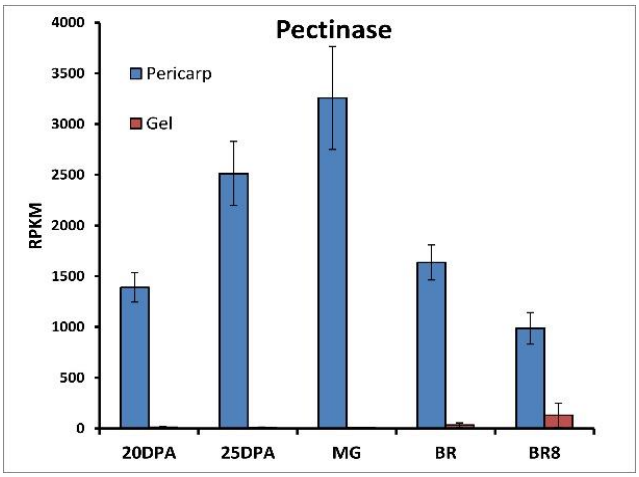


Figure 5.4. Expression levels of cell wall loosening genes in pericarp and locular gel tissues during fruit development as determined by RPKM value from RNA-seq.

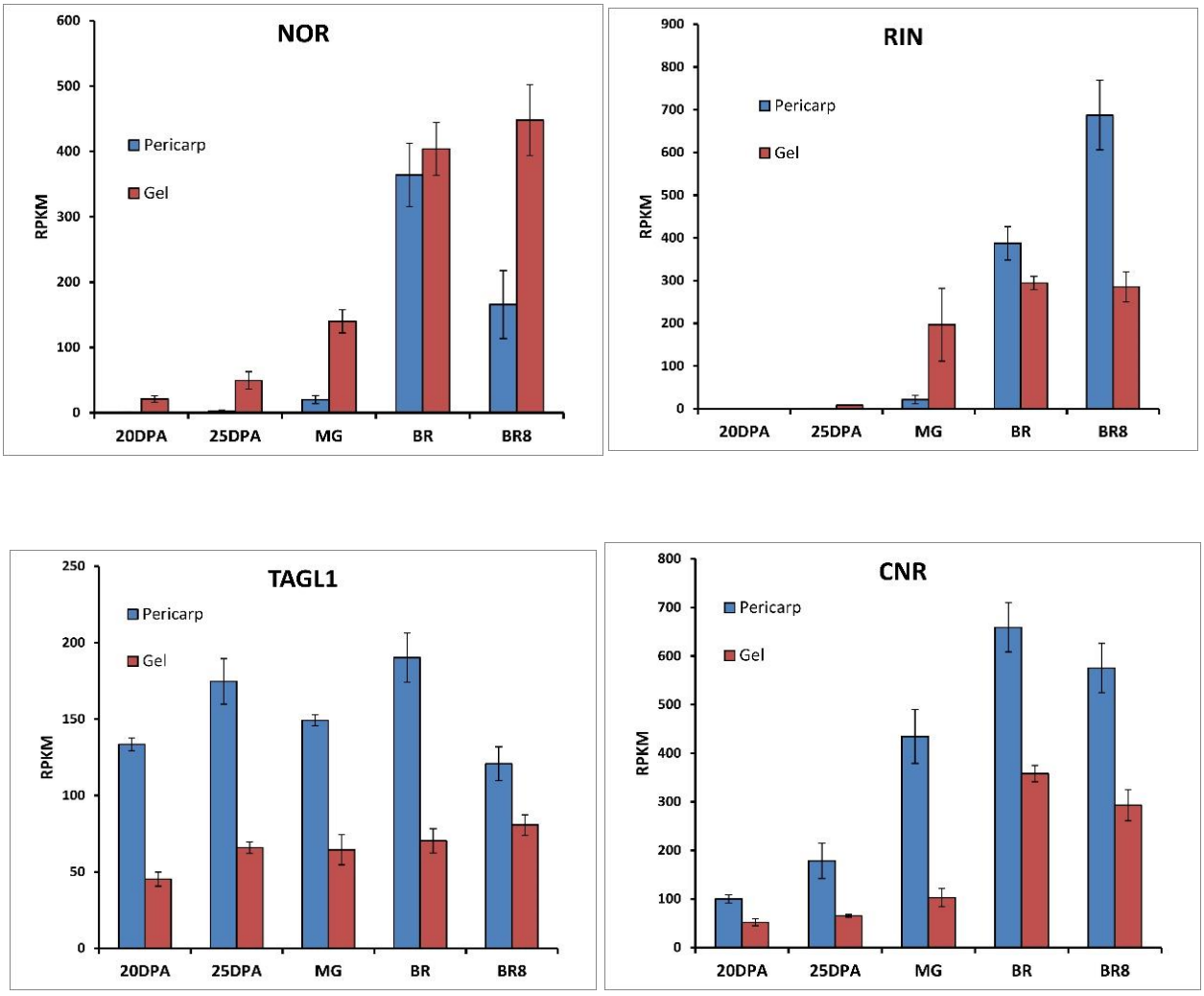


Figure 5.5. Expression levels of known ripening regulators during fruit development as determined by RPKM value from RNA-seq.

Some additional interesting tissue specific transcription factors were noted during this analysis. An AGAMOUS like MADS-box transcription factor, *Solyc06g064840*, is highly and specifically expressed in the locular tissue in all fruit stages. Homology analysis showed that this gene is the homolog of *SEEDSTICK* (*STK*) in *Arabidopsis*. *STK* has been well characterized in *Arabidopsis* and its main function is to promote ovule identity (Pinyopich et al., 2003). The fact the tomato *STK* exclusively expressed in the locular tissue suggests the conserved function of the gene in ovule development. Further transgenic analysis would help confirm the function of this gene and is in progress.

Solyc11g062060, a homolog of *NON TRANSMITTING TRACT (NTT)* gene in *Arabidopsis*, is also predominantly expressed in the locular tissue. *NTT* is a zing finger transcription factor and in *Arabidopsis*, this gene, controls the development of the transmitting tract and *ntt* mutants have reduced fertility due to severely inhibited pollen-tube movement (Crawford et al., 2007). Again, transgenic characterization of this gene in tomato would help elucidate its function.

Another interesting tissue specific transcription factor is the *YABBY* gene, *Solyc06g073920*, which in contrast to *STK*, is only expressed in the pericarp. YABBY proteins have conserved roles in specifying abaxial cell fate in lateral organs such as leaves, floral organs and ovules, and establishing the proper boundaries in meristems (Golz and Hudson, 1999; Bowman, 2000). In tomato, a mutation in a *YABBY* gene from the *facsiated* mutant results in a high locule number and more flat fruit shape (Cong et al., 2008). The homolog of *FAS* in *Arabidopsis*, *YABBY2*, however does not have the same role in regulating carpel number. It is interesting to note that although tomato *YABBY* genes involve in controlling locule number, they

are not expressed in the locular gel tissue, suggesting that the carpel number control may be independent from this gel tissue.

