#### SPATIAL AND TEMPORAL MECHANISMS UNDERLYING APPLE FRUIT GROWTH

by

#### SHAN JING

(Under the Direction of Anish Malladi)

#### ABSTRACT

Fruit size is an important economic trait that is dependent both on genetic and environmental factors. Apple (*Malus*  $\times$  *domestica Borkh*) fruit has a unique structure where the fleshy part of the fruit (cortex) is derived largely from non-ovarian tissue while the core (pith) is the true fruit derived from the ovary. In this study, we investigated the molecular and metabolic mechanisms regulating apple fruit development in a spatial and temporal manner. Results indicated that greater growth occurred in the cortex than the pith during fruit development due to more cell production and expansion. Targeted metabolite profiling revealed distinct characteristics of C and N metabolism between the pith and cortex that may be associated with different growth patterns in different tissue. Variations of metabolic profiles were also revealed between early and mid/late fruit development. A subset of organ growth regulating genes have been identified and characterized in model species. The potential functions of organ growth regulating genes from five gene families were investigated in apple through analysis of their transcript abundance: Fruit Weight 2.2/Cell Number Regulator, Growth Regulating Factor (GRF), GRF-interacting Factor (GIF), ARGOS/ARGOS-Like and KLUH. These data indicate that MdGRF7a and MdGIF3 may act as positive regulators of cell production, while *MdCNR5a* may function as a negative regulator of cell production during fruit development in apple. RNA-seq analysis was performed to understand spatiotemporal and fruit load reduction-related changes in the transcriptome during fruit development. The cortex and pith displayed distinctive transcriptome profiles during all three stages analyzed: cell production, transition from cell production, and cell expansion. Weighted gene co-expression network analysis (WGCNA) revealed a negative regulator of cell production: *TEOSINTE BRANCHED 1/CYCLOIDEA/PROLIFERATING CELL NUCLEAR ANTIGEN FACTOR1 (TCP)*. Transcript accumulation pattern of this transcription factor was consistent with negative regulation of spatiotemporal cell production patterns in the developing fruit. A regulatory cascade involving several *miRNAs* and this *TCP* has been previously implicated in regulation of cell production during Arabidopsis leaf growth. Data presented here are consistent with a similar role for components of this cascade in regulating cell production during apple fruit growth.

INDEX WORDS: apple, fruit development, cell production, metabolic profiling, RNA-seq

# SPATIAL AND TEMPORAL MECHANISMS UNDERLYING APPLE FRUIT GROWTH

by

# SHAN JING

B.S., Beijing Forestry University, China, 2009M.S., Beijing Forestry University, China, 2012

A dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2019

© 2019

Shan Jing

All Rights Reserved

# SPATIAL AND TEMPORAL MECHANISMS UNDERLYING APPLE FRUIT GROWTH

by

## SHAN JING

Major Professor: Committee: Anish Malladi Chang Hyun Khang Wolfgang Lukowitz Savithri Nambeesan Dayton Wilde

Electronic Version Approved:

Suzanne Barbour Dean of the Graduate School The University of Georgia May 2019

# DEDICATION

To my father Jing Youzhong,

who taught me to never give up on my dream.

To my mother Xu Zhengfeng,

who taught me to be optimistic whatever challenges we face in life.

To my husband Shangpeng,

for always making me laugh even if the dream does not come true.

#### ACKNOWLEDGEMENTS

It has been a long but pleasant journey as I have had the fortune to work with a group of people without whom none of this would have been possible.

I would like to thank my major advisor Dr. Anish Malladi who shares a great passion for plants, particularly, for apple. It has been a wonderful journey working with you. I am very grateful to you for sharing your knowledge of plant science in classes. I am especially thankful for your time and efforts you have provided for this project. Thank you for your patient guidance and providing all the best resources a graduate student could ever ask for. Thank you for encouraging and always being optimistic even if the experiments are failed. Thank you for understanding and helping me through the difficulties studying in United States as an international student. I have enjoyed every discussion with you about research and anything else and I will always cherish the time I worked with you. I look forward to many collaborations with you in the future.

Thank you to my committee member, Dr. Savithri Nambeesan for helping me about many matters, both professional and personal, Dr. Dayton Wilde for providing whatever is missing in our lab and never turning me away from your office when I had questions, Dr. Wolfgang Lukowitz for your patient help on the Arabidopsis transformation work, Dr. Chang Hyun Khang for sharing your experience of organizing literature and being persistent pursuing science. Your perspective and professionalism have been invaluable to me. Thank you to Dr. Ann Loraine at UNC for your patient training on RNA-seq data analysis.

V

Special thanks to the lab technique John Doyle, who is always there to help us, and comes up with great ideas to make our experiments easier.

Thank you to the undergraduate student Easter Perdido, who made all the excellent apple sections for two years.

Thank you to Winnie, Afia, Yi-Wen, Alex for friendship and making the graduate school life happier.

Thanks to the Department of Horticulture for financial assist during my doctoral study, Dr. Marc Van Iersel and Dr. Douglas Bailey for supporting my research, Mary Jane Scarborough and other staff for assisting in the administrative matters.

Thanks to my husband for continuing to love and support me.

Thank you to my parents for pushing me work hard in school since a young age, and for your sacrifices and efforts for providing us better education.

Thank you to Athens horticulture friends and all my friends back in China who helped, listened, encouraged whenever I needed.

Thank you God for all his blessings.

# TABLE OF CONTENTS

Page
ACKNOWLEDGEMENTSv
CHAPTER
1 INTRODUCTION AND LITERATURE REVIEW1
2 TISSUE SPECIFIC METABOLIC PROFILING REVEALS DISTINCT CARBON
METABOLISM DURING APPLE FRUIT DEVELOPMENT
3 EXPRESSION PROFILING OF PUTATIVE ORGAN SIZE REGULATORY
GENES DURING PRE-BLOOM, FRUIT SET AND FRUIT DEVELOPMENT IN APPLE70
4 TRANSCRIPTIONAL PROFILING REVEALS GENE NETWORKS
REGULATING APPLE FRUIT DEVELOPMENT
5 CONCLUSIONS

#### CHAPTER 1

#### INTRODUCTION AND LITERATURE REVIEW

Apple (*Malus* × *domestica Borkh*), is one of the most consumed fresh fruits in the United States, and also one of the most widely grown fruits around the world (USDA-NASS, 2013). As in all fruit crops, fruit size is one of the most important traits directly associated with fruit quality. It is also a trait of significant biological significance. Improving apple fruit size has been always of great interest to both growers and researchers. Cultivars consistently producing large fruit size are preferred by growers for economic benefits. The understanding of fruit growth and final size regulation has been greatly improved by progress in genomics, transcriptomics, proteomics and metabolomics recently. It is the aim of this review to summarize some of the key mechanisms in the regulation of fruit growth in model species and also in apple.

#### Apple fruit set and development

The juvenile phase of apple may last for 5-12 years and involves extensive vegetative growth prior to entering into the adult, reproductive phase (Yamagishi *et al.*, 2014). Floral buds in apple are borne on branches that are two years or older and are initiated during the previous year. Following bud break, there is extensive cell production inside the bud until around 1 week before bloom. The cessation of cell production has been observed in apple and is reinitiated upon fertilization (Malladi and Johnson, 2011). The cessation of cell production and growth before bloom may be an adaptation to reduce consumption of resources prior to fruit set and has been noted in other fruit

species such as tomato. Apple flowers develop in clusters, each inflorescence includes a king or apical flower and four lateral flowers (Botton *et al.*, 2011). In king flower dominant cultivars such as 'Golden Delicious', the king flower located in the center opens first, followed by the lateral flowers blooming in sequence (Ackerman and Samach, 2015). Under optimum conditions, more than one flower within the inflorescence can be pollinated. Multiple hormones including auxins, gibberellins, ethylene, and ABA interact with each other, resulting in the abscission of unfertilized flowers and resumption of cell production in the fertilized flowers (Bangerth, 2000). Although many fruit drop naturally after blooming due to unsuccessful pollination or competition for resources, a large number of fruit still need to be removed through chemical or manual fruit load reduction in commercial fruit production (Looney, 1985). When performed early in the season, fruit load reduction can reduce competition among fruit for limited resources. This can lead to increase in size of the remaining fruit. Further, fruit load reduction can also reduce the alternate bearing effect in some cultivars that are prone to display cycles of 'on' and 'off' years on flower and fruit production (Jonkers, 1979).

Apple fruit has a distinct fruit composition where the fleshy part of the fruit is the accessory fruit derived from the non-ovarian tissue. The nature of the true fruit in apple is debated (Pratt, 1988). According to the receptacular hypothesis, the true fruit consists of five drupe-like structures surrounded immediately by fleshy tissue. Further, the accessory tissue outside the vascular ring of five sepal and five petal vascular traces is an extension of the cortical region peripheral to the vascular tissue within the stem and is referred as the cortex. The tissue surrounding the five drupe-like structures and present inside the vascular ring is referred as the pith (MacDaniels, 1940). According to the appendicular hypothesis, the tissue peripheral to the vascular ring is constituted by the fused basal regions of the appendages: sepals, petals and stamen. Fleshy tissue inside of the

vascular ring is considered of ovarian origin (exocarp and mesocarp) with the cartilaginous tissue surrounding the seeds serving as the endocarp tissue. During fruit growth and development, more extensive growth occurs in the cortex compared to the pith (Tukey and Young, 1942). The cortex may constitute greater than 80% of the fruit volume at maturity. In mature fruit, cell size is largely different between the pith and cortex (Bain and Robertson, 1950). Further, development of void space is substantially different between the two fleshy tissue types (Herremans *et al.*, 2015). While porosity within the cortex can approach around 25% at maturity, this is substantially lower at around 15% within the pith tissue, further indicting differences in growth patterns between these two major fleshy tissues of the fruit.

#### Regulation of apple fruit growth by cell production related genes

Cell cycle regulation is conserved in all eukaryotes and includes the Gap 1 phase (G1), Synthesis phase (S), Gap 2 phase (G2) and Mitosis phase (Fowler *et al.*, 1998). Transitions from G1-to-S and G2-to-M are the two of the main checkpoints in cell cycle progression (Dewitte and Murray, 2003). Regulation of the cell cycle is facilitated by several conserved proteins including cyclins (CYCs), cyclin-dependent kinases (CDKs), cyclin-dependent kinase inhibitors (CKIs)/Kip-related protein (KRPs), retinoblastoma (RB)/retinoblastoma-related proteins(RBR), E2F transcription factors (Inzé and De Veylder, 2006). The CYC family proteins control the progression of cells through the cell cycle in association with CDKs. The CDK-CYC complexes can be inhibited by KRPs that respond to the growth-inhibiting signals (De Veylder *et al.*, 2007). In the plant genome, different CDK-CYC complex is activated during different phases of the cell cycle. For example, A-type of CDKs and CYCD are expressed during the S phase and are necessary for the initiation of DNA duplication, while B-type CDKs and CYCB are expressed during the G2-to-M transition

and promote the events of mitosis in Arabidopsis (De Veylder *et al.*, 2007). In the apple genome, 14 genes including *A2*, *B1*, *B2 CYCs*, and *CDKB*, a *WEE kinase (MdWEE1*) and an E2F transcription factor (*MdDEL1*) were found to display positive association with cell production across various stages of flower and fruit development (Malladi and Johnson, 2011). Treatments that alter cell production during fruit growth often involve changes in transcript abundance of cell cycle genes (Dash *et al.*, 2012, 2013). For example, transcript abundance of an A2 type cyclin transcript declined in shaded fruit by up to 4.6-fold in association with a decline in cell production and growth (Dash *et al.*, 2012).

In addition, five *MdKRPs* were negatively associated with cell production during apple fruit development. Both MdKRP4 and MdKRP5 displayed high transcript abundance prior to bloom, in non-pollinated flowers, and during the transition period from cell production to expansion, indicating their negative role in regulating cell production (Malladi and Johnson, 2011). The expression level of *MdKRP4* and *MdKRP5* was also elevated in shaded fruit further indicating the role as negative regulators of cell production in these fruit (Dash *et al.*, 2012). In addition to the cell cycle genes, transcription factors potentially regulating cell production have been identified in apple. Two apple AINTEGUMENTA genes, MdANT1 and MdANT2, may affect apple fruit growth by regulating cell production. The transcript abundance of the MdANTs is not only positively regulated during cell production, it is also correlated with that of the positive regulators of cell production such as MdCYCA2;3, MdCYCB1;1, MdCDKB1;2, MdCDKB2;1 and MdDEL1. The transcript abundance of the cell division repressor, MdKRP4, starts to decline in the cortex, while displaying a gradual increase in the pith tissue (Dash and Malladi, 2012), further suggesting that spatial differences in cell production may also be associated with differential transcript abundance of these key cell cycle regulators.

#### Regulation of apple fruit growth by cell expansion related genes

Much fewer genes have been associated with cell expansion-mediated regulation of fruit growth. During the cell expansion period, many cell wall modifying proteins mediate cell wall loosening and allow cell growth and expansion such as endo-1,4-β-D-glucanases (EGases), xyloglucan endotransglycosylases (XETs) and expansins (EXPs). In the apple genome, the transcript abundance of several  $\alpha$ -type expansin (*MdEXPA*) genes displayed potential association with cell expansion during apple fruit development (Dash et al., 2013). The transcript level of MdEXPA10;1 decreased in shaded fruit and was associated with a reduction in cell expansion (Dash *et al.*, 2012). COBRA(COB) and COBRA-LIKE(COBL) genes encode putative Glycosyl-Phosphatidyl Inositol (GPI) anchored proteins at the plasma membrane-cell wall interface, which were first identified as key regulators of oriented cell expansion in Arabidopsis (Roudier et al., 2005; Li et al., 2013). The transcript abundance of *MdCOB1* remains low during the cell production period, and increases greatly during the transition from cell production to cell expansion, suggesting its role in upregulating cell expansion in apple fruit development. Decreased transcript levels of *MdCOB1* in response to shading also suggest that *MdCOB1* may contribute to the decline in cell expansion (Dash et al., 2012, 2013).

#### Key regulators of organ growth and development in other species

Organ size is determined by two partially overlapping phases, cell production and cell expansion (Horiguchi *et al.*, 2006). A large number of factors regulating organ growth have now been described, and connections involving different regulators are starting to emerge from various studies. For example, *fruit weight (FW)2.2* is a quantitative trait locus on chromosome 2

responsible for approximately 30% of the fruit weight variation between domesticated tomatoes and their small-fruited wild relatives (Frary *et al.*, 2000). QTL mapping of more species in the *Solanaceae* family indicated that the role of *FW2.2* was conserved in eggplant (*Solanum melongena*) and pepper (*Capsium annuum*) (Chaim *et al.*, 2001; Guo and Simmons, 2011). The gene underlying *FW2.2*, was identified in maize and named as *Cell Number Regulator* (*CNR*) owing to its function in inhibiting cell production. Over expression of *ZmCNR1* reduced overall plant size and that of multiple organs, confirming its role as a cell production inhibitor (Guo *et al.*, 2010; Guo and Simmons, 2011). *CNR* family genes have also been identified in peach (*Prunus persica*) and sweet cherry (*Prunus avium*). Two of these in sweet cherry were found to co-localize with two fruit size QTLs (De Franceschi *et al.*, 2013).

In Arabidopsis, a regulatory cascade involving *miR319-TCP4-miR396-GRF/GIF* has been proposed to control leaf growth (Gonzalez *et al.*, 2012). *Growth Regulating Factors* (*GRFs*) belong to a family encoding a class of plant specific transcription factors. The GRF family display a wide range of functions including regulation of leaf, stem and floral organ development, flowering time, seed and root development, control of growth under adverse environmental conditions, and plant longevity (Frary *et al.*, 2000; Kim *et al.*, 2003; Kim and Lee, 2006; Liu *et al.*, 2014). In Arabidopsis, *GRF* triple mutants, *grf1*/2/3, display smaller and narrower leaves, while larger leaves and cotyledons are formed through overexpression of *AtGRF1* or *AtGRF2* (Kim *et al.*, 2003; Kim and Kende, 2004). GRFs do not function independently, rather as a transcription factor complex formed with the co-activators, GRF-interacting factors (GIFs). The *Arabidopsis* triple mutant, *gif1*/2/3 display extremely small plant size compared to single mutants (Kim and Lee, 2006; Lee *et al.*, 2009). In Arabidopsis, six *GRFs* transcripts are targeted by miR396. Overexpression of miR396 antagonizes the expression pattern of GRFs (Liu *et al.*, 2014). Expression of miR396 is regulated by а transcription factor: **TEOSINTE** BRANCHED1/CYCLOIDEA/PCF4 (TCP4) (Rodriguez et al., 2010). TCP4 directly binds to the promoter of miR396 and promotes its expression (Schommer et al., 2014). Increased activity of the transcription factor TCP4 inhibited cell production in Arabidopsis leaves through higher miR396 and lower GRF abundance (Liu et al., 2009; Rodriguez et al., 2010). In Arabidopsis, expression of miR396 resistant versions GRFs resulted in enhanced leaf growth (Rodriguez et al., 2010; Schommer et al., 2014). The expression of TCP4 is further regulated by another microRNA, miR319 (Palatnik et al., 2003). A point mutation at the target site of miR319 within TCP4, or the overexpression of TCP4 leads to the formation of smaller leaves (Rodriguez et al., 2010; Kalve et al., 2014). In addition, TCP4 may also directly bind to the promoter of KRP1 in Arabidopsis indicating an additional mechanism by which it can regulate cell production (Schommer et al., 2014).

Deep small RNA-seq identified both *miR396* and *miR319* in apple genome, and the expression level of *miR396* is much lower in young fruit at 15 days after anthesis than mature fruit. Inversely, miR319 expression level is higher in the young fruit than mature fruit (Xia *et al.*, 2012). Multiple apple *TCP* genes have been identified (Mimida *et al.*, 2011). MdTCP2 and MdTCP4 were demonstrated to interact with FLOWERING LOCUS T 1(MdFT1) and MdFT2 proteins to regulate flowering (Mimida *et al.*, 2011). In addition, target sites of *miR319* within the *TCP* genes and of *miR396* within the *GRFs* have been confirmed in apple, implicating this cascade in the regulation of cell production in apple.

Organ growth is closely regulated by phytohormones. *AUXIN-REGULATED GENE INVOLVED IN ORGAN SIZE (ARGOS)* is highly induced by auxin which plays a positively regulatory role of cell division in Arabidopsis (Hu *et al.*, 2003). Transgenic Arabidopsis expressing

7

greater or lower *ARGOS* display enlarged or reduced aerial organs, respectively, due to changes in cell number (Hu *et al.*, 2003). *ARGOS* may act upstream of the transcription factor *AINTEGUMENTA* (*ANT*) since the loss of function of *ANT* can block *ARGOS* effects on organ development (Hu *et al.*, 2003). In Arabidopsis, ANTs have been shown to be a positive regulator of cell production that plays an important role regulating organ growth. The apple orthologous ANT genes display a similar expression pattern as several cell cycle genes during apple fruit development, indicating a similar role as positive regulator of cell production. Another gene with sequence similar to *ARGOS*, *ARGOS-LIKE* (*ARL*), is also identified in Arabidopsis as a potential organ growth regulatory gene. It is suggested that *ARLs* are involved in the regulation of cell expansion rather than cell division during organ growth. Transcript levels of *ARL* increased by 2.2-fold with the treatment of epi-brassinolide, with a slight change induced by auxin or cytokinin (Hu *et al.*, 2006).

A mobile plant organ growth stimulator distinct from classical hormones has been described in *Arabidopsis*, and involves the activity of the *KLUH* gene (Eriksson *et al.*, 2010). It was proposed that a threshold for cell division signal is created by the KLUH protein: cell production ceases if the signal level within the cell is above the threshold level; whereas the cell production continues if the signal level is below its threshold. The Arabidopsis *kluh* mutant developed thinner stems, smaller flowers and leaves, lighter and smaller seeds than the wild type. Inversely, the over-expression of *KLUH* resulted in overgrowth of the organ due to cell number alteration (Anastasiou *et al.*, 2007; Adamski *et al.*, 2009). *KLUH* was also identified as the gene corresponding to a QTL associated with fruit size regulation in tomato (Chakrabarti *et al.*, 2013), indicating that this gene may play conserved roles across species in regulating organ growth.

#### Metabolic regulation underlying fruit growth

Apple fruit is predominantly heterotrophic and relies on translocated resources from source organs. The fruit not only compete with each other for C, N and other resources, but also with other sink organs such as newly growing leaves, since the growth of the first five to six leaves in apple is largely supported by reserves. As the leaves mature, they become source tissues exporting resources to the nearby fruit (Hansen, 1971). Early fruit development mediated by cell production creates great demands for resources owing to rapid cell growth, cell wall synthesis, and the energy requirements associated with these processes. Respiration rates are in fact highest during early fruit development (Blanke and Lenz, 1989). Post-mitotic cell expansion involves extensive enlargement of vacuolar volume with intake of water, which is coincident with the transitory starch accumulation. Owing to the vast differences in mechanisms that facilitate growth during different phases of fruit development, it is likely that fruit metabolic programs are also substantially different. In fact, glycolysis has been suggested as an important aspect of C metabolism during early fruit growth but may have limited significance at later stages (Beshir *et al.*, 2017).

Like many other fruits in the *Rosaceae* family, sorbitol and sucrose are the major photosynthetic products and translocated carbohydrates in apple (Webb and Burley, 1962). Sorbitol is synthesized in source tissue via reduction of glucose-6-phosphate (G6P) to sorbitol-6phosphate (S6P) by sorbitol-6-phosphate dehydrogenase (S6PDH). After translocation to fruit through sorbitol transporters (SOT), sorbitol is rapidly converted to fructose by sorbitol dehydrogenase (SDH). The concentration of sorbitol in apple fruit increases over two fold at 10 days after the fruit load reduction, suggesting more sorbitol is partitioned into the remaining fruits after fruit load reduction treatment (Dash *et al.*, 2013). Nine SDH genes have been identified in the apple genome, and five of them are expressed in the cortex during fruit development (Nosarzewski and Archbold, 2007). Apple seeds display higher SDH activity than the cortex during early fruit development (Nosarzewski and Archbold, 2007). The transcript abundance of MdSDH1 and MdSDH2 gene is elevated at 3 days after shading treatment, indicating rapid response of SDH activity to compensate for the decreased photosynthetic resource partitioning to the fruit (Dash et al., 2012). The Suc-Suc cycle, was reported to regulate sucrose metabolism and accumulation in apple fruit (Li et al., 2012). In this system, sucrose is either converted to fructose and glucose by cell wall invertase (CWINV), or travels into the cytosol via sucrose transporter (SUT), and is converted to fructose and glucose by neutral invertase (NINV), or to fructose and UDP-glucose by sucrose synthase (SUSY) inside the cell. Fructose and glucose are phosphorylated to G6P and F6P by hexokinase (HXK) and fructokinase (FK), respectively. F6P then enters glycolysis and TCA cycle to generate energy that is required during fruit development, while G1P is used to generate starch as the storage form of C. F6P and UDPG react with each other and can re-synthesize sucrose through sucrose phosphate synthase (SPS) activity. Substantial shifts in sugar concentration and activity of sugar metabolism related genes have been described during apple fruit development (Zhang et al., 2010; Li et al., 2012). The sorbitol concentration declines during apple fruit development, indicating it is quickly metabolized once translocated into fruit. The sucrose concentration, on the hand, increases during fruit development, suggesting substantial Suc-Suc cycle activity to generate sucrose. At maturity, fructose is the most abundant sugar in apple fruit, followed by sucrose and glucose. The enzymatic activity of SDH, CwINV, NINV, SUSY, and FK all declined during fruit development, while the SPS increased continuously, indicating the shifts in metabolite concentration is mediated by changes in the activity of sugar metabolism related enzymes (Zhang et al., 2010).

Malic acid and quinic acid are the two major organic acids in apple. The concentration of both acids increases during early fruit development, but decreases continuously afterward until maturation (Zhang et al., 2010). The extremely high level of organic acids during early fruit development may not only add sourness to discourage consumption before seed maturation, they also indicate high aerobic cellular respiration during early fruit development as acids such as malate are intermediates of the TCA cycle (Blanke and Lenz, 1989). Malic acid synthesis starts from the fixation of bicarbonate  $(HCO_3)$  by phosphoenolpyruvate carboxylase (PEPC) to generate oxaloacetic acid (OAA), and OAA is converted to malate by malate dehydrogenase (MDH) (Blanke and Lenz, 1989). The degradation of malate during later fruit development is mediated by cytosolic NADP-dependent malic enzyme (NADP-ME). One cytosolic MDH (cyMDH) and three PEPC genes were identified previously in apple (Shangguan et al., 2015). The overexpression of MdcyMDH contributed to malate accumulation in the apple callus (Yao et al., 2011). The enzymatic activity of PEPC is relatively high around 40 days after full bloom, and starts to decline abruptly afterward. On the opposite, the activity of NAD-MDH increases slightly during apple fruit development (Li et al., 2013).

Transitory accumulation of starch occurs during mid fruit development in apple (Berüter, 1985; Zhang *et al.*, 2010). The rate limiting step for starch synthesis is the conversion from Glc-1-phosphate (G1P) to ADP-Glc by ADP-glucose pyrophosphorylase (AGPase). Other enzymes involved in starch synthesis include starch branching enzyme (SBE) contributing to branching of the glucan chain, and starch synthase (SS), an enzyme involved in the addition of the Glc residue from ADP-Glc to the glucan chain. The transcript abundance of both *MdAGPase* and *MdSS* is elevated slightly before the accumulation of starch, and declines or remained stable as starch concentration declines during later fruit development (Janssen *et al.*, 2008). An *MdSBE* gene also displays increased transcript abundance during apple fruit development (Han *et al.*, 2007).

The main entry forms of N into fruit are predominately asparagine (Asn), aspartate (Asp), glutamine (Gln) and arginine (Tromp and Ovaa, 1971). Asn is the most abundant amino acid during early fruit development in apple, and its level decreases during mid and late fruit development (Zhang *et al.*, 2010). Within the fruit, metabolism of Asn is not characterized but may be likely metabolized by asparaginase (ASPA) releasing aspartate (Asp) and ammonium (Gaufichon *et al.*, 2016). Asp is subsequently used as the N source for the biosynthesis of other amino acids, while ammonium may be re-assimilated by glutamine synthetase (GS) using glutamate to yield glutamine.

Several aspects of fruit metabolism are not well characterized in apple: 1. Temporal changes in fruit metabolism are known in other fruit species such as tomato (Carrari *et al.*, 2006). However, this information is not well characterized in apple. In spite of the abundance of information on changes in metabolite concentrations during mid and later fruit development, relatively little is known regarding this aspect during early fruit development. 2. Spatially characteristic metabolic programs may be expected owing to the previously described distinct fruit anatomy of the apple fruit. However, such spatial differences in fruit metabolism have not yet been explored. 3. Fruit load reduction enhances growth and final size. However, relatively few studies have explored the changes in fruit metabolism associated with enhanced growth. 4. Finally, integration of fruit metabolism with growth in a spatio-temporal context is essential to better understand the interactions between these processes but has not yet been explored.

#### Significance and hypothesis

Despite the progress made in understanding organ growth and development in model plant species, there exists a significant gap in knowledge of the regulation of these processes in economically important crops, especially tree fruits such as apple. The unique fruit structure in apple involving tissues of diverse origins presents an opportunity to better characterize spatio-temporal characteristics of growth, development and metabolism. *The goal of this study is to understand fruit growth in apple in a spatial and temporal context.* The specific objectives of this studies include:

1. To understand tissue and stage specific metabolic regulation of apple fruit growth

2. To identify and characterize organ size regulatory genes in apple

3. To investigate tissue specific transcriptome-wide changes during apple fruit development and identify potential genes and networks regulating fruit growth.

### References

Ackerman, M. and A. Samach. 2015. Doubts regarding carbohydrate shortage as a trigger toward abscission of specific Apple (Malus domestica) fruitlets. <u>New Negatives in Plant Science</u> 1: 46-52.

Adamski, N. M., E. Anastasiou, S. Eriksson, C. M. O'Neill and M. Lenhard. 2009. Local maternal control of seed size by KLUH/CYP78A5-dependent growth signaling. <u>Proceedings of the National Academy of Sciences</u> **106**: 20115-20120.

Anastasiou, E., S. Kenz, M. Gerstung, D. MacLean, J. Timmer, C. Fleck and M. Lenhard. 2007. Control of plant organ size by KLUH/CYP78A5-dependent intercellular signaling. Developmental cell 13: 843-856.

**Bain, J. M. and R. N. Robertson.** 1950. The physiology of growth in apple fruits I. Cell size, cell number, and fruit development. <u>Australian Journal of Biological Sciences</u> **4**: 75-91.

**Bangerth, F.** 2000. Abscission and thinning of young fruit and thier regulation by plant hormones and bioregulators. <u>Plant Growth Regulation</u> **31**: 43-59.

**Berüter, J.** 1985. Sugar accumulation and changes in the activities of related enzymes during development of the apple fruit. Journal of Plant physiology **121**: 331-341.

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

Botton, A., G. Eccher, C. Forcato, A. Ferrarini, M. Begheldo, M. Zermiani, S. Moscatello, A. Battistelli, R. Velasco and B. Ruperti. 2011. Signaling pathways mediating the induction of apple fruitlet abscission. <u>Plant physiology</u> 155: 185-208.

Chaim, A. B., I. Paran, R. Grube, M. Jahn, R. Van Wijk and J. Peleman. 2001. QTL mapping of fruit-related traits in pepper (*Capsicum annuum*). <u>Theoretical and Applied Genetics</u> **102**: 1016-1028.

**Dash, M., L. K. Johnson and A. Malladi**. 2012. Severe shading reduces early fruit growth in apple by decreasing cell production and expansion. Journal of the American Society for <u>Horticultural Science</u> **137**: 275-282.

**Dash, M., L. K. Johnson and A. Malladi.** 2013. Reduction of fruit load affects early fruit growth in apple by enhancing carbohydrate availability, altering the expression of cell production-related genes, and increasing cell production. Journal of the American Society for Horticultural Science 138: 253-262.

**Dash, M. and A. Malladi.** 2012. The AINTEGUMENTA genes, MdANT1 and MdANT2, are associated with the regulation of cell production during fruit growth in apple (*Malus* × *domestica Borkh*.). <u>BMC plant biology</u> **12**: 98.