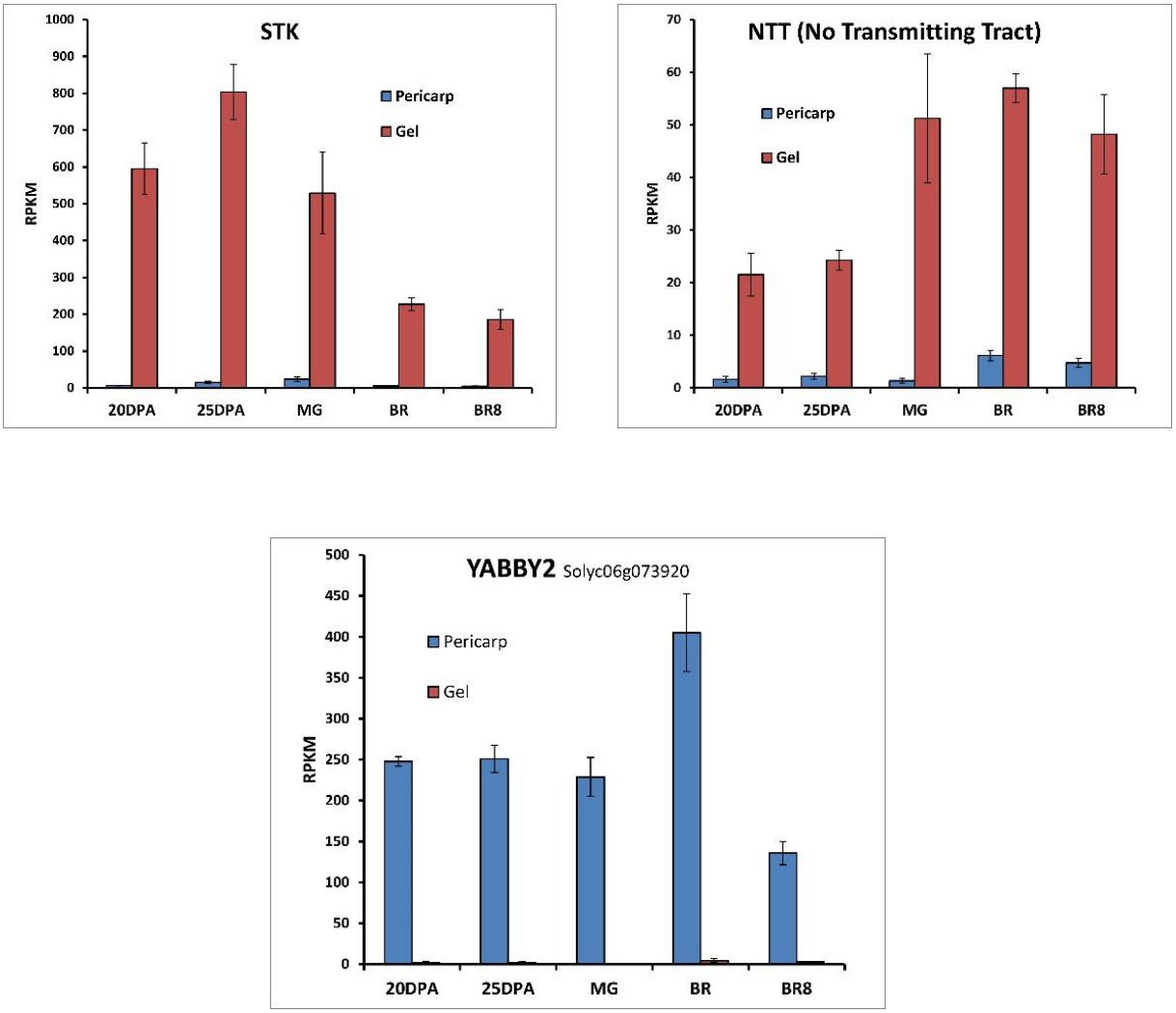


Figure 5.6. Expression levels of some tissue specific transcription factors during fruit development as determined by RPKM value from RNA-seq.

The fleshy fruit is a complex organ with different cell types and developmental processes. The transcriptome analysis described here comparing gene expression in locule tissue and pericarp presents new information about tissue-specific expression in each tissue. While ethylene synthesis appears to occur first in the pericarp, based on expression of synthesis genes in the context of the developmental stages analyzed, key transcription factors including RIN and NOR are interestingly expressed first in locule tissues and prior to obvious pericarp ripening phenotypes (color, ethylene, softening). It is noteworthy that tissue specific genes identified in this analysis could be used for isolation of promoters to create tissue-specific gene expression tools.

Table 5.2. Genes that expressed 5 fold or higher in locule compared to pericarp in all 5 stages

Solyc01g016970	Unknown Protein
Solyc01g067300	UPF0497 membrane protein 13 , trans-membrane plant subgroup
Solyc01g100880	MtN21 nodulin protein-like Protein of unknown function DUF6, transmembrane
Solyc01g107800	Expressed protein of unknown function DUF579, plant
Solyc01g107840	Immediate-early salicylate-induced glucosyltransferase
Solyc01g107850	UDP-glucosyltransferase
Solyc01g111970	Laccase 1a Multicopper oxidase, type 1
Solyc02g071050	Purine permease family protein of unknown function DUF250
Solyc02g085660	UDP-glucosyltransferase
Solyc02g086310	Lipid transfer protein
Solyc02g092680	Subtilisin-like protease Peptidase S8, subtilisin-related
Solyc03g034140	Flavoprotein wrbA
Solyc03g097580	MtN3-like protein RAG1-activating protein 1 homologue
Solyc03g098100	Reductase 2 Aldo/keto reductase subgroup
Solyc03g116390	Late embryogenesis abundant protein
Solyc03g118540	Protein TIFY 6B
Solyc04g064610	RAG1-activating protein 1 homolog
Solyc04g077220	Homeobox-leucine zipper protein 22
Solyc04g078460	N(4)-(Beta-N-acetylglucosaminyl)-L-asparaginase Peptidase T2, asparaginase 2
Solyc05g009520	Genomic DNA chromosome 3 P1 clone MPE11
Solyc05g025680	Respiratory burst oxidase protein B Ferric reductase, NAD binding
Solyc05g025690	Respiratory burst oxidase-like protein E NADPH oxidase Respiratory burst
Solyc05g052860	Os12g0581300 protein Protein of unknown function DUF620
Solyc05g052870	UDP-glucosyltransferase family 1 protein
Solyc06g064840	Agamous MADS-box transcription factor
Solyc06g073750	Beta-D-glucosidase Glycoside hydrolase, family 3, N-terminal
Solyc06g073760	Beta-D-glucosidase Glycoside hydrolase, family 3, N-terminal
Solyc07g052260	Unknown Protein (AHRD V1)
Solyc07g063880	Beta-glucosidase Glycoside hydrolase, family 1
Solyc08g078970	Pleckstrin homology domain-containing family A member 3
Solyc09g008530	RRP1 Prefoldin
Solyc09g082810	Unknown Protein Glycine rich
Solyc09g090350	Oxidoreductase Long-chain fatty alcohol dehydrogenase
Solyc10g017960	F-box protein PP2-B1 Cyclin-like F-box
Solyc11g008260	Cysteine proteinase cathepsin F Peptidase C1A, papain
Solyc11g062060	Zinc finger protein Zinc finger, C2H2-type
Solyc12g014570	Glycerophosphoryl diester phosphodiesterase family protein