**De Franceschi, P., T. Stegmeir, A. Cabrera, E. Van Der Knaap, U. Rosyara, A. Sebolt, L. Dondini, E. Dirlewanger, J. Quero-Garcia and J. Campoy**. 2013. Cell number regulator genes in Prunus provide candidate genes for the control of fruit size in sweet and sour cherry. <u>Molecular Breeding</u> **32**: 311-326.

**De Veylder, L., T. Beeckman and D. Inzé.** 2007. The ins and outs of the plant cell cycle. Nature Reviews Molecular Cell Biology **8**: 655-665.

**Dewitte, W. and J. A. Murray.** 2003. The plant cell cycle. <u>Annual review of plant biology</u> **54**: 235-264.

**Eriksson, S., L. Stransfeld, N. M. Adamski, H. Breuninger and M. Lenhard.** 2010. KLUH/CYP78A5-dependent growth signaling coordinates floral organ growth in Arabidopsis. <u>Current biology</u> **20**: 527-532.

Fowler, M., S. Eyre, N. Scott, A. Slater and M. Elliott. 1998. The plant cell cycle in context. Molecular biotechnology 10: 123-153.

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

Gaufichon, L., A. Marmagne, T. Yoneyama, T. Hase, G. Clément, M. Trassaert, X. Xu, M. Shakibaei, A. Najihi and A. Suzuki. 2016. Impact of the disruption of ASN3-encoding asparagine synthetase on Arabidopsis development. <u>Agronomy</u> **6**: 12.

**Gonzalez, N., H. Vanhaeren and D. Inzé.** 2012. Leaf size control: complex coordination of cell division and expansion. <u>Trends in plant science</u> **17**: 332-340.

Guo, M., M. A. Rupe, J. A. Dieter, J. Zou, D. Spielbauer, K. E. Duncan, R. J. Howard, Z. Hou and C. R. Simmons. 2010. Cell Number Regulator1 affects plant and organ size in maize: implications for crop yield enhancement and heterosis. <u>The Plant Cell</u> **22**: 1057-1073.

**Guo, M. and C. R. Simmons.** 2011. Cell number counts–the fw2. 2 and CNR genes and implications for controlling plant fruit and organ size. <u>Plant Science</u> **181**: 1-7.

Han, Y., K. Gasic, F. Sun, M. Xu and S. S. Korban. 2007. A gene encoding starch branching enzyme I (SBEI) in apple (*Malus* × *domestica*) and its phylogenetic relationship to Sbe genes from other angiosperms. <u>Molecular phylogenetics and evolution</u> **43**: 852-863.

**HANSEN, P.** 1971. <sup>14</sup>C-studies on apple trees. VII. The early seasonal growth in leaves, flowers and shoots as dependent upon current photosynthates and existing reserves. <u>Physiologia</u> <u>Plantarum</u> **25**: 469-473.

Herremans, E., P. Verboven, M. Hertog, D. Cantre, M. van Dael, T. De Schryver, L. Van Hoorebeke and B. Nicolai. 2015. Spatial development of transport structures in apple (Malus× domestica Borkh.) fruit. Frontiers in Plant Science 6: 679.

Horiguchi, G., A. Ferjani, U. Fujikura and H. Tsukaya. 2006. Coordination of cell proliferation and cell expansion in the control of leaf size in Arabidopsis thaliana. Journal of plant research 119: 37-42.

Hu, Y., H. M. Poh and N. H. Chua. 2006. The Arabidopsis ARGOS-LIKE gene regulates cell expansion during organ growth. <u>The Plant Journal</u> **47**: 1-9.

Hu, Y. X., O. Xie and N. H. Chua. 2003. The Arabidopsis auxin-inducible gene ARGOS controls lateral organ size. <u>Plant Cell</u> 15: 1951-1961.

Inzé, D. and L. De Veylder. 2006. Cell cycle regulation in plant development 1. <u>Annu. Rev.</u> <u>Genet.</u> 40: 77-105.

Janssen, B. J., K. Thodey, R. J. Schaffer, R. Alba, L. Balakrishnan, R. Bishop, J. H. Bowen, R. N. Crowhurst, A. P. Gleave and S. Ledger. 2008. Global gene expression analysis of apple fruit development from the floral bud to ripe fruit. <u>BMC Plant Biology</u> **8**: 16.

**Jonkers, H.** 1979. Biennial bearing in apple and pear: a literature survey. <u>Scientia Horticulturae</u> **11**: 303-317.

Kalve, S., D. De Vos and G. T. Beemster. 2014. Leaf development: a cellular perspective. <u>Frontiers in plant science</u> **5**: 362.

Kim, J. H., D. Choi and H. Kende. 2003. The AtGRF family of putative transcription factors is involved in leaf and cotyledon growth in Arabidopsis. <u>The Plant Journal</u> **36**: 94-104.

**Kim, J. H. and H. Kende.** 2004. A transcriptional coactivator, AtGIF1, is involved in regulating leaf growth and morphology in Arabidopsis. <u>Proceedings of the National Academy of Sciences of the United States of America</u> **101**: 13374-13379.

**Kim, J. H. and B. H. Lee.** 2006. GROWTH-REGULATING FACTOR4 of Arabidopsis thaliana is required for development of leaves, cotyledons, and shoot apical meristem. <u>Journal of Plant</u> <u>Biology</u> **49**: 463-468.

Lee, B. H., J.-H. Ko, S. Lee, Y. Lee, J.-H. Pak and J. H. Kim. 2009. The Arabidopsis GRF-INTERACTING FACTOR gene family performs an overlapping function in determining organ size as well as multiple developmental properties. <u>Plant physiology</u> **151**: 655-668.

Li, M., F. Feng and L. Cheng. 2012. Expression patterns of genes involved in sugar metabolism and accumulation during apple fruit development. <u>PloS one</u> 7: e33055.

Li, P., F. Ma and L. Cheng. 2013. Primary and secondary metabolism in the sun-exposed peel and the shaded peel of apple fruit. <u>Physiologia plantarum</u> 148: 9-24.

Li, S., F. R. Ge, M. Xu, X. Y. Zhao, G. Q. Huang, L. Z. Zhou, J. G. Wang, A. Kombrink, S. McCormick and X. S. Zhang. 2013. Arabidopsis COBRA-LIKE 10, a GPI-anchored protein, mediates directional growth of pollen tubes. <u>The Plant Journal</u> 74: 486-497.

Liu, D., Y. Song, Z. Chen and D. Yu. 2009. Ectopic expression of miR396 suppresses GRF target gene expression and alters leaf growth in Arabidopsis. <u>Physiologia plantarum</u> 136: 223-236.

Liu, H., S. Guo, Y. Xu, C. Li, Z. Zhang, D. Zhang, S. Xu, C. Zhang and K. Chong. 2014. OsmiR396d-regulated OsGRFs function in floral organogenesis in rice through binding to their targets OsJMJ706 and OsCR4. <u>Plant physiology</u> **165**: 160-174.

**Looney**, N. E. 1985. Chemical thinning of apple: some new strategies and important refinements to old procedures. V International Symposium on Growth Regulators in Fruit Production 179.

**MacDaniels, L. H.** 1940. The morphology of the apple and other pome fruits. <u>Memoirs. Cornell</u> <u>Agricultural Experiment Station 230</u>.

**Malladi, A. and P. M. Hirst.** 2010. Increase in fruit size of a spontaneous mutant of 'Gala' apple (Malus x domestica Borkh.) is facilitated by altered cell production and enhanced cell size. <u>J Exp</u> <u>Bot</u> **61**: 3003-3013.

**Malladi, A. and L. K. Johnson.** 2011. Expression profiling of cell cycle genes reveals key facilitators of cell production during carpel development, fruit set, and fruit growth in apple (Malus× domestica Borkh.). Journal of experimental botany **62**: 205-219.

Mimida, N., S.-I. Kidou, H. Iwanami, S. Moriya, K. Abe, C. Voogd, E. Varkonyi-Gasic and N. Kotoda. 2011. Apple FLOWERING LOCUS T proteins interact with transcription factors implicated in cell growth and organ development. <u>Tree Physiology</u> **31**: 555-566.

**Nosarzewski, M. and D. D. Archbold.** 2007. Tissue-specific expression of sorbitol dehydrogenase in apple fruit during early development. Journal of experimental botany **58**: 1863-1872.

Palatnik, J. F., E. Allen, X. Wu, C. Schommer, R. Schwab, J. C. Carrington and D. Weigel. 2003. Control of leaf morphogenesis by microRNAs. <u>Nature</u> **425**: 257.

**Pratt, C.** 1988. Apple flower and fruit: morphology and anatomy. <u>Horticultural Reviews</u> **10**: 273-308.

Rodriguez, R. E., M. A. Mecchia, J. M. Debernardi, C. Schommer, D. Weigel and J. F. Palatnik. 2010. Control of cell proliferation in Arabidopsis thaliana by microRNA miR396. Development 137: 103-112. Roudier, F., A. G. Fernandez, M. Fujita, R. Himmelspach, G. H. Borner, G. Schindelman, S. Song, T. I. Baskin, P. Dupree and G. O. Wasteneys. 2005. COBRA, an Arabidopsis extracellular glycosyl-phosphatidyl inositol-anchored protein, specifically controls highly anisotropic expansion through its involvement in cellulose microfibril orientation. <u>The Plant Cell</u> **17**: 1749-1763.

Shangguan, L., X. Sun, C. Zhang, Q. Mu, X. Leng and J. Fang. 2015. Genome identification and analysis of genes encoding the key enzymes involved in organic acid biosynthesis pathway in apple, grape, and sweet orange. <u>Scientia Horticulturae</u> 185: 22-28.

**Tromp, J. and J. Ovaa.** 1971. Phloem translocation of storage nitrogen in apple. <u>Physiologia</u> <u>Plantarum</u> **25**: 407-413.

**Tukey, H. and J. O. Young.** 1942. Gross morphology and histology of developing fruit of the apple. <u>Botanical Gazette</u> **104**: 3-25.

Webb, K. L. and J. W. A. Burley. 1962. Sorbitol translocation in apple. Science 137: 766-766.

Xia, R., H. Zhu, Y.-q. An, E. P. Beers and Z. Liu. 2012. Apple miRNAs and tasiRNAs with novel regulatory networks. <u>Genome biology</u> 13: R47.

**Yamagishi, N., R. Kishigami and N. Yoshikawa.** 2014. Reduced generation time of apple seedlings to within a year by means of a plant virus vector: a new plant-breeding technique with no transmission of genetic modification to the next generation. <u>Plant biotechnology journal</u> **12**: 60-68.

**Yao, Y.-X., M. Li, H. Zhai, C.-X. You and Y.-J. Hao.** 2011. Isolation and characterization of an apple cytosolic malate dehydrogenase gene reveal its function in malate synthesis. <u>Journal of plant physiology</u> **168**: 474-480.

**Zhang, Y., P. Li and L. Cheng.** 2010. Developmental changes of carbohydrates, organic acids, amino acids, and phenolic compounds in 'Honeycrisp'apple flesh. <u>Food Chemistry</u> **123**: 1013-1018.

# CHAPTER 2

# TISSUE SPECIFIC METABOLIC PROFILING REVEALS DISTINCT CARBON METABOLISM DURING APPLE FRUIT DEVELOPMENT

Jing, S. and Malladi, A. To be submitted to Journal of Experimental Botany.

#### Abstract

Apple fruit (*Malus* × *domestica* Borkh) consist of cortex and pith tissues with differential growth and developmental features. Apple fruit development relies on the reserved C and N resources, especially during early fruit development. The spatiotemporal C and N metabolic regulation of apple fruit development is investigated in this study. Fruit load reduction resulted to larger fruits due to enhanced growth in the cortex specifically. Targeted metabolic profiling revealed spatial characteristics of fruit metabolism. The main translocated forms of C and N, Sor and Asn, both displayed higher concentration in the pith than the cortex, especially during early fruit development. The C and N metabolism related genes were also differentially expressed between cortex and pith. Temporally, nine stages were clustered into early (0-26 DAT), mid (26-47 DAT) and late fruit development (47-118 DAT) indicating distinctive features of fruit metabolism across different developmental stages. In addition, Sor and Asn content was positively associated suggesting C and N resources are not independent with each other during fruit development. These results indicate characteristic C and N metabolism within the pith and cortex tissues that may explain different growth and development patterns between the two types of tissue.

**Keywords:** apple fruit, metabolic profiling, fruit development, carbon metabolism, nitrogen metabolism, tissue-specific

#### Introduction

Growth and development of fleshy fruits, such as apple, involves multiple phases: cell production, cell expansion, maturity and ripening (Gillaspy *et al.*, 1993; Giovannoni, 2004). Progression through these diverse phases requires coordinated changes in fruit metabolism, which involves resource import, metabolic inter-conversions, metabolite compartmentation, and storage. Many fleshy fruits including apple display some photosynthetic capacity but are predominantly heterotrophic (Blanke and Lenz, 1989; Carrari and Fernie, 2006). Hence, fruit metabolism is substantially dependent on import of carbon (C), and other resources such as nitrogen (N). In apple, sorbitol (Sor) and sucrose (Suc) are the major translocated C forms (Bieleski, 1982; Yamaki and Ino, 1992). Nitrogen demands of apple fruit growth are likely met through remobilization of stored reserves during early development and through new uptake during later stages (Malaguti *et al.*, 2001). Remobilized and newly acquired N is primarily translocated as amino acids with Asn (Asn) serving as the predominant form (Malaguti *et al.*, 2001; Guak *et al.*, 2004). The majority of fruit structural and non-structural components are subsequently derived from these imported resources.

Temporally dynamic metabolic programs have been characterized recently in multiple fruits (Carrari *et al.*, 2006; Moco *et al.*, 2007). These studies indicate metabolic signatures characteristic of specific developmental stages (Carrari and Fernie, 2006). Temporal patterns of apple fruit metabolism have not achieved sufficient resolution, particularly during early fruit development. Fruit metabolite content has been analyzed across multiple stages of development in apple but only a few (1 or 2) stages of early fruit development have been investigated (Zhang *et al.*, 2010; Beshir *et al.*, 2017). Early growth following bud break in apple is largely supported by remobilization of stored C, N, and other reserves (Hansen, 1971; Malaguti *et al.*, 2001). From around bloom, photosynthetic activity of current vegetative growth supports the C demands of developing fruit (Byers and Carbaugh, 1991; Bepete and Lakso, 1998). Dependence of the fruit during a part of its initial growth and development on remobilized reserves and subsequently on alternative sources is likely to involve characteristic transitions in metabolic programs. Further, metabolism during early stages may be substantially distinct from that at later stages owing to specific processes mediating growth: intensive cell production during early growth *vs* cell expansion during subsequent stages. These processes are likely supported by distinct metabolic programs. Hence, temporal analysis of fruit metabolism, particularly during early development, and investigation of its relationship to growth are essential to better understand fruit development.

Fruits are morphologically and anatomically diverse, but in all cases are constituted by multiple tissue types with differential growth and developmental features. It is likely that these tissues display characteristic metabolic programs. In tomato, comparison of mesocarp and locule tissues during the cell expansion phase of early fruit development revealed distinctly different metabolic and transcriptional programs associated with their specific growth characteristics (Mounet et al., 2009). In strawberry, achenes originating from the ovary, and fleshy fruit originating from the receptacle displayed differences in C, N, and secondary metabolism (Fait et al., 2008). In apple, majority of the fleshy part of the fruit (cortex) is derived from accessory tissue, likely from fused basal regions of sepals, petals and anthers (Pratt, 1988; Yao et al., 2016). The true fruit originating from the ovary is contained within the interior region (pith) surrounding seed locules (MacDaniels, 1940; Pratt, 1988). The pith and cortex display differences in growth during development: while they contribute similarly to flower size at bloom, the cortex constitutes around 80% of the fruit volume at maturity, suggesting preferentially intensive growth in this region (Tukey and Young, 1942). It may be expected that such preferential growth is supported by differential metabolism. Evaluation of spatial differences in metabolism during development is

essential to better understand its contribution to growth, but has not been explored in apple fruit previously.

Fruit load (the extent of fruit formed and retained) influences competition for available resources and can influence multiple fruit traits (Prudent *et al.*, 2009). Fruit growth and metabolism are dynamically altered under conditions of limited resource availability (Kromdijk *et al.*, 2014). In tomato, severe limitation of C availability altered utilization of C reserves, and N remobilization (Baldet *et al.*, 2002). In peach, altering source-sink balance altered growth, and Sor and Suc metabolism (Morandi *et al.*, 2008). In apple, reducing flower load prior to bloom resulted in substantial changes in carbohydrate storage(Berüter, 1990); altering resource availability during early fruit development transiently increased Sor and fructose (Fru) concentrations (Dash *et al.*, 2013); and altering resource availability during mid fruit development resulted in diversion of C from starch to Suc and glucose (Glc) accumulation (Berüter and Feusi, 1997), emphasizing stage-dependent effects of resource availability on metabolism. However, how resource availability affects growth and metabolism in a spatiotemporal manner remains poorly characterized.

The goal of the current study was to determine temporal, spatial, and resource availabilityrelated characteristics of apple fruit growth and metabolism. To address this objective, pith and cortex tissues of the fruit were analyzed at multiple stages of development, with emphasis on early development. Resource availability was manipulated by fruit load reduction during early development. Spatiotemporal growth patterns were determined in relation to resource availability. To identify the major features of metabolism, major sugars, sugar alcohols, organic acids, amino acid, and starch concentrations were quantified. Further, transcript accumulation of multiple genes associated with metabolism of above metabolites was quantified to gain insights into the regulation of fruit metabolism.

#### **Materials and Methods**

#### **Plant Material**

'Golden Delicious Smoothee' apple trees at the Mountain Research and Education Center, University of Georgia, Blairsville, GA maintained following commercial production practices were used in this study. In 2015, four trees were randomly selected and subjected to the fruit load reduction treatment while another four trees were untreated and used as controls (control fruit load - CL). Each tree was treated as an experimental unit. For the reduced fruit load (RL) treatment, all fruitlets except the king fruitlet within a cluster were manually removed at 11 days after full bloom (DAFB). No chemical thinning agents were applied during this study. Fruit diameter and length were measured on 10 tagged king fruit from each tree at the following stages: 0, 8, 12, 19, 26, 33, 47, 77, 118 days after treatment (DAT). At each stage, four king fruit from each tree were sampled and longitudinally cut in half. One half was fixed in CRAF III (Chromic acid: Acetic acid: formalin) fixative for morphometric analysis, and the other half was used for metabolite and transcript abundance profiling. For samples intended for the latter use, pith and cortex tissues were separated using a biopsy punch (1-10 mm), immediately frozen in liquid nitrogen, and stored at - 80 °C.

#### Measurement of fruit growth and development in pith and cortex

Images of the longitudinal fruit profile were obtained using a flatbed scanner (V600, Epson). ImageJ software (National Institutes of Health, USA) was used to outline and measure locule, core (marked by sepal and petal vascular traces) and total fruit sectional areas. Sectional area of the core was subtracted from that of the fruit to obtain cortex sectional area. Area of locule was subtracted from that of the core to obtain pith sectional area.

#### Metabolic profiling using gas chromatography (GC)

Soluble sugars, organic acids and related compounds were extracted and analyzed according to (Chapman and Horvat, 1989) with some modifications. Fruit tissues were ground to fine powder in liquid nitrogen. Around 50 - 100 mg of tissue powder was extracted in 1.2 mL of 80% methanol containing phenyl β-D-glucoside as an internal standard. After centrifuging for 40 min at 14,000 g at 4 °C, 100  $\mu$ L of the supernatant was transferred to a 300  $\mu$ L glass insert in a 2 mL GC vial. The solvent was evaporated by drying under a stream of nitrogen. Metabolites were first converted to their oxime derivatives by adding 25  $\mu$ L hydroxylamine and heating to 50 °C for 30 min, and then converted to their tri-methyl silyl (TMS) derivatives by adding 50  $\mu$ L of BSTFA (Bis(trimethylsilyl)trifluoroacetamide) and heating to 50 °C for 30 min. One  $\mu$ L of this mixture was injected and analyzed using a GC (GC-2014; Shimadzu, Japan) equipped with an HP-5 capillary column and a flame ionization detector. Helium was used as the carrier gas. The oven temperature program consisted of the following steps: 1 min hold at 150 °C, 4 °C/min ramp to 190°C, 0.5 min hold at 190°C, 1.5°C/min ramp to 210°C, 0.5 min hold at 210°C, 10°C/min ramp to 260°C, 10 min hold at 260°C. Standard solutions of known concentration were prepared for all metabolites analyzed and derivatized using the same method described above. Standard curves were generated and used for metabolite quantification. The metabolites analyzed in this study were: six sugars and sugar alcohols [Sor, Suc, Glc, Fru, xylose (Xyl), myo-inositol (Ino)], four organic acids (malate, quinate, citrate, succinate), and one amino acid (Asn).

#### Starch quantification

Starch level was determined as mg Glc equivalents g<sup>-1</sup> fresh weight following (Smith and Zeeman, 2006). Around 50 - 100 mg of tissue powder was extracted three times in 80% ethanol at 80 °C for 10 min. The pellet was retained and digested with 35 units of amyloglucosidase at pH 4.8 and at 55 °C for 36 h. The Glc concentration was measured using an enzymatic assay in which hexokinase and Glc-6-phosphate (G6P) dehydrogenase were used. NADH generated during the conversion of Glc to 6-phosphogluconate was monitored spectrophotometrically at 340 nm. A standard curve of Glc was used to determine the equivalents.

#### RNA Extraction, cDNA Synthesis, and qPCR

RNA was extracted from the pith and cortex using the CTAB based extraction buffer method described in Vashisth *et al.* (2011). Synthesis of cDNA was performed using 1  $\mu$ g of total RNA. ImProm II reverse transcriptase (Promega, USA) was used for reverse transcription in a total volume of 20  $\mu$ L. The cDNA was diluted 6-fold, and 1 $\mu$ L of diluted cDNA was used for quantitative RT-PCR following the method described previously in Dash *et al.* (2013), with the exception of using PowerUp SYBR green master mix (ThermoFisher, USA). The Stratagene Mx3005P (Agilent Technologies, USA) quantitative real-time PCR instrument was used for this analysis. Melt-curve analyses were performed at the end of the PCR amplification to determine primer specificity. Control reactions without template and without reverse transcriptase were included in the analyses. Two reference genes were used for normalization of target gene expression, *ACTIN* and *GAPDH*. Selection of genes for analysis was based on Li *et al.* (2012) and on highest abundance within the gene family based on RNA-Seq data (Jing and others, *In preparation*). List of genes and the primer sequences for qRT-PCR are presented in Supplementary

Table 2.1. Efficiencies of the qPCR reactions were determined using LinRegPCR (Ruitjers *et al.*, 2009). Relative quantity (RQ) values were determined following efficiency correction and normalized using the geometric mean of RQs of the reference genes to generate normalized RQs (NRQs). Data analysis were performed on the NRQ values after  $log_2$  transformation. Standard errors were determined as described in Rieu and Powers (2009). Expression of all genes are presented as fold change in relation to mean transcript abundance of target gene in RL fruit cortex at 0 DAT.

#### Statistical analyses

Statistical analyses and graph preparation were performed using RStudio (Version 1.0.143) and Inkscape (Version 0.92.3). Fruit diameter and length were compared between CL and RL fruit using *Student's t-test* ( $\alpha = 0.05$ ) at each stage. Paired *t-tests* ( $\alpha = 0.05$ ) were used to analyze differences in metabolite concentration and transcript abundance between pith and cortex tissues within a given stage. Comparison of metabolite concentration and transcript abundance between CL and RL treatments at a given stage were performed using the *Student's t-test* ( $\alpha = 0.05$ ). Principal components analysis (PCA) was performed to assess associations among concentrations of metabolites using the *prcomp* function in RStudio. The first two principal components were displayed using the plot function in RStudio. Pearson's correlation analysis was used to determine correlations among metabolites.
### Results

# Fruit load reduction increases growth specifically in the cortex

Fruit growth was quantified by measuring the diameter and length of the tagged king fruit, and the pith and cortex area in longitudinal section of sampled apple fruit during different stages (Figure 2.1). Fruit growth was enhanced by fruit load reduction. Fruit in the RL treatment displayed greater diameter (9.4%) and length (9.7%) at 19 DAT which continued until 118 DAT (8.5% and 9%, respectively) (Figure 2.1B and C). Pith and cortex displayed similar tissue areas at 0 DAT (Figure 2.1D and E). Cortex area increased by around 90-fold during fruit development reaching up to 36 cm<sup>2</sup> (Figure 2.1E). Pith area increased slightly from 0.2 cm<sup>2</sup> to 0.3 cm<sup>2</sup> (1.5-fold) between 0 and 19 DAT, and subsequently to ~2 cm<sup>2</sup> (6-fold) by 118 DAT (Figure 2.1D). Fruit load reduction increased cortex area from around 26 DAT resulting in 17% higher area in RL fruit at 118 DAT (Figure 2.1E). It did not alter pith area except transiently at 26 DAT (Figure 2.1D). These data indicate more intensive growth in the cortex during fruit development, and that fruit load reduction specifically enhanced cortex growth.

# PCA reveals temporal and spatial characteristics of fruit metabolism

Concentration of the 11 metabolites (Sor, Suc, Glc, Fru, Xyl, Ino, malate, quinate, citrate, succinate, Asn) and starch were quantified at nine stages of fruit development, in pith and cortex tissues, and in response to fruit load reduction. These data were subjected to PCA to determine the major components contributing to variance in fruit metabolism. Around 77% of variance was explained by two principal components (PC), PC1 (52.3%) and PC2 (24.7%; Figure 2.2). The nine stages of fruit development were clearly separated along PC1, suggesting that majority of variation was associated with temporal patterns in metabolite accumulation. In the cortex, early stages of fruit

development (0-26 DAT) were well separated from the mid (33 and 47 DAT) and late (77 and 118 DAT) stages along PC1, indicating distinct temporal metabolite accumulation patterns. A similar pattern was also displayed in the pith, especially with early and late fruit development samples forming well separated groups. Based on PCA, three phases of fruit development are defined: early fruit development (EFD; 0 - 26 DAT); mid-fruit development (MFD; 26 - 47 DAT); and late fruit development (LFD; 47 - 118 DAT). Cortex and pith samples were clearly separated during EFD, partly during MFD, and to a lesser extent during LFD along PC2. These data indicate clear spatially and temporally specific patterns in metabolite accumulation.

# Temporal and spatial patterns of metabolite accumulation during fruit growth

# Sugars and sugar alcohols

### <u>Sorbitol</u>

Following apoplastic unloading (Zhang *et al.*, 2004), Sor is transported into fruit cells and rapidly metabolized by sorbitol dehydrogenase (SDH) to Fru (Nosarszewski *et al.*, 2004). In this study, Sor concentration declined during fruit development. During EFD, it was reduced by around 54% in the cortex but remained relatively less altered in the pith (Figure 2.3A). During MFD, it decreased greatly in both tissues, especially in the pith (around 2-fold). Sor concentration and content (Figure S2.1) decreased during LFD. Sor concentration was consistently higher in the pith than in the cortex, particularly during EFD (by > 3-fold). Fruit load reduction did not affect Sor concentration.

Transcript abundance of four genes coding for Sor dehydrogenase (*MdSDH1*, *MdSDH2*, *MdSDH5* and *MdSDH9*) was investigated. *MdSDH1* and *MdSDH9* displayed an increase in transcript accumulation during EFD (Figure 2.4A and D). *MdSDH1* abundance declined during

MFD and increased slightly during LFD. *MdSDH9* transcript abundance in the cortex was higher by almost 2-fold than in the pith during MFD and declined subsequently. *MdSDH2* and *MdSDH5* transcript levels were not substantially altered during EFD and MFD (Figure 2.4B and C). But during LFD, *MdSDH2* transcript abundance increased by up to 5-fold while that of *MdSDH5* decreased slightly. Fruit load reduction resulted in a few major changes in *SDH* transcript abundance: in RL fruit cortex, *MdSDH1* abundance was around 1.5-fold higher 8 DAT; and *MdSDH5* and *MdSDH9* transcript levels were lower in RL fruit pith at 8 DAT.

# <u>Sucrose</u>

Following its unloading, Suc may be metabolized by cell wall invertase (CwINV), transported into cells and metabolized by neutral invertase (NINV) or sucrose synthase (Susy), or transported across the tonoplast and metabolized by vacuolar invertase (VINV) or stored. Products of invertase-mediated Suc metabolism are Fru and Glc, while that of Susy-mediated metabolism are Fru and UDP-Glc. Sucrose can also be synthesized in fruit cells by sucrose phosphate synthase (SPS) from Fru-6-phosphate (F6P) and UDP-Glc (Berüter and Feusi, 1997; Li *et al.*, 2012). Sucrose concentration increased in the cortex during EFD but decreased initially in the pith (between 0 and 8 DAT) and started increasing slightly after that (Figure 2.3B). It was lower in the cortex than in the pith (by up to 8-fold) at 0 DAT and until 19 DAT, especially within CL fruit. During MFD, Suc concentration in both tissues was similar, increased slightly between 26 and 33 DAT and then remained unaltered until 47 DAT. During LFD, it increased by more than 2-fold in both tissues. It decreased by 13% in RL fruit cortex at 12 DAT but was otherwise not affected by fruit load reduction.

Transcript abundance of genes coding for enzymes associated with Suc metabolism was determined using qRT-PCR. During most of fruit development *MdCwINV* transcript abundance in

the pith was consistently higher by up to 15-fold (RL fruit at 47 DAT) (Figure 2.5A), but was substantially low and relatively unaltered in the cortex. MdCwINV transcript levels in the pith decreased during EFD, increased abruptly during MFD, and declined during LFD to the same level as in the cortex at the end of the experiment. MdCwINV transcript accumulation was not affected by fruit load reduction at any stage. Transcript abundance of MdNINV3 increased in the cortex during EFD, after which it decreased during MFD and LFD. It was relatively unaltered in the pith during EFD and MFD but declined rapidly during LFD (Figure 2.5B). It was higher in the pith than in the cortex (up to 1.4-fold) until around 19 DAT but was largely similar thereafter. Transcript abundance of MdNINV4 declined between 0 and 8 DAT and was not substantially altered after 19 DAT (Figure 2.5C). It was slightly higher in the pith during MFD. MdNINV6 transcript accumulation pattern was similar to that of MdNINV4 (Figure 2.5D). Transcript abundance of all three NINV genes analyzed was lower in response to fruit load reduction by up to 2-fold in the cortex at 19 DAT. A vacuolar invertase gene, MdVINV3, displayed higher transcript abundance during EFD, and remained low thereafter (Figure 2.5E). Its transcript abundance was higher in the cortex at 0 and 8 DAT but was not affected by fruit load reduction. Transcript accumulation of a gene coding for sucrose synthase, MdSUSY3, slightly declined during fruit development but was not different between the tissues except at 47 DAT, and was not affected by fruit load reduction (Figure 2.5F).