Table 5.3. Genes that expressed 5 fold or higher in pericarp compared to locule in all 5 stages

Genes that expressed 5 fold higher in pericarp compared to locule	
Solyc02g081190	1-aminocyclopropane-1-carboxylate oxidase (AHRD V1 ***- Q9ZWP2_SOLLC); contains Interpro domain(s) IPR005123 Oxoglutarate and iron-dependent oxygenase
Solyc09g089730	1-aminocyclopropane-1-carboxylate oxidase-like protein (AHRD V1 **-- Q9LTH7_ARATH); contains Interpro domain(s) IPR005123 Oxoglutarate and iron-dependent oxygenase
Solyc09g010020	1-aminocyclopropane-1-carboxylate oxidase-like protein (AHRD V1 **-* Q9LTH8_ARATH); contains Interpro domain(s) IPR005123 Oxoglutarate and iron-dependent oxygenase
Solyc08g081550	1-aminocyclopropane-1-carboxylate synthase (AHRD V1 **** Q96579_SOLLC); contains Interpro domain(s) IPR004839 Aminotransferase, class I and II
Solyc03g005090	4-coumarate-coa ligase (AHRD V1 *-*- B9MY03_POPTR); contains Interpro domain(s) IPR000873 AMP-dependent synthetase and ligase
Solyc08g075430	ABC transporter G family member 14 (AHRD V1 ***- AB14G_ARATH); contains Interpro domain(s) IPR013525 ABC-2 type transporter
Solyc08g076980	Acetylornithine deacetylase or succinyl-diaminopimelate desuccinylase (AHRD V1 *-*- D4MFM7_9ENTE); contains Interpro domain(s) IPR002933 Peptidase M20
Solyc08g076970	Acetylornithine deacetylase or succinyl-diaminopimelate desuccinylase (AHRD V1 **-- D4MFM7_9ENTE); contains Interpro domain(s) IPR011650 Peptidase M20, dimerisation
Solyc08g067160	Acyl-protein thioesterase 2 (AHRD V1 **-- B6T1C9_MAIZE); contains Interpro domain(s) IPR003140 Phospholipase/carboxylesterase
Solyc02g087110	Alpha-dioxygenase (AHRD V1 **-- Q5GQ66_PEA); contains Interpro domain(s) IPR002007 Haem peroxidase, animal
Solyc04g077050	Amino acid permease 6 (AHRD V1 **** Q7Y076_BRANA); contains Interpro domain(s) IPR013057 Amino acid transporter, transmembrane
Solyc05g006910	AMP deaminase (AHRD V1 **** D3BRS6_POLPA); contains Interpro domain(s) IPR006329 AMP deaminase
Solyc09g005860	Ankyrin repeat family protein (AHRD V1 **-* Q9LVG7_ARATH); contains Interpro domain(s) IPR002110 Ankyrin
Solyc11g013110	Anthocyanidin synthase (AHRD V1 **** B2M0X8_GINBI); contains Interpro domain(s) IPR005123 Oxoglutarate and iron-dependent oxygenase
Solyc02g062140	Armadillo/beta-catenin repeat family protein (AHRD V1 ***- D7MDH4_ARALY); contains Interpro domain(s) IPR011989 Armadillo-like helical
Solyc05g056040	Auxin response factor 14 (AHRD V1 *-*- D9HNU1_MAIZE); contains Interpro domain(s) IPR003340 Transcriptional factor B3
Solyc11g013310	Auxin transporter-like protein 3 (AHRD V1 ***- B6TPT6_MAIZE); contains Interpro domain(s) IPR013057 Amino acid transporter, transmembrane
Solyc02g021260	B3 domain-containing protein LOC_Os12g40080 (AHRD V1 *-*- Y1208_ORYSJ); contains Interpro domain(s) IPR003340 Transcriptional factor B3
Solyc02g080290	Beta-glucosidase 47 (AHRD V1 **** D7MEP5_ARALY); contains Interpro domain(s) IPR001360 Glycoside hydrolase, family 1
Solyc07g020960	BHLH transcription factor-like protein (AHRD V1 *-** Q5ZAK6_ORYSJ); contains Interpro domain(s) IPR011598 Helix-loop-helix DNA-binding
Solyc08g068150	BURP domain-containing protein (AHRD V1 **-- B2ZPK7_SOLLC); contains Interpro domain(s) IPR004873 BURP
Solyc05g005560	BURP domain-containing protein (Fragment) (AHRD V1 *--- C1PI44_ARATH); contains Interpro domain(s) IPR004873 BURP
Solyc12g056860	BZIP transcription factor (AHRD V1 **** A7Y1S1_BRAJU); contains Interpro domain(s) IPR011616 bZIP transcription factor, bZIP-1
Solyc10g078390	Calcium-dependent protein kinase-like (AHRD V1 ***- Q9LET1_ARATH); contains Interpro domain(s) IPR002290 Serine/threonine protein kinase
Solyc12g094620	Catalase (AHRD V1 ***- Q2PYW5_SOLTU); contains Interpro domain(s) IPR018028 Catalase related subgroup
Solyc07g053010	Cc-nbs-lrr, resistance protein
Solyc05g053550	Chalcone synthase (AHRD V1 ***- C5IWS6_NICAL); contains Interpro domain(s) IPR011141 Polyketide synthase, type III
Solyc07g064020	class I heat shock protein 1 (AHRD V1 ***- B6TTC8_MAIZE); contains Interpro domain(s) IPR008978 HSP20-like chaperone
Solyc01g066880	Copper chaperone (AHRD V1 *--- B6TJ24_MAIZE); contains Interpro domain(s) IPR006121 Heavy metal transport/detoxification protein