Transcript abundance of two genes coding for sucrose phosphate synthase (*MdSPS2* and *MdSPS3*), an enzyme involved in sucrose synthesis was evaluated. *MdSPS2* transcript levels were not substantially altered during most of fruit development in the cortex. Higher transcript levels of *MdSPS2* were observed in the pith than in the cortex by up to 3 fold during EFD and by up to 18 fold at 47 DAT (Figure 2.5G). Fruit load reduction did not affect *MdSPS2* transcript accumulation

except for a 1.5-fold decrease in the pith at 19 DAT. *MdSPS3* accumulation was highest at 0 DAT, declined greatly by 8 DAT and remained low throughout the rest of fruit development. *MdSPS3* accumulation was not different between the pith and cortex and was not affected by fruit load reduction (Figure 2.5H).

### Fructose and glucose

Fructose concentration in the cortex increased continuously from 0 to 118 DAT by up to 9-fold (Figure 2.3C). A similar trend was noted in the pith except for a transient pause during MFD. Fructose concentration was generally higher in the cortex than in the pith except at 26 DAT, especially in RL fruit. Glucose concentration increased by over 4-fold during EFD and was generally higher in the cortex than in the pith during this period (30% at 19 DAT; Figure 2.3D). During MFD, it decreased slightly in all tissues, and then increased again during LFD by around 1.7-fold. Fruit load reduction did not affect Fru and Glc concentrations.

Transcript accumulation of several *FRUCTOKINASE* genes (*MdMDFK1*, *MdMDFK3*, *MdMDFK4*) and one *HEXOKINASE* (*MdHXK3*) gene, putatively encoding enzymes involved in phosphorylation of Fru and Glc, to yield F6P and G6P respectively, were analyzed. Transcript abundance of the three *FK* genes in the cortex was generally higher during EFD (Figure 2.6). In the pith, *MdMDFK1* transcript abundance declined during the initial part of EFD (0-8 DAT), increased subsequently until 47 DAT, and then declined sharply by 118 DAT (Figure 2.6A). Transcript accumulation of *MDFK1* was 2 to 3-fold greater in the pith than in the cortex during most of fruit development. *MDFK3* transcript abundance in the pith, was unaltered during early stages, but declined from 19 DAT. It was higher (up to 1.5-fold) in the cortex than in the pith at 0 and 8 DAT (Figure 2.6B). *MDFK4* levels decreased gradually during EFD in both tissues but increased during MFD in the pith (Figure 2.6C). *MDFK4* transcript accumulation was generally

higher in the pith (up to > 3-fold). Fruit load reduction did not substantially alter transcript abundance of *FK* genes. Transcript abundance of *HXK3* steadily increased during EFD by around 2-fold, and then gradually declined throughout MFD and LFD (Figure 2.6D). It was 1.65 fold higher in the cortex of RL than in CL fruit at 118 DAT.

### Myo-inositol and Xylose

Myo-inositol concentration increased by 4-fold and 2-fold in the cortex and pith tissues during EFD (Figure 2.3E). It was higher in the pith than in the cortex during most of EFD by up to 3-fold. During MFD, it was not substantially altered and remained higher in the pith, especially in CL fruit. Between 47 and 77 DAT, it increased again in both tissues by almost 2-fold, and later decreased to similar levels in both tissues. Fruit load reduction resulted in 43% lower Ino concentration in the cortex at 8 DAT. Xylose concentration decreased slightly during the initial stages but increased at later stages of EFD (Figure 2.3F). Its concentration was generally higher in the pith during this period. Later, its concentration declined until 47 DAT after which little change was noted. It was not affected by the fruit load reduction during EFD and was only slightly higher in the RL fruit cortex at 47 DAT.

# Organic acids

#### <u>Malic acid</u>

Malate was the most abundant organic acid identified in this study, consistent with multiple previous reports (Zhang *et al.*, 2010; Li *et al.*, 2012). During EFD, its concentration increased steadily by almost 3-fold in the cortex and by over 8-fold within the pith (Figure 2.7A). Malate concentration was consistently higher in the cortex than in the pith by > 2.5-fold during this period. During MFD, malate concentration declined gradually by >50% in both cortex and pith, but malate

concentration was still >2-fold higher in the cortex than in the pith during this period. During LFD, malate concentration continued to decline to similar levels between the pith and cortex by 118 DAT. However, tissue malate content was not substantially altered during MFD and declined only slightly during LFD (Figure S2.1). Fruit load reduction did not affect malate concentration.

Malate accumulation in the fruit is largely a result of its metabolism within the organ (Walker and Famiani, 2018). Malate synthesis is mediated by sequential activities of phosphoenol pyruvate (PEP) carboxylase (PEPC) and malate dehydrogenase (MDH), which convert PEP to oxaloacetic acid (OAA), and OAA to malate, respectively. Transcript abundance of *MdPEPC1* was only slightly altered during EFD after which it declined by around 50% (Figure 2.8A). *MdPEPC1* transcript abundance was 1.2-fold higher in pith of RL compared to that of CL at 8 DAT. MdPEPC2 transcript abundance gradually increased during EFD concomitant with increase in malate concentration and was not greatly altered thereafter (Figure 2.8B). It was decreased by 30% in the cortex at 0 DAT but increased by 16% in the pith at 26 DAT due to fruit load reduction. MdMDH2 transcript abundance decreased initially during EFD (between 0 and 8 DAT), increased subsequently and was variably altered during MFD across different tissues (Figure 2.8C). Fruit load reduction transiently decreased its transcript abundance by 52% in the cortex at 19 DAT. Transcript abundance of *MdMDH4* was relatively constant during EFD and MFD but increased slightly during LFD (Figure 2.8D). Transcript levels of MDH4 were increased slightly in the cortex by 14% and 37% at 8 and 47 DAT, respectively, while it was increased by 73% and 17% in RL pith at 0 and 26 DAT, respectively, in response to fruit load reduction.

# <u>Quinic acid</u>

Quinate concentration increased by up to 1.5-fold in the cortex (RL fruit) between 0 and 12 DAT and then declined (Figure 2.7B). Its concentration in the pith increased by more than 2-fold

between 0 and 19 DAT, especially in RL fruit. It was higher in the cortex than in the pith between 0 and 12 DAT by up to 1.5 fold, but was similar between these tissues by 19 DAT. Its concentration and content (Figure S2.1) decreased rapidly during MFD and remained low thereafter. In CL fruit, its concentration was slightly lower in the pith than in the cortex at 33 DAT and then higher at 77 and 118 DAT. Fruit load reduction resulted in higher quinate concentration in the cortex at 0 and 8 DAT by around 15% and 25%, respectively. It also resulted in a 10% and 14% increase in concentration at 33 DAT, in the cortex and pith, respectively.

### Citric acid and succinic acid

Citrate concentration was highest at 0 DAT and declined during EFD (Figure 2.7C). It was generally higher in the pith than in the cortex during this period. It declined further during MFD and slightly increased during LFD in both cortex and pith. The pith displayed slightly higher citrate concentration at multiple stages of MFD and LFD, especially in CL fruit. In cortex, it was increased by around 34% and 50% due to fruit load reduction at 8 DAT and 77 DAT, respectively. Succinate concentration followed a pattern similar to that of citrate, declining during EFD and MFD and then increasing slightly during LFD (Figure 2.7D). During EFD, its concentration in the pith generally tended to be higher than in the cortex by up to 3.2-fold (at 12 DAT in RL fruit). Fruit load reduction resulted in 51% higher succinate concentration in the cortex at 77 DAT.

#### Amino acid

#### <u>Asparagine</u>

As concentration in the cortex increased slightly in CL fruit and gradually decreased in RL fruit during EFD (Figure 2.9A). It declined during MFD and was subsequently unaltered during LFD. As tissue content declined during MFD and increased slightly during later stages on LFD (Figure S2.1). In the pith, Asn concentration increased by over 2-fold, reaching the peak level by 19 DAT, then declined sharply during MFD by more than 7-fold, and remained unaltered during LFD. Asn concentration in the pith was around 2-fold and 3-fold higher than that in the cortex at 0 DAT and 19 DAT, respectively. It decreased during EFD in response to fruit load reduction: in the pith at 8 DAT (by 19%), in the cortex at 12 and 19 DAT (by 35% and 37%, respectively), and in both tissues at 26 DAT (around 51% in the cortex and 41% in the pith).

As is metabolized by asparaginase (ASPA) releasing aspartate (Asp) and ammonium  $(NH_4^+; Gaufichon et al., 2016)$ . Subsequently,  $NH_4^+$  may be assimilated by glutamine synthetase (GS) using glutamate to yield glutamine. Transcript abundance of *MdASPA1* increased during EFD by around 2-fold in both tissues (Figure 2.9B). During MFD, it declined in the cortex tissue but increased in the pith. At 47 DAT, MdASPA1 transcript levels in the pith were more than 1.5-fold higher than in the cortex. Transcript abundance of MdASPA4 was generally lower in the cortex compared to the pith throughout fruit development (Figure 2.9C). Although, it still displayed low abundance during EFD in both cortex and pith, it was about over 10-fold higher in the pith than in the cortex. At 118 DAT, MdASPA4 transcript abundance was over 100-fold higher in the pith compared to the cortex. Its abundance increased dramatically in the pith by over 25-fold during MFD concomitant with a sharp decline in Asn concentration. Neither of the ASPA transcript levels were altered in response to fruit load reduction. MdGS1 transcript abundance pattern was similar to that of *MdASPA1*, increasing during EFD by > 2- fold by 19 DAT and remaining largely unchanged during the rest of fruit development in the cortex (Figure 2.9D). While the pattern was similar in the pith during EFD, *MdGS1* transcript abundance continued to increase during MFD and declined later. Fruit load reduction reduced MdGS1 transcript levels by 19% in the cortex at 19 DAT. MdGS3 which codes for a potential plastid form of glutamine synthetase, displayed higher transcript abundance during EFD decreasing during the rest of fruit development. Its abundance was >2-fold higher in the cortex than in the pith during most of EFD, suggesting greater  $NH_4^+$  re-assimilation in the cortex (Figure 2.9E).

# <u>Starch</u>

Starch concentration could not be quantified at 0 and 8 DAT due to limited tissue availability. Its concentration in the cortex and pith was very low (< 0.7 mg/g) at 12 and 19 DAT (Figure 2.10A). At 12 DAT, its concentration was generally low in both cortex and pith although marginally higher in the pith. Little starch accumulation occurred in the pith during EFD, while the concentration increased gradually in the cortex, especially in the thinned fruit by around 17 fold. During MFD, starch accumulation continued to increase in the cortex, by around 5 fold between 26 to 47 DAT. While a similar trend was observed in the pith, the levels were still much lower than in the cortex. AT 47 DAT, starch concentration in the cortex was 2.4- and 2.8-fold higher than in the pith in RL and CL fruit, respectively. During LFD, starch concentration continuously declined in the cortex. In the pith, it was unchanged until 77 DAT and then declined by 118 DAT. Tissue starch content also declined between 77 and 118 DAT (Figure S2.1). Fruit load reduction increased starch concentration in the pith by 46% at 12 DAT and by 48% in the cortex at 26 DAT.

Transcript abundance of three genes coding for ADP-glucose pyrophosphorylase (*MdAGPase3*, *MdAGPase4* and *MdAGPase5*), an enzyme catalyzing the synthesis of ADP-Glc from Glc-1-phosphate (G1P), the committing step for starch synthesis, were analyzed. *MdAGPase3* transcript abundance was not substantially altered during the early stages of EFD but increased by over 2-fold between 19 and 47 DAT, coincident with the sharp increase in starch concentration (Figure 2.10B). *MdAGPase4* and *MdAGPase5* transcript accumulation patterns

were similar and displayed little variation during EFD. Their transcript abundance increased slightly during MFD and remained similar or decreased slightly during LFD (Figure 2.10C and D). Transcript abundance of none of these genes was substantially affected by fruit load reduction, except for 32% lower *MdAGPase4* abundance in the cortex at 26 DAT. Transcript abundance of a gene coding for starch branching enzyme (*MdSBE2*), an enzyme involved in branching of the glucan chain, generally increased during EFD, especially in the cortex by over 3-fold between 0 and 26 DAT (Figure 2.10E). Subsequently, it declined slightly in the cortex during MDF and remained at a similar level until 118 DAT. Inversely, *MdSBE2* expression level increased slightly in the pith during MDF and declined dramatically during LFD. The *MdSBE2* expression level was not affected by fruit load reduction at any stage. Transcript accumulation of a gene coding for starch synthase (*MdSS1*), an enzyme involved in the addition of the Glc residue from ADP-Glc to the glucan chain, was higher during EFD but variable across the tissues (Figure 2.10F). It declined between 19 and 26 DAT and remained low thereafter. Fruit load reduction resulted in 2-fold lower transcript abundance in the cortex at 26 DAT.

#### Pearson correlation analysis reveals associations among different metabolites

Significant correlation among metabolites with a correlation coefficient < -0.65 and > 0.65 are discussed here. The two major C and N sources translocated into the fruit, Sor and Asn, displayed a strong positive correlation (0.9), the highest noted among all metabolites in this study (Figure 2.11). Sor accumulation was negatively correlated with that of Fru (- 0.7) and positively with citrate (0.72). Sucrose displayed strong positive correlations with Glc (0.7) and Fru (0.88), which were in turn positively correlated (0.83). The two major organic acids, malate and quinate displayed a positive correlation (0.67). Quinate was negatively correlated with both Suc (- 0.7) and

Fru (- 0.71). The TCA cycle intermediates, citrate and succinate were strongly correlated (0.85), and the two were correlated with Asn (0.84 and 0.66, respectively).

# Discussion

### Preferential growth in the fruit cortex

Clear spatial differences in fruit growth were noted in this study with lesser growth in the pith, especially during EFD, but also at later stages. These data suggest developmental programs that result in preferential cortex growth allowing for its establishment as the dominant fruit sink tissue. Further, fruit load reduction enhanced growth primarily in the cortex. This indicates that additional resources are specifically partitioned to the cortex to support its growth. Such differential growth is likely supported by spatiotemporally specific metabolic programs.

# Metabolism and Growth During EFD

EFD is associated with rapid growth in the cortex but substantially lesser growth in the pith. The majority of cortex growth during this period is associated with rapid cell production, lasting until around 3-4 weeks after bloom (Dash and Malladi, 2012; Dash *et al.*, 2013). Such rapid cell production-mediated growth requires synthesis of new cell wall components, membranes and cell content, and substantial energy inputs. This needs to be supported by intensive metabolism of imported C and N. Concentrations of the C sources, Sor and Suc, and imported N in the form of Asn, were several-fold lower in the cortex than in the pith during EFD, indicating their rapid consumption in the cortex. As the initial size of these tissues were similar (Figure 2.1D and E), it is likely that resources are equitably partitioned to the pith and cortex during initial stages of flower and fruit growth. Higher metabolism associated with a developmentally programmed greater

growth potential may lead to greater consumption of these resources, lowering their concentrations in the cortex. Subsequently, higher sink strength of the cortex owing to greater sink activity (higher metabolism) and size may allow for preferential resource partitioning. Such preferential resource partitioning and metabolism likely supports continued structural and metabolic demands of cortex growth.