Solyc11g062100	Copper-transporting ATPase (AHRD V1 ***- C1H876_PARBA); contains Interpro domain(s) IPR001757 ATPase, P-type, K/Mg/Cd/Cu/Zn/Na/Ca/Na/H-transporter
Solyc08g074480	Cortical cell-delineating protein (AHRD V1 **-- B6UGA2_MAIZE); contains Interpro domain(s) IPR013770 Plant lipid transfer protein and hydrophobic protein, helical
Solyc06g068800	CRAL/TRIO domain containing protein expressed (AHRD V1 **-- Q75GU9_ORYSJ); contains Interpro domain(s) IPR001251 Cellular retinaldehyde-binding/triple function, C-terminal
Solyc12g013900	CT099 (Fragment) (AHRD V1 **-- Q30GX2_SOLPI); contains Interpro domain(s) IPR003245 Plastocyanin-like
Solyc02g089770	Dihydroflavonol-4-reductase (AHRD V1 ***- B6TK03_MAIZE); contains Interpro domain(s) IPR016040 NAD(P)-binding domain
Solyc11g069930	Disease resistance protein R3a-like protein (AHRD V1 ***- B3F572_SOLDE)
Solyc04g015520	DNA binding protein (AHRD V1 **** B6TND2_MAIZE); contains Interpro domain(s) IPR005516 Remorin, C-terminal region
Solyc06g071280	Enhanced disease susceptibility 1 (Fragment) (AHRD V1 **-- B2BDU7_ARALY); contains Interpro domain(s) IPR002921 Lipase, class 3
Solyc05g051200	Ethylene-responsive transcription factor 1A (AHRD V1 *-** A9P6A4_MEDTR); contains Interpro domain(s) IPR001471 Pathogenesis-related transcriptional factor and ERF, DNA-binding
Solyc09g010860	Expansin (AHRD V1 ***- Q9ZP32_SOLLC); contains Interpro domain(s) IPR007112 Expansin 45, endoglucanase-like IPR007117 Pollen allergen/expansin, C-terminal
Solyc02g085870	Fatty acid elongase 3-ketoacyl-CoA synthase (AHRD V1 **** Q6DUV5_BRANA); contains Interpro domain(s) IPR012392 Very-long-chain 3-ketoacyl-CoA synthase
Solyc08g067260	Fatty acid elongase 3-ketoacyl-CoA synthase (AHRD V1 **** Q6DUV6_BRANA); contains Interpro domain(s) IPR012392 Very-long-chain 3-ketoacyl-CoA synthase
Solyc08g068390	Fatty acid oxidation complex subunit alpha (AHRD V1 ***- FADJ_ECOHS); contains Interpro domain(s) IPR006176 3-hydroxyacyl-CoA dehydrogenase, NAD binding
Solyc02g083860	Flavanone 3-hydroxylase (AHRD V1 ***- A9ZMI9_TOBAC); contains Interpro domain(s) IPR005123 Oxoglutarate and iron-dependent oxygenase
Solyc12g094650	Genomic DNA chromosome 5 P1 clone MFC19 (AHRD V1 ***- Q9FHI2_ARATH); contains Interpro domain(s) IPR007658 Protein of unknown function DUF594
Solyc01g101120	Glucan endo-1 3-beta-glucosidase 1 (AHRD V1 *-*- B4FGE4_MAIZE); contains Interpro domain(s) IPR012946 X8
Solyc04g080460	Glucose transporter 8 (AHRD V1 ***- Q2KKJ3_SOLIN); contains Interpro domain(s) IPR003663 Sugar/inositol transporter
Solyc09g011600	Glutathione S-transferase-like protein (AHRD V1 **** A8DUB0_SOLLC); contains Interpro domain(s) IPR004046 Glutathione S-transferase, C-terminal
Solyc07g017780	H-ATPase (AHRD V1 **** Q8L6I3_ORYSJ); contains Interpro domain(s) IPR006534 ATPase, P-type, plasma-membrane proton-efflux
Solyc02g091930	Homeobox-leucine zipper protein 22 (AHRD V1 ***- D7MA03_ARALY); contains Interpro domain(s) IPR003106 Leucine zipper, homeobox-associated
Solyc11g065070	Hydroxymethylglutaryl-CoA lyase (AHRD V1 ***- B6U7B9_MAIZE); contains Interpro domain(s) IPR000891 Pyruvate carboxyltransferase
Solyc12g056450	Hydroxymethylglutaryl-CoA synthase (AHRD V1 ***- A9ZMZ7_HEVBR); contains Interpro domain(s) IPR010122 Hydroxymethylglutaryl-CoA synthase, eukaryotic
Solyc06g068520	Hydroxyproline-rich systemin (AHRD V1 ***- HSY1_SOLLC)
Solyc00g009110	Inositol-1 4 5-trisphosphate 5-Phosphatase-like protein (AHRD V1 *-*- Q9M1W5_ARATH); contains Interpro domain(s) IPR000300 Inositol polyphosphate related phosphatase
Solyc04g015690	Insulin degrading enzyme (AHRD V1 **-- Q93YG9_SOLLC); contains Interpro domain(s) IPR011237 Peptidase M16, core
Solyc01g099170	Lipoxygenase (AHRD V1 ***- Q9FT17_SOLLC); contains Interpro domain(s) IPR000907 Lipoxygenase
Solyc08g014000	Lipoxygenase (AHRD V1 **** Q43190_SOLTU); contains Interpro domain(s) IPR001246 Lipoxygenase, plant
Solyc01g099160	Lipoxygenase (AHRD V1 **** Q9FT17_SOLLC); contains Interpro domain(s) IPR001246 Lipoxygenase, plant
Solyc03g114840	MADS box transcription factor 11 (AHRD V1 **** D9IFM5_ONCHC); contains Interpro domain(s) IPR002100 Transcription factor, MADS-box IPR002487 Transcription factor, K-box
Solyc02g084630	MADS-box transcription factor (AHRD V1 ***- Q4LEZ4_ASPOF); contains Interpro domain(s) IPR002100 Transcription factor, MADS-box