Sorbitol metabolism is chiefly mediated by SDH (Yamaki and Ishikawa, 1986; Archbold, 1999) as indicated by the inverse correlation between Sor and Fru concentrations (Figure 2.3A and C). Transcript accumulation of *MdSDH1* and *MdSDH9* increased during EFD in a pattern complimentary to the decline in Sor concentration, implicating their gene products in its metabolism (Figure 2.4 A and D). Consistently, SDH activity is high during EFD (Archbold, 1999; Nosarszewski *et al.*, 2004), likely contributing to fruit sink strength (Archbold, 1999). Further, Sor concentration was over 2-fold lower in the cortex than in the pith. Together, these data indicate that Sor metabolism supports rapid growth during EFD (Figure 2.3A).

During EFD, Suc concentration was generally lower in the cortex, and declined transiently in response to fruit load reduction, suggesting that higher C demand for growth was also supported by Suc metabolism (Figure 2.3B). During EFD, transcript abundance of *MdNINV3* and to some extent that of *MdNINV6* increased in the cortex, while that of *MdVINV3* was higher in the cortex suggesting that these invertases are involved in Suc metabolism (Figure 2.5C, D and E). Consistently, acid invertase activity was reportedly higher during EFD (Beruter, 1985; Yamaki and Ino, 1986). In the pith, *MdCwINV* transcript abundance was multiple-fold higher during EFD, suggesting greater apoplastic Suc catabolism in this tissue (Figure 2.5A). During this period, Fru and Glc concentrations were lower in the pith, while that of Suc was higher. These data suggest

greater Suc-Suc cycle activity in the pith, which is also supported by higher *MdSPS2* transcript abundance during this period (Figure 2.5G).

Fructose in fruit cells is derived from Sor and Suc metabolism. It can subsequently enter primary C metabolism, contribute to Suc synthesis, or accumulate in the vacuole. Fructose was proposed to be primarily allocated to storage based on labeling studies (Beruter et al., 1997), but these data were obtained during MFD. During EFD, while Fru accumulated in fruit tissues, several lines of evidence suggest that it was also allocated to metabolism for supporting growth. Its concentration in the cortex was higher than in the pith, but the magnitude of this difference was lesser than the complimentary difference in Sor concentration, suggesting that a proportion of Fru derived from Sor was allocated to further metabolism in the cortex (Figure 2.3A and C). If Fru was largely allocated to storage during EFD, its concentration may be expected to be substantially higher than that of Glc which is derived primarily from Suc metabolism during this period. However, Fru and Glc concentrations during EFD were similar, suggesting that Fru was substantially allocated to metabolism (Figure 2.3C and D). Transcript abundance of three FK genes, MdFK1, MdFK3 and MdFK4 was generally high during EFD and declined at later stages, suggesting higher fructokinase activity during this period, and consistent with decline in fructokinase activity over fruit development (Li et al., 2012). Fructokinases phosphorylate Fru to F6P which can subsequently enter glycolysis (Granot et al., 2013), be converted to Glc phosphates through phosphoglucose isomerase (PGI) and phosphoglucomutase (PGM) activities and subsequently enter primary metabolism, or generate Suc through SPS activity. In the cortex, Suc concentration was relatively lower than that in the pith during the early part of EFD, suggesting that substantial F6P was allocated to primary metabolism. Hence, Fru contribution to primary C metabolism, likely supports growth during EFD, especially in the cortex. Glucose concentration

increased by almost 5-fold during EFD and was consistently higher in the cortex (Figure 2.3D). An initial step in Glc metabolism is its phosphorylation to G6P by hexokinases (Granot et al., 2013). Transcript abundance of *MdHXK3* increased in both tissues during EFD, suggesting substantial flux of Glc to G6P. Supply of labelled Glc to apple fruit discs during EFD (30 d after bloom) resulted in higher label incorporation into G6P and downstream glycolysis and tricarboxylic acid (TCA) cycle metabolites, consistent with high respiration rates during this period (Beshir et al., 2017). In the current study, TCA cycle metabolites citrate and succinate displayed higher concentration primarily during EFD, and lower concentration in the cortex, features consistent with higher TCA cycle activity in this tissue (Figure 2.7C and D). Myo-inositol, derived from G6P, also increased in concentration during EFD, was lower in the cortex, and decreased in RL fruit cortex transiently (Figure 2.3E). Myo-inositol contributes to various cellular components including membranes, and cell wall polysaccharides through the Ino oxidation pathway (Loewus and Murthy, 2000). Its accumulation pattern during EFD is consistent with its increased metabolism to synthesize cellular components for supporting enhanced growth. Further, the hexose-phosphate pool also contributes greatly to cell wall synthesis through synthesis of UDP-Glc (from G1P), a precursor for the majority of cell wall polysaccharides, and through synthesis of GDP-Mannose and GDP-Fucose (from F6P; Verbancic et al., 2018). This serves as an additional sink during EFD owing to intensive cell production. Together, data from this study suggest higher metabolism of Fru and Glc during EFD to support structural and metabolic demands of rapid growth, particularly in the cortex.

Malate and quinate increased in concentration multiple-fold, and were higher in the cortex during EFD (Figure 2.7A and B). Transcript accumulation of *MdPEPC2*, increased during EFD while that of *MdPEPC1* was generally higher during this period, suggesting substantially high

PEPC activity contributing to the conversion of PEP to OAA (Figure 2.8A and B). Transcript abundance of two MDH genes was not greatly altered during fruit development. However, owing to high *MDH* transcript abundance during EFD (Jing *et al.*, unpublished results), it is likely that MDH is not limiting for malate synthesis. Similarly, MDH transcript abundance was not correlated with malate concentration (Yao et al., 2011). Hence, malate synthesis from PEP may be regulated by PEPC abundance. PEP, along with erythrose-4-phosphate (E4P), also serves as a substrate for synthesis of dehydroquinate, a precursor of quinate and shikimate (Walker and Famiani, 2018). Further, PEP also serves as substrate for pyruvate synthesis which allows for C entry into the TCA cycle. Hence, PEP metabolism likely serves as a key branching point during EFD. Higher glycolytic flux from F6P and G6P during EFD may result in increased synthesis of PEP, especially in the cortex, which may allow for its increased allocation towards multiple competing metabolic pathways. Higher malate and quinate concentration in the cortex during EFD, and increased quinate concentration in the cortex in response to fruit load reduction, support this possibility. PEP metabolism to malate and quinate may function as a passive overflow storage process, allowing for fine-tuning of respiratory flux to meet dynamic metabolic and growth demands. Alternatively, PEP partitioning towards these organic acids may reflect a metabolic program that allows for C storage during EFD to meet energy and C skeleton demands at later stages.

Concentration of Asn generally decreased slightly in the cortex (RL fruit) and mostly increased in the pith during EFD. Further, it was lower in the cortex, and in RL fruit during most of EFD (Figure 2.9A). Cell production-mediated growth in the cortex during EFD is likely to require high N inputs. Increasing N supply enhanced cell production and fruit growth(Xia *et al.,* 2009). High N demands during EFD, particularly in response to fruit load reduction, may result in intensive Asn metabolism. Consistent with this, transcript accumulation of *MdASPG1* and *MdGS1* 

increased in a coordinated manner, while that of *MdGS3* was specifically higher in the cortex. Hence, it is likely that Asn catabolism and subsequent  $NH_4^+$  assimilation are enhanced during EFD in the cortex to support intensive growth during this period. Hence, similar to C metabolism, N metabolism is also elevated during EFD, underlining its contribution to early fruit growth.

#### Metabolism and Growth During MFD

MFD is associated with post-mitotic cell expansion-mediated growth (Dash *et al.*, 2013). Cell expansion is supported by accumulation of various metabolites such as Suc, Fru, Glc, and malate, which can function as osmolytes (Guillet *et al.*, 2002). Their accumulation may therefore facilitate cell expansion-mediated growth during MFD.

A characteristic aspect of MFD was the rapid increase in starch concentration. Cell production, a major resource sink, ceases prior to MFD (Malladi and Hirst, 2010; Dash and Malladi, 2012; Dash *et al.*, 2013). Consequently, a substantial proportion of imported C during MFD may be re-allocated to biosynthesis of starch, a transient C storage form (Beruter, 1989; Beruter *et al.*, 1997). Transcript accumulation of *MdAGPase3*, *MdAGPase4* and *MdSBE2* increased during this period suggesting transcriptional regulation of starch synthesis (Figure 2.10B, C and E). The C route for starch synthesis was proposed to primarily involve Suc metabolism by SUSY and subsequent G1P generation by UDP-Glc pyrophosphorylase (Beruter and Feusi, 1997; Beruter *et al.*, 1997). In the current study, the steady increase in Glc concentration observed during EFD was abruptly halted at the onset of MFD and subsequently, it declined slightly (Figure 2.3D). Similarly, Beruter (1989) reported a temporary decrease in Glc concentration during the starch accumulation period. Also, increase in Suc concentration observed during EFD was temporarily halted, especially between 33 and 47 DAT (Figure 2.3B). The tissue contents of these sugars increased

during MFD but at a rate lower than that during EFD. These data are consistent with the idea that diversion from Glc accumulation and Suc metabolism support starch synthesis. Starch concentration in the cortex increased at a higher rate than that in the pith. This may be supported by relatively higher C partitioning to this tissue during this period.

In the cortex, Sor concentration declined slightly while its content remained largely unchanged. As the extent of Sor import is similar during most of fruit development (Yamaki and Ishikawa, 1986), it is likely that it continues to be rapidly metabolized to Fru. Fructose concentration and content continued to increase, suggesting that much of the Fru synthesized from Sor was allocated to storage, as previously proposed (Beruter *et al.*, 1997). Hence, contribution of Sor to starch synthesis may be limited in this tissue. In the pith, Sor concentration declined substantially and its content declined slightly, while Fru concentration remained largely unchanged, suggesting alternative allocation of Fru. Transcript levels of *MdFK1* and *MdFK4* increased in the pith during this period, suggesting higher F6P synthesis, increase in the hexose-phosphate pool (Figure 2.6A and C). As Suc concentration was not substantially altered, hexose-phosphates derived from Fru may contribute to starch synthesis, suggesting a role for Sor metabolism in this process specifically in the pith.

Increase in starch content coincided with a sharp reduction in malate concentration, but the tissue malate content remained largely unaltered. These data suggest that C was re-allocated from malate to starch synthesis during MFD. Together, these data indicate a metabolic shift during MFD where C allocation is preferentially diverted towards starch synthesis from multiple other metabolic routes. Such a shift may be associated with a decrease in energy and C skeleton demands during MFD, as growth in this phase is mediated largely by post-mitotic cell expansion.

Quinate concentration and content also declined substantially during MFD. Quinate may be converted to shikimate, directly *via* dehydroshikimate, or *via* dehydroquinate (Walker and Famiani, 2018). Shikimate metabolism supports biosynthesis of aromatic amino acids such as phenylalanine, a precursor for lignin and many secondary metabolites such as flavonoids and anthocyanins, through the phenylpropanoid metabolic pathway (Vogt, 2010). Phenylalanine concentration increases during MFD (Zhang *et al.*, 2010; Beshir *et al.*, 2017), suggesting that quinate accumulation during EFD may allow for it to serve as a C source to support synthesis of various secondary metabolites *via* shikimate and phenylpropanoid metabolism during MFD, as proposed previously in kiwifruit(Marsh *et al.*, 2009; Walker and Famiani, 2018).

As concentration and content declined rapidly during MFD indicating increased N metabolism, potentially coupled with a decrease in fruit N import or a change in the N source. Interestingly, a greater decline in its concentration was noted in the pith than in the cortex. This was associated with a steep increase in MdASPG4 transcript abundance and higher transcript abundance of MdASPG1 and MdGS1, suggesting higher N metabolism in this tissue. Nitrogen requirement either for growth or for differential synthesis of secondary metabolites in the pith may result in enhanced Asn catabolism and  $NH_4^+$  re-assimilation during this period.

### Metabolism and Growth During LFD

LFD is associated with growth mediated by post-mitotic cell expansion, and ripening (Malladi and Hirst, 2010; Malladi and Johnson, 2011). There was no further increase in starch concentration in both tissues during LFD. Its concentration and content decreased dramatically by 118 DAT, especially in the cortex. Spatial differences in starch accumulation across the fruit at maturity were suggested to be associated with differential starch accumulation during fruit development

(Doerflinger *et al.*, 2015). Here, the rate of starch degradation was higher in the cortex than in the pith indicating that degradation rates also contribute to spatial patterns of starch abundance. Starch degradation likely contributes to the resumption of, and substantial increase in Glc and Suc concentrations, partly to meet the metabolic requirements of a respiratory climacteric, as proposed previously (Berüter and Feusi, 1997). Fructose concentration continued to increase during LFD, supported by Sor metabolism as Sor concentration and content declined greatly. Increase in *MdSDH2* transcript accumulation as associated activity may contribute to Sor metabolism and preferential accumulation of Fru during LFD. Continued accumulation of Suc, Fru and Glc during LFD may allow for maintenance of vacuolar osmotic concentration needed for continued cell expansion and also contribute to fruit quality.

# Effects of resource availability on fruit metabolism

Early reduction of fruit load affects fruit metabolism but does not dramatically alter metabolite concentrations in apple, consistent with previous reports (Beruter *et al.*, 1997; Dash *et al.*, 2013). Entry of additional resources into the fruit may lead to their increased allocation toward growth through cell production (Dash *et al.*, 2013), while maintaining a threshold concentration of free metabolites. However, severely restricting C supply during MFD by girdling resulted in multiple metabolic changes indicating re-mobilization of C from stored reserves to meet growth demands (Berüter and Feusi, 1997). Also in tomato, C starvation substantially altered metabolite concentrations and expression of genes associated with growth, and C and N metabolism (Baldet *et al.*, 2002). Reponses to severe C limitation may require extensive metabolic re-programming. However, relatively milder alteration of resource availability through fruit load reduction and early timing of this treatment may elicit different responses.

# References

**Archbold DD.** 1999. Carbohydrate availability modifies Sordehydrogenase activity of apple fruit. Physiologia Plantarum **105**, 391-395.

**Baldet P, Devaux C, Chevalier C, Brouquisse R, Just D, Raymond P.** 2002. Contrasted responses to carbohydrate limitation in tomato fruit at two stages of development. Plant, Cell & Environment **25**, 1639-1649.

Bepete M, Lakso AN. 1998. Differential effects of shade on early-season fruit and shoot growth rates in empire' apple. HortScience **33**, 823-825.

**Berüter J.** 1990. Carbohydrate partitioning and changes in water relations of growing apple fruit. Journal of Plant physiology **135**, 583-587.

**Berüter J, Feusi MES.** 1997. The effect of girdling on carbohydrate partitioning in the growing apple fruit. Journal of plant physiology **151**, 277-285.

**Beshir WF, Mbong F, Hertog ML, Geeraerd AH, Van den Ende W, Nicolaï BM.** 2017. Dynamic labeling reveals temporal changes in carbon re-allocation within the central metabolism of developing apple fruit. Frontiers in plant science **8**, 1785.

**Bieleski R.** 1982. Sugar alcohols. Plant carbohydrates I, Springer: 158-192. **Byers RE, Carbaugh DH.** 1991. Effect of chemical thinning sprays on apple fruit set. HortTechnology 1, 41-48.

**Carrari F, Baxter C, Usadel B,** *et al.* 2006. Integrated analysis of metabolite and transcript levels reveals the metabolic shifts that underlie tomato fruit development and highlight regulatory aspects of metabolic network behavior. Plant Physiology **142**, 1380-1396.

**Carrari F, Fernie AR.** 2006. Metabolic regulation underlying tomato fruit development. Journal of Experimental Botany **57**, 1883-1897.

**Chapman GW Jr, Horvat RJ.** 1989. Determination of nonvolatile acids and sugars from fruits and sweet potato extracts by capillary GLC and GLC/MS. Journal of Agricultural and Food Chemistry **37**, 947-950.

**Dash M, Johnson LK, Malladi A.** 2013. Reduction of fruit load affects early fruit growth in apple by enhancing carbohydrate availability, altering the expression of cell production-related genes, and increasing cell production. Journal of the American Society for Horticultural Science 138, 253-262.

**Dash M, Malladi A.** 2012. The AINTEGUMENTA genes, MdANT1 and MdANT2, are associated with the regulation of cell production during fruit growth in apple (Malus× domestica Borkh.). BMC plant biology **12**, 98.

**Doerflinger FC, Miller WB, Nock JF, Watkins CB.** 2015. Relationships between starch pattern indices and starch concentrations in four apple cultivars. Postharvest Biology and Technology **110**, 86-95.

**Fait A, Hanhineva K, Beleggia R, Dai N, Rogachev I, Nikiforova VJ, Fernie AR, Aharoni A.** 2008. Reconfiguration of the achene and receptacle metabolic networks during strawberry fruit development. Plant physiology **148**, 730-750.

Gillaspy G, Ben-David H, Gruissem W. 1993. Fruits: a developmental perspective. <u>The Plant</u> Cell 5, 1439.

**Giovannoni JJ.** 2004. Genetic regulation of fruit development and ripening. The plant cell **16**, S170-S180.

**Granot D, Kelly G, Stein O, David-Schwartz R.** 2013. Substantial roles of hexokinase and fructokinase in the effects of sugars on plant physiology and development. Journal of Experimental Botany **65**, 809-819.

**Guak S, Neilsen D, Millard P, Looney NE.** 2004. Leaf absorption, withdrawal and remobilization of autumn-applied urea-N in apple. Canadian journal of plant science **84**, 259-264.

Guillet C, Just D, Bénard N, Destrac-Irvine A, Baldet P, Hernould M, Causse M, Raymond P, Rothan C. 2002. A fruit-specific phosphoenolpyruvate carboxylase is related to rapid growth of tomato fruit. Planta 214, 717-726.

**HANSEN P.** 1971. <sup>14</sup>C-studies on apple trees. VII. The early seasonal growth in leaves, flowers and shoots as dependent upon current photosynthates and existing reserves. Physiologia Plantarum **25**, 469-473.

**Kromdijk J, Ubierna N, Cousins AB, Griffiths H.** 2014. Bundle-sheath leakiness in C<sub>4</sub> photosynthesis: a careful balancing act between CO<sub>2</sub> concentration and assimilation. Journal of experimental botany **65**, 3443-3457.

Li M, Feng F, Cheng L. 2012. Expression patterns of genes involved in sugar metabolism and accumulation during apple fruit development. PloS one 7, e33055.

**MacDaniels LH.** 1940. Morphology Of The Apple And Others Pome Fruits, Cornell University, Agricultural Experiment Station; Ithaca, New York.

**Malaguti D, Millard P, Wendler R, Hepburn A, Tagliavini M.** 2001. Translocation of amino acids in the xylem of apple (*Malus × domestica* Borkh.) trees in spring as a consequence of both N remobilization and root uptake. Journal of Experimental Botany **52**, 1665-1671.

**Malladi, A. and P. M. Hirst.** 2010. Increase in fruit size of a spontaneous mutant of 'Gala' apple (*Malus*  $\times$  *domestica* Borkh.) is facilitated by altered cell production and enhanced cell size. Journal of experimental botany 61, 3003-3013.

**Malladi A, Johnson LK.** 2011. Expression profiling of cell cycle genes reveals key facilitators of cell production during carpel development, fruit set, and fruit growth in apple (*Malus*  $\times$  *domestica* Borkh.). Journal of experimental botany **62**, 205-219.

Marsh KB, Boldingh HL, Shilton RS, Laing WA. 2009. Changes in quinic acid metabolism during fruit development in three kiwifruit species. Functional Plant Biology **36**, 463-470.

Moco S, Capanoglu E, Tikunov Y, Bino RJ, Boyacioglu D, Hall RD, Vervoort J, De Vos RC. 2007. Tissue specialization at the metabolite level is perceived during the development of tomato fruit. Journal of Experimental Botany 58, 4131-4146.

Morandi B, Grappadelli LC, Rieger M, Bianco RL. 2008. Carbohydrate availability affects growth and metabolism in peach fruit. Physiologia Plantarum 133, 229-241.

Mounet F, Moing A, Garcia V, *et al.* 2009. Gene and metabolite regulatory network analysis of early developing fruit tissues highlights new candidate genes for the control of tomato fruit composition and development. Plant Physiology **149**, 1505-1528.

Nosarszewski M, Clements AM, Downie AB, Archbold DD. 2004. Sorbitol dehydrogenase expression and activity during apple fruit set and early development. Physiologia Plantarum 121, 391-398.

**Pratt C.** 1988. Apple flower and fruit: morphology and anatomy. Horticultural Reviews **10**, 273-308.

**Prudent M, Causse M, Génard M, Tripodi P, Grandillo S, Bertin N.** 2009. Genetic and physiological analysis of tomato fruit weight and composition: influence of carbon availability on QTL detection. Journal of experimental botany **60**, 923-937.

Smith AM, Zeeman SC. 2006. Quantification of starch in plant tissues. Nature\_protocols 1, 1342.

**Tukey HB, Young JO.** 1942. Gross morphology and histology of developing fruit of the apple. Botanical Gazette **104**, 3-25.

**Vashisth T, Johnson LK, Malladi A.** 2011. An efficient RNA isolation procedure and identification of reference genes for normalization of gene expression in blueberry. Plant cell reports **30**, 2167-2176.

Vogt T. 2010. Phenylpropanoid biosynthesis. Molecular plant 3, 2-20.

Walker RP, Famiani F. 2018. Organic acids in fruits: metabolism, functions and contents. Hortic Rev.

Xia G, Cheng L, Lakso A, Goffinet M. 2009. Effects of nitrogen supply on source-sink balance and fruit size of 'Gala'apple trees. Journal of the American Society for Horticultural Science 134, 126-133.

**Yamaki S, Ino M.** 1992. Alteration of cellular compartmentation and membrane permeability to sugars in immature and mature apple fruit. Journal of the american Society for Horticultural Science **117**, 951-954.

**Yao JL, Tomes S, Xu J, Gleave AP.** 2016. How microRNA172 affects fruit growth in different species is dependent on fruit type. Plant signaling & behavior **11**, 417-427.

Zhang LY, Peng YB, Pelleschi-Travier S, Fan Y, Lu YF, Lu YM, Gao XP, Shen YY, Delrot S, Zhang DP. 2004. Evidence for apoplasmic phloem unloading in developing apple fruit. Plant Physiology 135, 574-586.

**Zhang YZ, Li P, Cheng L.** 2010. Developmental changes of carbohydrates, organic acids, amino acids, and phenolic compounds in 'Honeycrisp' apple flesh. Food Chemistry **123**, 1013-1018.

Gene	Accession Number	Primer Orientation	Primer Sequence (5'-3')
MdACTIN	EB127077	Forward	ACCATCTGCAACTCATCCGAACCT
		Reverse	ACAATGCTAGGGAACACGGCTCTT
MdGAPDH	EB146750	Forward	TGAGGGCAAGCTGAAGGGTATCTT
		Reverse	TCAAGTCAACCACACGGGTACTGT
MdSDH1	MDP0000932467	Forward	GAGTCTTGGCGCAGATGCAGT
		Reverse	ACAGTCGAAGGTTACATCCACTCCATT
MdSDH2	MDP0000874667	Forward	CATTGCCAGCAGTGCAAAGGC
		Reverse	GGCAATTTAAAGCACAGATCCGCG
MdSDH5	MDP0000250546	Forward	GTTAGAGATGTCAAACCTGTGGAGA
		Reverse	GCAAATGCCGACAGCCTTAATT
MdSDH9	MDP0000188052	Forward	CCTGCAATGGCATGGTTAGACAA
		Reverse	CACAAATGCCGACAGCCTTG
MdCwINV	MDP0000275150	Forward	CCTCATCAATTGGGAAGCTCTTGAG
		Reverse	GATAGGGGTCCGACGCATTTTTC
MdNINV3	MDP0000186866	Forward	GTACTCCATGATCCTGTCCGGAATAGT
		Reverse	CATACCCTTCTGGCATTCAGGCAG
MdNINV4	MDP0000652278	Forward	GTCTTGGCCAATCTGGGATAATAT
		Reverse	ATTCACACGGGTCTCAATTGAC
MdNINV6	MDP0000261740	Forward	AATGCCCAATTGTGTACAAATGCG
		Reverse	GGTATCGCTTTATCTGATTCGTCTACA
MdVINV3	MDP0000377084	Forward	CCCTGACGGCCAAATCATAATGT
		Reverse	GAAATCAGTGGATCCGATCCCG
MdSUSY3	MDP0000126946	Forward	GGAAAAGAATACTGCAGCCGCACG
		Reverse	GAACTTCGCTGAAAGGTCCGGT
MdSPS2	MDP0000288684	Forward	GCATCACAAGCAATCAGATGTACCT
		Reverse	CAAGCCACAGGTTTTTGTCTCCT
MdSPS3	MDP0000331376	Forward	CGAGGGAGAGAAGGGAGATTTG
		Reverse	TCTGCTGACTAATCCATGTCTCCATT
MdFK1	MDP0000173131	Forward	GACTGGTGGTGATGATCCTTGC
		Reverse	CCCGCCCATGGAATTTCTGT

Table 2.1. List of the apple genes and the sequence of primers used in quantitative RT-PCR analyses.

Gene	Accession Number	Primer Orientation	Primer Sequence (5'-3')
MdFK3	MDP0000309723	Forward	AAGCATTTGCAGGAGATGTGCTAT
		Reverse	GCTTCAGCCAATGAAAGTCCATTAGTA
MdFK4	MDP0000765663	Forward	TCTAGCTGCTATGAAAATTGCCAAGG
		Reverse	CCAGTCAGGAATCTAATTTCATCCTCG
MdHXK3	MDP0000643891	Forward	GGCAAGATGTAGTGGCTGAATTG
		Reverse	CCTCCAGCTAATGTCCCAACC
MdAGPase3	MDP0000203812	Forward	AACTCTTCATCCAAGTTGGCCAG
		Reverse	GATTATACTGCTGGAAATGCATGCA
MdAGPase4	MDP0000394192	Forward	GAAACGTTCTGGTGGACGAGACT
		Reverse	ATGCTGCTGGAAATGCAAGCG
MdAGPase5	MDP0000323050	Forward	GAGGCTCACAAAAAATATGGTGGG
		Reverse	AGTGGAAACACCTTCAACGGC
MdSBE2	MDP0000214735	Forward	AAACAATGCAGATGGCTCACCTTC
		Reverse	AACATGTGCCTCATAAATGCGAAGTG
MdSS1	MDP0000842179	Forward	GACTGACAGCTCTGTGCTTGTTC
		Reverse	GAGCTCTTTCCAATCGCATGAAG
MdPEPC1	MD03G1242000	Forward	CTCAAAGCGAAAACCTAGTGGTGGT
		Reverse	GCTCCAAGGCCTAACCACACA
MdPEPC2	MD17G1230800	Forward	TCCAAATGCTTCGGGAGATGTATAATCAG
		Reverse	CTGACACGAGAAGCTTGTCATACAGAGAG
MdMDH2	MD07G1073300	Forward	ACTCTCACCGCCGTGTGG
		Reverse	AACACGCCGAAGCACTCAGG
MdMDH4	MD16G1219000	Forward	GGCCTTGGGCCAGGTTTCT
		Reverse	ACAAGCTCACGGACACCTTTCTCC
MdASPA1	MD06G1205500	Forward	GCTAGGGAACAGGGTGTTGAGACT
		Reverse	TCACCATCAGCATCTGGTGTTTCTTCT
MdASPA4	MD08G1092600	Forward	AGGCGTAGAGAGTCCGTTGGT
		Reverse	CAGCGCAACGCCCTTGGC
MdGS1	MD17G1268700	Forward	CATCAACCTGGATCTCTCAGGCTCT
		Reverse	TTCACTATCTTCTCCAGGAGCTTGACC
MdGS3	MD13G1180400	Forward	GCAGATGAAGATTACAAGGAGCTCAACC
		Reverse	GTTCAGTAAACCCTCTAGCCTGTTGATT



Figure 2.1. Effects of fruit load reduction on fruit growth.

Diagram of apple fruit in longitudinal section showing cortex, pith and seed locule (A). Diameter (B), length (C), and longitudinal sectional areas of pith (D) and cortex (E) of apple fruit are presented. Mean  $\pm$  S.E (n = 4) are displayed here. Asterisk indicates significant difference between the control (CL) and reduced fruit load (RL) treatments at  $\alpha = 0.05$ .



**Figure 2.2.** *Principal components analysis (PCA) reveals spatial and temporal characteristics of fruit metabolism.* 

The first and second components explained 77% of variation in data and are displayed here. The ovals display three clusters divided based on the temporal variation of fruit metabolism in the cortex. The square display two clusters divided based on the spatial variation of fruit metabolism. The numbers above the symbols indicate the days after treatment. Letters next to the numbers indicate treatment and tissue type. CC: Cortex of control fruit load (CL) fruit; CP: Pith of CL fruit; RC: Cortex of reduced fruit load (RL) fruit; RP: Pith of RL fruit.



**Figure 2.3.** The concentrations of major sugars/sugar alcohols in the cortex and pith tissues in response to fruit load reduction.

Shaded regions in the background indicate early (dark grey), mid (light grey), and late fruit development (white) periods. The mean and standard error of the mean (n = 4) are displayed. CC: Control fruit load cortex; CP: Control fruit load pith; RC: Reduced fruit load cortex; RP: Reduced fruit load pith. The asterisk indicates significant difference between control and reduced fruit load treatments in the cortex while the same is displayed with dots for the pith ( $\alpha = 0.05$ ).



**Figure 2.4.** *Relative expression of sorbitol metabolism related genes in the cortex and pith tissues in response to fruit load reduction.* 

Shaded regions in the background indicate early (dark grey), mid (light grey), and late fruit development (white) periods. The mean and standard error of the mean (n = 4) are displayed. CC: Control fruit load cortex; CP: Control fruit load pith; RC: Reduced fruit load cortex; RP: Reduced fruit load pith. Asterisks and † indicates the significant difference between control and reduced fruit load treatments in the cortex and pith, respectively ( $\alpha = 0.05$ ). All expression data are in reference to mean expression at 0 d after treatment in RC tissues.