Solyc08g082120	Methanol inducible protein (AHRD V1 ***- D1MAX5_NICBE)
Solyc01g005080	Microtubule-associated protein MAP65-1a (AHRD V1 **** C0J9Z7_ORYNI); contains Interpro domain(s) IPR007145 MAP65/ASE1
Solyc03g117540	Multidrug resistance protein ABC transporter family (AHRD V1 **-- B9I191_POPTR); contains Interpro domain(s) IPR001140 ABC transporter, transmembrane region
Solyc05g014380	Multidrug resistance protein ABC transporter family (AHRD V1 ***- B9HTC0_POPTR); contains Interpro domain(s) IPR001140 ABC transporter, transmembrane region
Solyc12g006800	Myb family transcription factor-like (AHRD V1 **-- Q6Z869_ORYSJ); contains Interpro domain(s) IPR006447 Myb-like DNA-binding region, SHAQKYF class
Solyc08g076400	Myb family transcription factor-like protein (AHRD V1 *--- Q84UP8_ORYSJ); contains Interpro domain(s) IPR006447 Myb-like DNA-binding region, SHAQKYF class
Solyc09g083020	Myrosinase-binding protein-like protein (AHRD V1 **-* Q9M5W9_ARATH); contains Interpro domain(s) IPR001229 Mannose-binding lectin
Solyc11g012930	Nodulin family protein (AHRD V1 ***- A7L2Z5_GOSHI); contains Interpro domain(s) IPR000620 Protein of unknown function DUF6, transmembrane
Solyc08g083060	Oligopeptide transporter (Fragment) (AHRD V1 **** B0DAK3_LACBS); contains Interpro domain(s) IPR004813 Oligopeptide transporter OPT superfamily
Solyc05g008090	Os02g0200800 protein (Fragment) (AHRD V1 *-*- Q0E303_ORYSJ); contains Interpro domain(s) IPR009675 Targeting for Xklp2
Solyc02g077080	Os06g0207500 protein (Fragment) (AHRD V1 **-- Q0DDQ9_ORYSJ); contains Interpro domain(s) IPR004253 Protein of unknown function DUF231, plant
Solyc08g005800	Pectinacetylesterase like protein (Fragment) (AHRD V1 **-- Q56WP8_ARATH); contains Interpro domain(s) IPR004963 Pectinacetylesterase
Solyc07g064190	Pectinesterase (AHRD V1 ***- B9RXQ4_RICCO); contains Interpro domain(s) IPR000070 Pectinesterase, catalytic
Solyc07g064170	Pectinesterase (AHRD V1 ***- B9RXQ4_RICCO); contains Interpro domain(s) IPR018040 Pectinesterase, active site IPR000070 Pectinesterase, catalytic
Solyc07g064180	Pectinesterase (AHRD V1 ***- B9RXQ4_RICCO); contains Interpro domain(s) IPR018040 Pectinesterase, active site IPR000070 Pectinesterase, catalytic
Solyc12g009270	Pectinesterase (AHRD V1 **** C0PST8_PICSI); contains Interpro domain(s) IPR006501 Pectinesterase inhibitor
Solyc04g071890	Peroxidase 4 (AHRD V1 **** B7UCP4_LITCN); contains Interpro domain(s) IPR002016 Haem peroxidase, plant/fungal/bacterial
Solyc02g031740	phloem lectin (AHRD V1 ***- Q8L5A9_CUCSA)
Solyc02g069020	Phloem protein 2 (AHRD V1 **-- D0R6I5_MALDO)
Solyc03g077920	Phosphoribosylanthranilate transferase (Fragment) (AHRD V1 *--- Q43085_PEA); contains Interpro domain(s) IPR013583 Phosphoribosyltransferase C-terminal, plant
Solyc01g065500	Phosphoribosylanthranilate transferase (Fragment) (AHRD V1 *--- Q43085_PEA); contains Interpro domain(s) IPR013583 Phosphoribosyltransferase C-terminal, plant
Solyc05g051750	Prosystemin (AHRD V1 ***- Q5MBK9_BETVU); contains Interpro domain(s) IPR009966 Prosystemin
Solyc01g103370	Ras-related protein Rab-1A (AHRD V1 ***- C0HBJ4_SALSA); contains Interpro domain(s) IPR003579 Ras small GTPase, Rab type
Solyc09g055180	Receptor like kinase, RLK
Solyc05g056370	Receptor like kinase, RLK
Solyc04g005910	Receptor like kinase, RLK
Solyc09g061930	Receptor like kinase, RLK
Solyc12g039080	Receptor-like kinase (AHRD V1 ***- A7VM37_MARPO)
Solyc07g005140	Response regulator 9 (AHRD V1 ***- Q94LW1_MAIZE); contains Interpro domain(s) IPR017053 Response regulator, plant B-type
Solyc05g008390	RING zinc finger-containing protein (AHRD V1 *--- Q54MR8_DICDI); contains Interpro domain(s) IPR011016 Zinc finger, RING-CH-type
Solyc05g009010	Serine/threonine protein kinase family protein (AHRD V1 **** C6ZRM5_SOYBN); contains Interpro domain(s) IPR002290 Serine/threonine protein kinase
Solyc01g112220	Serine/threonine protein kinase-like (AHRD V1 **** Q5XWQ1_SOLTU); contains Interpro domain(s) IPR002290 Serine/threonine protein kinase

Solyc03g081260	Subtilisin-like protease-like protein (AHRD V1 **-- Q9T0B5_ARATH); contains Interpro domain(s) IPR015500 Peptidase S8, subtilisin-related
Solyc07g005400	Transcription factor (AHRD V1 *-** D6MKM4_9ASPA); contains Interpro domain(s) IPR011598 Helix-loop-helix DNA-binding
Solyc01g066740	Transducin family protein (Fragment) (AHRD V1 *--- D7L607_ARALY); contains Interpro domain(s) IPR017986 WD40 repeat, region
Solyc01g014230	U-box domain-containing protein 4 (AHRD V1 **-- PUB4_ARATH); contains Interpro domain(s) IPR011989 Armadillo-like helical
Solyc08g082190	Unknown Protein (AHRD V1)
Solyc02g086460	Unknown Protein (AHRD V1)
Solyc06g059870	Unknown Protein (AHRD V1)
Solyc01g044240	Unknown Protein (AHRD V1)
Solyc03g112200	Unknown Protein (AHRD V1)
Solyc04g040210	Unknown Protein (AHRD V1)
Solyc02g076800	Unknown Protein (AHRD V1)
Solyc02g063440	Unknown Protein (AHRD V1)
Solyc09g074040	Unknown Protein (AHRD V1)
Solyc05g013630	Unknown protein (Fragment) (AHRD V1 ***- Q43520_SOLLC)
Solyc04g055050	VQ motif family protein expressed (AHRD V1 *-*- Q10KE8_ORYSJ); contains Interpro domain(s) IPR008889 VQ
Solyc06g073920	YABBY-like transcription factor CRABS CLAW-like protein (AHRD V1 **-* Q6SRZ7_ANTMA); contains Interpro domain(s) IPR006780 YABBY protein
Solyc07g008180	YABBY-like transcription factor CRABS CLAW-like protein (AHRD V1 **-* Q6SRZ7_ANTMA); contains Interpro domain(s) IPR006780 YABBY protein
Solyc06g082010	Zinc finger CCCH domain-containing protein 66 (AHRD V1 ***- C3H66_ARATH); contains Interpro domain(s) IPR002110 Ankyrin

III. Material and Methods

Plant materials

Plants were grown in greenhouses at the Boyce Thompson Institute for Plant Research, Ithaca, NY with a 16 h light (27^{o}C) and 8 h dark (19^{o}C) cycle. Tomato fruit were tagged at 7-8 days post anthesis (dpa) when they were 1 cm in diameter and reached breaker stage at 37 to 38 dpa.

Transcriptome sequencing

Fruit were harvested at the stages indicated in the text from three to six biological replicates, each replicate contained 4-5 combined fruit of the same developmental stage and from the same plant. Fruit were cross-sectioned and pericarp and locular tissues were collected. Total RNA from pericarp and locular tissues was used to construct strand-specific RNA libraries as described (Zhong et al., 2011). Three to six biological replicates for each sample were sequenced using a HiSeq2000 Sequencing System (Illumina, CA) according to the manufacturer's instructions.

RNA seq data processing

RNA-Seq reads were first processed by removing barcode and adaptor sequences. The resulting reads were aligned to ribosomal RNA (rRNA) and transfer RNA (tRNA) sequences using Bowtie allowing up to three mismatches (Langmead et al., 2009) to remove potential contaminating reads. The resulting cleaned reads were then aligned to the tomato genome using Tophat allowing one segment mismatch (Trapnell et al., 2009) . Following alignments, raw counts of mapped reads for each tomato gene model were derived and then normalized to reads per kilobase of exon model per million mapped reads (RPKM). Differentially expressed genes

between transgenic and control tissue were identified using the DESeq package (Anders and Huber, 2010). Raw p values were adjusted for multiple testing using the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995). Dr. Zhangjun Fei helped analyzed the RNA-seq data.

REFERENCES

Alexander, L., and Grierson, D. (2002). Ethylene biosynthesis and action in tomato: a model for climacteric fruit ripening. Journal of Experimental Botany 53: 2039-2055.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biology 11: R106.

Atta-Aly, M.A., Brecht, J.K., and Huber, D.J. (2000). Ripening of tomato fruit locule gel tissue in response to ethylene. Postharvest Biology and Technology 19: 239-244.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society. Series B (Methodological) 57: 289-300.

Bowman, J.L. (2000). The YABBY gene family and abaxial cell fate. Current Opinion in Plant Biology 3: 17-22.

Brecht, J. (1987). Locular gel formation in developing tomato fruit and the initiation of ethylene production. HortScience, 22:476-479.

Cheng, G.W., and Huber, D.J. (1996). Alterations in structural polysaccharides during liquefaction of tomato locule tissue. Plant Physiology 111: 447-457.

Cong, B., Barrero, L.S., and 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.

Crawford, B.C., Ditta, G., and Yanofsky, M.F. (2007). The NTT gene is required for transmitting-tract development in carpels of Arabidopsis thaliana. Current Biology 17: 1101-1108.

Gapper, N.E., McQuinn, R.P., and Giovannoni, J.J. (2013). Molecular and genetic regulation of fruit ripening. Plant Molecular Biology 82: 575-591.

Gillaspy, G., Ben-David, H., and Gruissem, W. (1993). Fruits: A developmental perspective. Plant Cell 5: 1439-1451.

Giovannoni, J.J. (2004). Genetic regulation of fruit development and ripening. Plant Cell 16: S170-180.

Giovannoni, J.J. (2007). Fruit ripening mutants yield insights into ripening control. Current Opinion in Plant Biology 10: 283-289.

Golz, J.F., and Hudson, A. (1999). Plant development: YABBYs claw to the fore. Current Biology 9: R861-R863.

Klee, H.J., and Giovannoni, J.J. (2011). Genetics and control of tomato fruit ripening and quality attributes. Annual Review of Genetics 45: 41-59.

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

Liljegren, S.J., Ditta, G.S., Eshed, Y., Savidge, B., Bowman, J.L., and Yanofsky, M.F. (2000). SHATTERPROOF MADS-box genes control seed dispersal in Arabidopsis. Nature 404: 766-770.

Matas, A.J., Yeats, T.H., Buda, G.J., Zheng, Y., Chatterjee, S., Tohge, T., Ponnala, L., Adato, A., Aharoni, A., Stark, R., Fernie, A.R., Fei, Z.J., Giovannoni, J.J., and Rose, J.K.C. (2011). Tissue-and cell-type specific transcriptome profiling of expanding tomato fruit provides insights into metabolic and regulatory specialization and cuticle formation. Plant Cell 23: 3893-3910.

Pinyopich, A., Ditta, G.S., Savidge, B., Liljegren, S.J., Baumann, E., Wisman, E., and Yanofsky, M.F. (2003). Assessing the redundancy of MADS-box genes during carpel and ovule development. Nature 424: 85-88.

Rose, J.K., Lee, H.H., and Bennett, A.B. (1997). Expression of a divergent expansin gene is fruit-specific and ripening-regulated. Proceedings of the National Academy of Sciences 94: 5955-5960.

Sampedro, J., and Cosgrove, D.J. (2005). The expansin superfamily. Genome biology 6.

Seymour, G.B., Ostergaard, L., Chapman, N.H., Knapp, S., and Martin, C. (2013). Fruit development and ripening. Annual Review of Plant Biology 64: 219-241.

Smith, C.J.S., Watson, C.F., Ray, J., Bird, C.R., Morris, P.C., Schuch, W., and Grierson, D. (1988). Antisense RNA inhibition of polygalacturonase gene expression in transgenic tomatoes. Nature 334: 724-726.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25: 1105-1111.

Wilkinson, J.Q., Lanahan, M.B., Yen, H.C., Giovannoni, J.J., and Klee, H.J. (1995). An ethylene-inducible component of signal transduction encoded by never-ripe. Science 270: 1807-1809.

Zhong, S., Joung, J.G., Zheng, Y., Chen, Y.R., Liu, B., Shao, Y., Xiang, J.Z., Fei, Z., and Giovannoni, J.J. (2011). High-throughput illumina strand-specific RNA sequencing library preparation. Cold Spring Harbor Protocols 2011: 940-949.

Zhong, S., Fei, Z., Chen, Y.R., Zheng, Y., Huang, M., Vrebalov, J., McQuinn, R., Gapper, N., Liu, B., Xiang, J., Shao, Y., and Giovannoni, J.J. (2013). Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening. Nature Biotechnology 31: 154-159.

SUMMARY AND FUTURE DIRECTIONS

Chloroplast manipulation

The area I would consider the highlight of my thesis is isolating the gene underlying the green shoulder/uniform ripening (*u*) phenotype. The gene encodes a *Golden 2-like* transcription factor (*GLK*) and functional characterization of this fruit-related gene (*SlGLK2*), and its leaf-expressed family member (*SlGLK1*), through a combination of mutant, transgenic, biochemical and transcriptomic analyses revealed their critical roles in fruit and leaf plastid accumulation, respectively. The results of these efforts have shed light on the regulation of chlorophyll and chloroplast development during fruit maturation and confirm the important role of plastid organelles as related to fruit nutritional quality traits. This work further highlights the unintended consequences of extensive breeding for *u* in commercial tomato gemplasm and suggests strategies toward enhancing fruit quality, both in a general sense via targeting elevated plastid accumulation, and in a direct sense via the development of nearly isogenic breeding lines harboring the *u* allele and tightly linked DNA markers.

The results of this project reveal a promising strategy to manipulate the levels of metabolites in fruit. In short, instead of changing the rate of metabolic flux by replacing or manipulating rate-limiting factors in biosynthetic pathways, an optimal strategy may prove to be targeting change in the size or number of the organelles or compartments in which reactions of interest occur. In plants, chloroplasts are the sites of photosynthesis in addition to synthesis of many nutritionally relevant secondary metabolites. Indeed, most of the precursors of flavor-related volatiles are synthesized within chloroplasts. This organelle can also serve as the precursor of chromoplasts, the site of carotenoid accumulation. Thus, rather than focusing on

individual pathways, an advantage of manipulating chloroplasts levels is that it can simultaneously affect multiple nutritionally relevant metabolite compounds.

Over-expression of *GLKs* results in dark green immature fruit with elevated chlorophyll and chloroplast content in addition to elevated sugars, carotenoids and ascorbic acid in ripe fruit. Recently, Pan et al., (2013) deployed a network analysis approach to identify a transcription factor related to the *ARABIDOPSIS PSEUDO RESPONSE REGULATOR2-LIKE* gene (*APRR2-like*) that also influences pigmentation and ripening in tomato fruit. Transgenic tomato lines over-expressing the *APRR2-like* gene produced fruits with larger and more numerous plastids and consequently higher chlorophyll levels in immature unripe fruits. They also presented elevated carotenoid accumulation in ripe fruits. While neither I nor these authors measured flavor volatiles, they are likely also elevated as many are carotenoid or plastid derived. Confirmation of flavor volatile effects would further demonstrate the value of manipulating chloroplasts toward attaining biofortification goals.

Additional logical future objectives would include quantification of fruit flavor volatile compounds in transgenic/mutant plants that have perturbations in expression (over- or under-expression) of genes such as *GLKs*, *APRR2-like*, and *hp1* (*high pigment* 1 mutation) and/or combinations of the same. Transgenic plants overexpressing such genes, as well as plants with combined gene perturbations through crossing, could be tested for interaction/additivity. It would also be useful to assess the transcriptomes of unripe and ripe fruit from such lines (e.g. by RNA sequencing) to gain insight into the molecular basis of observed metabolic changes.

The results obtained in pursuit of this project can also be applied to other fruit species. For example, Brand et al (2014) recently showed that pepper *GLK2* also regulates natural variation in chlorophyll content and fruit color of pepper fruit, confirming conserved function of

the gene in fruit development in different species. It would be interesting and useful to assess the genetic variation of GLKs in additional fruit species and ask whether or not said variation shows any correlation with accumulation of nutrient compounds and quality-associated metabolites, especially those know to be plastid derived.

Finally, as indicated in chapter 3, the results presented herein suggest benefits for both growers and consumers in using the *U* allele, especially under growth conditions that limit the off-color shoulder phenotype. Practical greenhouse and field trials in different areas using multiple *U/U* versus *u/u* cultivars would help to optimize cultural-practice/genotype combinations for utilizing the *U* allele toward enhanced fruit quality.

Gene expression gradients

The green shoulder phenotype and *GLK* gradient expression represent a very interesting phenomenon manifesting along the latitudinal (top – bottom) axis of the fruit and whose molecular basis is poorly understood. Transcriptome profiling reported here revealed that the gradient includes genes associated with general ripening activities beyond plastid development and thus is likely to influence ripening polarity, chlorophyll distribution and sugar and carotenoid accumulation. Interestingly, no previously described ripening control genes such as *RIN*, *NOR*, or *CNR* display such patterns, suggesting that other factors regulate the formation of the gradient during fruit development and ripening of tomato.

These observations allow one to begin to ask a number of additional questions. For example: How early during fruit development are such gradients established? Is there a single gradient mechanism in the developing fruit or are there multiple gradient systems? The fact that *u/u* fruit still display a carotenoid gradient during ripening suggests that multiple gradient regulators operate in the maturing fruit. In this study, fruit were collected as early as 10 days

post-anthesis and the green shoulder phenotype was clearly visible at this time point. However, personal observations suggest that the green shoulder phenotype can be observed as early as when the ovary is 1-2 mm in diameter (just 1 to several days post-pollination), although whether it can be detected before pollination remains to be investigated. This would be relatively easy to determine using appropriate microscopy techniques. If the gradient is visible (or as early as it is visible), it would be helpful to perform transcriptome analysis to get closer to causative molecular events. If the transition is visible at all stages including when hand sectioning is difficult, microdissection methods including laser capture microdissection (LCM) could be a feasible approach to explore transcriptome variation along the gradient (Matas et al., 2011).

Based on the results obtained in this thesis project we do have some ideas as to which genes regulate fruit development gradients. Though questions remain regarding developing a full understanding of gradient biology in the maturing fruit, RNA-seq results showed that several transcription factors including members of the *KNOX*, *MADS*, *NGATHA* (*ARF*-related) and *WRKY* families displayed gradient expression in both green and red fruit. In the case of the KNOX gene *TKN4*, Nadakuduti et al., (2014) functionally confirmed this candidate as acting upstream of *GLK2*. This *knotted*-like gene, and its family members, would be logical targets for further analysis to ascertain whether they affect other genes besides *GLK2* and whether or not additional KNOX genes influence fruit gradient biology.

Beyond *TKN4* there are certainly additional factors acting upstream in the regulatory hierarchy of plastid gradient formation. Targeted transcriptome analysis on *u* and *ug* mutant fruit tissues (as compared to normal controls) could reveal additional candidates in this regulatory hierarchy. It is interesting to note that tomato cultivars with elevated anthocyanin content also demonstrate a higher concentration of metabolites at the shoulder/stem end of the fruit (Gonzali

et al., 2009) and these cultivars could be additional substrates for similar RNA-seq analysis. Such transcriptome studies coupled with QTL mapping could quickly localize and isolate candiate genes that control these anthocyanin accumulation phenotypes and may lead to clues for further upstream factors controlling gradient formation.

Fruit ripening gradient

An obvious question resulting from this project is, what are the factors that control the visually obvious ripening gradient? In tomato, the ripening process (at least as defined by pigment accumulation), usually does not occur uniformly throughout the fruit but initiates at the stylar end (bottom) of the fruit. This ripening gradient occurs in *u* mutant fruit as well suggesting distinction between the chlorophyll and ripening gradients. Besides color changes, total soluble solid (^{o}Brix) levels, as well as expression of the sucrose hydrolysis invertase gene (*sucr*), showed strong gradients from bottom to top, consistent with ripening progression. As noted above, previously characterized ripening control genes do not display expression gradients consistent with ripening gradient control. Finding the underlying regulatory factors would definitely contribute important information to our understanding of the fruit ripening processes.

As a start, several candidate genes from the *NGATHA* and *WRKY* transcription factor families were targeted for functional analyses. Transgenic plants altered in expression of several of these genes have been created as a first step toward better understanding their functions. For example, the pointed and elongated fruit phenotype in some *NGA1* RNAi lines is consistent with the model in which *NGATHA* genes are positive regulators of auxin, contributing to an auxin gradient along the apical-basal axis of the gynoecium. Preliminary analysis indicated a clear function in fruit shape, but not a clear ripening phenotype.

Further work to fully characterize these genes, includes measuring auxin levels in the *NGATHA* knockdown fruit and creating more stable knockouts in all 3 members of the gene family, would help identify specific and redundant functions in this family of genes displaying clear fruit gradient expression patterns. Research in Arabidopsis showed that the strongest phenotypes were obtained in this family with quadruple mutants. Recently, CRISPR, a new tool for gene editing is emerging as a powerful approach to create complete loss-of-function mutations and has been successful in a wide range of species, including tomato (Cong et al., 2013; Brooks et al., 2014). CRISPR could be deployed to create mutations in all three *NGATHA* genes and then single, double and triple mutants could be assessed for function in fruit morphology and ripening phenotypes.

In the case of the *SlWRKY51* gene, no *RNAi* plants were recovered from two independent transformation experiments, suggesting lethality. Tissue-specific expression (for example, fruit-specific expression) of repression constructs, allowing recovery of viable transgenic plants, would be the appropriate way forward to functionally assess this particular gene.

Characterization of locular gel and pericarp tissues

In contrast to the molecular gradients along the fruit latitudinal axis, where differentially expressed genes occur in the same tissue (pericarp), the locular gel is distinct from the pericarp, with the locular gel originating from the placenta and the pericarp from the ovary walls. The locular tissue is highly parenchymous and gradually solubilizes during fruit maturation. The locular tissue is of interest as it shows some of the earliest signs of change prior to, or at the onset of, ripening. RNA-seq analysis revealed the increase of expression of ethylene responsive genes such as *E4* and *E8*, as well as master ripening regulators including *RIN* and *NOR*, occurs earlier in locular tissue compared to the pericarp, suggesting that ripening phenomena may well

initiate in the locule. Surprisingly, while others have reported ethylene to elevate first in the locules (Brecht, 1987), ethylene biosynthetic genes in general are expressed higher in the pericarp compared to the locule. It is possible that the stages examined here (e.g. MG and Breaker) bracket the point where initial ethylene synthesis may occur in the locule and the optimal point of induction was missed. Indeed, in tomato, the mature green stage can extend for several days. To address this issue it would be useful to further divide the mature green stage into narrower time points and characterize transcriptomic activity during these more refined periods. For example, in *Alisa Craig* varieties, the mature green stage is reached at approximately 35 days after anthesis so fruit might be collected each day from 30 days until breaker stage (~42 dpa) and subject to RNA-seq analysis. Transcriptome characterization of locular tissues at even earlier fruit stages could be performed to determine the exact time of initial expression of some of the earliest ripening control genes including *NOR* and *RIN*.

Some additional interesting tissue-specific transcription factors were noted during this analysis, for example the homologs of Arabidopsis *STK* and *NTT*. These genes are highly and specifically expressed in the locular tissue in all fruit stages. In Arabidopsis, their main functions are to promote ovule identity. Transgenic analysis would help confirm the function of these genes in tomato.

Pericarp and locular tissues also display differences in expression levels of several cell wall-related genes including *Polygalacturonase* (*PG*) and *Expansin1* (*LeExp1*) and may contribute to the differences in cell wall composition of each tissue. More work can be focused in this direction as well to make or break correlations between differential gene expression and specific cell wall differences. Characterizing both transcriptome activity and cell wall composition at multiple time points as the locule and pericarp matures, would provide useful

data for identifying candidate genes from transcriptome analysis that may confer the observed cell wall changes.

Finally, locular and pericarp tissues have differences in acid as well as sugar composition. The locular tissue has higher titratable acidicity and citric acid than pericarp tissue, while the pericarp contains more reducing sugars and glucose (Stevens and Rick, 1986). The levels of these compounds may reflect changes in cell wall metabolism, and together with volatile compounds, represent the three major components of tomato flavor. Thus the balance in the sizes of locular versus pericarp tissues is important for maintaining flavor quality as well as other characteristics including firmness. These would be interesting phenotypes to observe and quantitate in segregating populations, diversity collections or fixed genetic populations derived, for example, from wild species crosses (Eshed and Zamir, 1995; Lin et al., 2014).

REFERENCES

Brand, A., Borovsky, Y., Hill, T., Rahman, K.A.A., Bellalou, A., Van Deynze, A., and Paran, I. (2014). CaGLK2 regulates natural variation of chlorophyll content and fruit color in pepper fruit. Theoretical and Applied Genetics 127: 2139-2148.

Brecht, J. (1987). Locular gel formation in developing tomato fruit and the initiation of ethylene production. HortScience, 22:476-479.

Brooks, C., Nekrasov, V., Lippman, Z.B., and Van Eck, J. (2014). Efficient Gene Editing in Tomato in the First Generation Using the Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-Associated9 System. Plant Physiology 166: 1292-1297.

Cong, L., Ran, F.A., Cox, D., Lin, S.L., Barretto, R., Habib, N., Hsu, P.D., Wu, X.B., Jiang, W.Y., Marraffini, L.A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science 339: 819-823.

Eshed, Y., and Zamir, D. (1995). An introgression line population of Lycopersicon pennellii in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. Genetics 141: 1147-1162.

Gonzali, S., Mazzucato, A., and Perata, P. (2009). Purple as a tomato: towards high anthocyanin tomatoes. Trends in Plant Science 14: 237-241.

Lin, T., Zhu, G.T., Zhang, J.H., Xu, X.Y., Yu, Q.H., Zheng, Z., Zhang, Z.H., Lun, Y.Y., Li, S., Wang, X.X., Huang, Z.J., Li, J.M., Zhang, C.Z., Wang, T.T., Zhang, Y.Y., Wang, A.X., Zhang, Y.C., Lin, K., Li, C.Y., Xiong, G.S., Xue, Y.B., Mazzucato, A., Causse, M., Fei, Z.J., Giovannoni, J.J., Chetelat, R.T., Zamir, D., Stadler, T., Li, J.F., Ye, Z.B., Du, Y.C., and Huang, S.W. (2014). Genomic analyses provide insights into the history of tomato breeding. Nature Genetics 46: 1220-1226.

Matas, A.J., Yeats, T.H., Buda, G.J., Zheng, Y., Chatterjee, S., Tohge, T., Ponnala, L., Adato, A., Aharoni, A., Stark, R., Fernie, A.R., Fei, Z.J., Giovannoni, J.J., and Rose, J.K.C. (2011). Tissue-and cell-type specific transcriptome profiling of expanding tomato fruit provides insights into metabolic and regulatory specialization and cuticle formation. Plant Cell 23: 3893-3910.

Martinez-Fernandez, I., Sanchis, S., Marini, N., Balanza, V., Ballester, P., Navarrete-Gomez, M., Oliveira, A.C., Colombo, L., and Ferrandiz, C. (2014). The effect of NGATHA altered activity on auxin signaling pathways within the Arabidopsis gynoecium. Frontiers in Plant Science 5.

Nadakuduti, S.S., Holdsworth, W.L., Klein, C.L., and Barry, C.S. (2014). KNOX genes influence a gradient of fruit chloroplast development through regulation of GOLDEN2-LIKE expression in tomato. Plant Journal 78: 1022-1033.

Pan, Y., Bradley, G., Pyke, K., Ball, G., Lu, C., Fray, R., Marshall, A., Jayasuta, S., Baxter, C., van Wijk, R., Boyden, L., Cade, R., Chapman, N.H., Fraser, P.D., Hodgman, C., and Seymour, G.B. (2013). Network inference analysis identifies an APRR2-like gene linked to pigment accumulation in tomato and pepper fruits. Plant Physiology 161: 1476-1485.

Stevens, M.A., and Rick, C.M. (1986). Genetics and Breeding. In The tomato crop, J.G. Atherton and J. Rudich, eds (London: Chapman and Hall), 35-100.