**Figure 2.5.** *Relative expression of sucrose metabolism related genes in the cortex and pith tissues in response to fruit load reduction.* 

Shaded regions in the background indicate early (dark grey), mid (light grey), and late fruit development (white) periods. The mean and standard error of the mean (n = 4) are displayed. CC: Control fruit load cortex; CP: Control fruit load pith; RC: Reduced fruit load cortex; RP: Reduced fruit load pith. Asterisks and † indicates the significant difference between control and reduced fruit load treatments in the cortex and pith, respectively ( $\alpha = 0.05$ ). All expression data are in reference to mean expression at 0 d after treatment in RC tissues.



**Figure 2.6.** *Relative expression of fructose and hexose metabolism related genes in the cortex and pith tissues in response to fruit load reduction.* 

*FK*: *FRUCTOKINASE*; *HXK*: *HEXOKINASE*. Shaded regions in the background indicate early (dark grey), mid (light grey), and late fruit development (white) periods. The mean and standard error of the mean (n = 4) are displayed. CC: Control fruit load cortex; CP: Control fruit load pith; RC: Reduced fruit load cortex; RP: Reduced fruit load pith. Asterisks and † indicates the significant difference between control and reduced fruit load treatments in the cortex and pith, respectively ( $\alpha = 0.05$ ). All expression data are in reference to mean expression at 0 d after treatment in RC tissues.



**Figure 2.7.** The concentrations of major organic acids in the cortex and pith tissues in response to fruit load reduction.

Shaded regions in the background indicate early (dark grey), mid (light grey), and late fruit development (white) periods. The mean and standard error of the mean (n = 4) are displayed. CC: Control fruit load cortex; CP: Control fruit load pith; RC: Reduced fruit load cortex; RP: Reduced fruit load pith. Asterisks and † indicates the significant difference between control and reduced fruit load treatments in the cortex and pith, respectively ( $\alpha = 0.05$ ).



**Figure 2.8.** The relative expression of malate metabolism-related in the cortex and pith tissues in response to fruit load reduction.

*PEPC*: *PHOSPHOENOL CARBOXYLASE*; *MDH*: *MALATE DEHYDROGENASE*. Shaded regions indicate early (dark grey), mid (light grey), and late fruit development (white) periods. The mean and standard error of the mean (n = 4) are displayed. CC: Control fruit load cortex; CP: Control fruit load pith; RC: Reduced fruit load cortex; RP: Reduced fruit load pith. Asterisks and  $\dagger$  indicates the significant difference between control and reduced fruit load treatments in the cortex and pith, respectively ( $\alpha = 0.05$ ). All expression data are in reference to mean expression at 0 d after treatment in RC tissues.


**Figure 2.9.** As *n* concentration and the expression of As *n* metabolism related gene in the cortex and pith tissues in response to fruit load reduction.

Shaded regions indicate early (dark grey), mid (light grey), and late fruit development (white) periods. The mean and standard error of the mean (n = 4) are displayed. CC: Control fruit load cortex; CP: Control fruit load pith; RC: Reduced fruit load cortex; RP: Reduced fruit load pith. *ASPA: ASPARAGINASE*; *GS: GLUTAMINE SYNTHETASE*. Asterisks and † indicates the significant difference between control and reduced fruit load treatments in the cortex and pith, respectively ( $\alpha = 0.05$ ). All expression data are in reference to mean expression at 0 d after treatment in RC tissues.



Time after Treatment (d)

**Figure 2.10.** *Starch concentration and the expression of starch metabolism related gene in the cortex and pith tissues in response to fruit load reduction.* 

Shaded regions indicate early (dark grey), mid (light grey), and late fruit development (white) periods. The mean and standard error of the mean (n = 4) are displayed. CC: Control fruit load cortex; CP: Control fruit load pith; RC: Reduced fruit load cortex; RP: Reduced fruit load pith. *AGPase: ADP GLUCOSE PYROPHOSPHORYLASE; SBE: STARCH BRANCHING ENZYME; SS: STARCH SYNTHASE.* Asterisks and † indicates the significant difference between control and reduced fruit load treatments in the cortex and pith, respectively ( $\alpha = 0.05$ ). All expression data are in reference to mean expression at 0 d after treatment in RC tissues.



**Figure 2.11.** A correlation matrix among different metabolites represented as color-keyed correlation coefficients (upper triangle) with heat map (lower triangle).

Positive correlations are indicated in blue and negative correlations are in red.

# **Supplement Data**



**Figure S2.1.** *Effects of fruit load reduction on estimated metabolites amounts in the cortex and pith tissues of apple fruit.* 

Shaded regions indicate early (dark grey), mid (light grey), and late fruit development (white) periods. The mean and standard error of the mean (n = 4) are displayed. CC: Control fruit load cortex; CP: Control fruit load pith; RC: Reduced fruit load cortex; RP: Reduced fruit load pith. Asterisks and  $\dagger$  indicates the significant difference between control and reduced fruit load treatments in the cortex and pith, respectively ( $\alpha = 0.05$ ).

# CHAPTER 3

# EXPRESSION PROFILING OF PUTATIVE ORGAN SIZE REGULATORY GENES DURING PRE-BLOOM, FRUIT SET AND FRUIT DEVELOPMENT IN APPLE

Jing, S. and Malladi, A. To be submitted to journal of American society for horticultural science

#### Abstract

Apple is the most consumed fresh fruit among U.S. consumers. Fruit size is one of the most important fruit quality traits. Fruit growth is facilitated by cell production and cell expansion. A subset of genes regulating cell production have been functionally characterized in model plant species. In this study, homologous genes of five organ growth regulatory genes: FW2.2/CNR, GRFs, GIFs, ARGOS/ARGOS-Like and KLUH were identified in the apple genome: 11, 14, 4, 2, and 4 homologous genes in each of the gene family. Transcript abundance profiling of these genes was performed before full bloom, during fruit set and the whole developmental process in apple. Three apple *GRF* and one *GIF* genes displayed transcript abundance patterns consistent with roles as positive regulators of cell production during fruit development, while three members of CNR displayed transcript abundance patterns suggesting a role as a cell production inhibitor. The four members of KLUH displayed very too expression to be quantified in this study. Two apple ARGOS were likely to be negative regulators of cell production. The expression profiling of the homologous genes were investigated further in response to fruit load reduction and shading. Transcript abundance of the apple MdGRF1, MdGRF7a, MdGRF8 and MdGIF3 genes was upregulated due to fruit load reduction but downregulated by shading. In contrast, the expression of MdARG1 and MdARG2 and CNR5a, CNR6 and CNR8 was downregulated in thinned fruits but upregulated in shaded fruits. Together, these data suggest the some of the organ growth regulating genes from model plant species may play similar roles in regulating apple fruit growth.

Key words: Cell production, FW2.2, GRF, GIF, ARGOS, KLUH, fruit development

#### Introduction

Apple flowers are induced and initiate during the year prior to flowering. After bud break, there is extensive growth in the bud mediated by cell production until around one week before bloom (Malladi and Johnson, 2011). The cessation of cell production has been observed in both apple and tomato, and is reinitiated upon fertilization(Malladi and Hirst, 2010). Under optimum conditions, more than one flower within each flower cluster is pollinated. Multiple hormones including auxins, gibberellins, ethylene, and ABA potentially interact with each other, resulting in the abscission of unfertilized flowers and resumption of cell production in fertilized flowers (Bangerth, 2000). During the first 3-4 weeks after pollination (bloom), the number of cells in the cortex increases by 8- to 10-fold (Dash and Malladi, 2012), and the total cell volume increases by around 20 fold before cell production ceases. During the later stage of development, the majority of fruit growth is associated with cell expansion. The total cell volume increases by up to 1500 fold before harvest (Dash and Malladi, 2012), and contributes most to increase in size of the organ.

The molecular mechanisms regulating organ growth have been extensively studied in model species including Arabidopsis and tomato. Multiple factors regulating organ growth have now been described, and connections involving different regulators are starting to emerge from various studies, primarily in the model species Arabidopsis(Breuninger and Lenhard, 2010). Some of the key regulators of organ growth identified in model species are summarized below.

#### FW2.2/Cell Number Regulator

*Fruit weight (FW)2.2* was initially identified as a quantitative trait locus on chromosome 2 responsible for approximately 30% of the fruit weight variation between domesticated tomatoes and their small-fruited wild relatives (Frary *et al.*, 2000). Higher transcript abundance of *FW2.2* is

72

correlated with a reduction of cell division in carpels of the small fruited near isogenic lines (NILs) of tomato, indicating that FW2.2 functions as a negative regulator of cell division (Frary et al., 2000). The role of FW2.2 in regulating fruit weight appears conserved across different species including eggplant (Solanum melongena), pepper (Capsium annuum) and Physalis floridana(Li and He, 2014). Homologs of FW2.2 were identified in maize and named as Cell Number Regulators (CNRs)(Guo et al., 2010). Thirteen CNR gene family members were identified in the maize genome (Zea mays). Overexpression of ZmCNR1 in maize resulted in a reduction of overall plant size and that of multiple organs (Guo et al., 2010; Guo and Simmons, 2011). CNR genes have also been identified in several species in Prunus family including peach (Prunus persica) and sweet cherry (Prunus avium) through a genome-wide search(De Franceschi et al., 2013). While the role of regulating plant and organ size through the alteration of cell number of CNR may be conserved across species, how the cell number is regulated by CNR is not very clear. CNR may interact with CKII regulatory subunit  $\beta$  (Frary *et al.*, 2000). As CKIIs are likely involved in the regulation of the cell cycle, it was proposed that the interaction between CNR and kinase components may regulate progression of the cell cycle (Cong and Tanksley, 2006). Further, CNR was also found to interact with AGAMOUS (AG), which interacted with a D-type CYCLIN (CYCD2;1) promoter in P. floridana(Li and He, 2014), further suggesting a potential functional link to regulation of the cell cycle. Potential homologs of CNR in apple have not yet been reported.

#### **GRF** and **GIF**

*Growth Regulating Factors* (*GRFs*) belong to a family encoding a class of plant specific transcription factors. The *GRF* family has been identified experimentally or *in silico* in many species including rice (*Oryza sativa*), *Arabidopsis*, maize (*Zea mays*), soybean (*Glycine max*),

Chinese cabbage (Brassica rapa) and almost all land plant genomes sequenced to date (Omidbakhshfard et al., 2015). Most GRF proteins share two domains, QLQ (glutamine, leucine, glutamine) and WRC (tryptophan, arginine, cysteine). The QLQ domain is involved in proteinprotein interaction with the GRF-interacting factor (GIF), while the WRC domain is relevant for DNA binding and TF targeting. Early studies identified the function of GRF family as regulators of leaf, stem and floral organ development (Frary et al., 2000; Kim et al., 2003; Horiguchi et al., 2005; Kim and Lee, 2006), but recent studies indicated its potential role in central developmental processes including regulation of flower time, seed and root development, control of growth under adverse environment, and plant longevity (Frary et al., 2000; Kim et al., 2003; Kim and Lee, 2006; Liu et al., 2014). The Arabidopsis genome harbors 9 AtGRF members. The Arabidopsis triple mutant, grf1/2/3, developed smaller and narrower leaves with shorter petioles, while overexpression of AtGRF1 or AtGRF2 resulted in larger leaves and cotyledons (Kim et al., 2003; Kim and Kende, 2004). Furthermore, the interaction between GRFs and GIFs have been identified in Arabidopsis, maize and rice(Omidbakhshfard et al., 2015). GIF proteins have been identified as regulators which control both cell division rate and duration. The Arabidopsis triple mutant, gif1/2/3, displayed a reduction in cell number and extremely smaller plants compared to the gif2 or gif3 single mutants(Kim and Lee 2006, Lee; Ko et al., 2009). The GRF-GIF complex is potentially involved in the regulatory cascade that includes another transcription factor and several micro RNAs: miR319-TCP4-miR396-GRF/GIF which has been proposed to control cell production in Arabidopsis (Rodriguez et al., 2010; Kalve et al., 2014).

#### AUXIN-REGULATED GENE INVOLVED IN ORGAN SIZE (ARGOS)

*ARGOS* is a auxin-inducible gene that controls organ growth and cell production by potentially regulating the transcription factor, AINTEGUMENTA (ANT) (Hu *et al.*, 2003). On the other hand, *ARGOS*-like (ARL), a gene sharing high sequence similarity with *ARGOS*, controls organ size by affecting cell expansion and acts downstream of brassinosteroids (Hu *et al.*, 2006). *ARGOS* was initially identified in a search for genes responsive to NAA (Naphthyl acetic acid) treatment in roots of young *Arabidopsis* seedlings. Transgenic *Arabidopsis* over-expressing *ARGOS* displayed enlarged aerial organs with increased cell number and extended expression of the cell cycle gene, *CYCD3;1* (Hu *et al.*, 2003). Overexpression of the orthologous genes *OsARGOS* in Arabidopsis resulted to larger organ size with increased cell production and expansion (Wang *et al.*, 2009).

# KLUH

A mobile plant organ growth stimulator involving the activity of *KLUH* gene product was initially described in *Arabidopsis(Anastasiou et al., 2007)*. *KLUH* encodes a putative microsomal cytochrome P450 monooxygenase CYP78A5 and is thought to be plant specific (Anastasiou *et al.,* 2007). In *Arabidopsis*, the loss-of-function of *KLUH* led to thinner stems, smaller flowers and leaves, and lighter and smaller seeds than the wild type. Over-expression of *KLUH* resulted in increased growth of the organ due to alteration of the cell number, indicating that it functions a positive regulator of growth (Anastasiou *et al.,* 2007; Adamski *et al.,* 2009). *KLUH* expression is not restricted to actively dividing regions. For example, the KLU protein is detected at the base and around the periphery of developing petals when most of the cell division is occurring at the center of petal (Anastasiou *et al.,* 2007). *KLUH* is thought to determine the time when cell

production ceases potentially based on a threshold. Additionally, in tomato, a KLUH homolog was identified as the candidate gene underlying a fruit size-related QTL (Chakrabarti *et al.*, 2013). Members of the above gene families have been demonstrated to play important roles in organ size control in several plant species. Hence, it is hypothesized that corresponding members of these gene families in apple may play similar roles in regulating fruit growth. The objective of this study was to identify the potential homologs of *FW2.2/CNR*, *GRF*, *GIF*, *ARGOS/ARL*, and *KLUH* in apple. Further, to initially characterize the functions of these genes and to identify candidates for further functional studies, the expression profiling of multiple gene family members was performed before bloom, during fruit set and fruit development, and in response to fruit load reduction and shading treatments.

#### **Materials and Methods**

#### **Plant Material**

Total RNA extracted from various experiments previously in Malladi's lab was utilized in this study. The brief set up of each experiment is described as follows.

1) Fruit developmental study

The king flowers of 'Gala' was collected at 26, 17, 7 and 0 days before full bloom (DBFB) (n=4). Total RNA was extracted from the floral tube region. After full bloom, fruit was sampled at 11 developmental stages including 0, 8, 11, 14, 18, 21,24, 43, 56, 79 and 123 days after full bloom(DAFB) and used for total RNA extraction (n=4).

2) Fruit set study

A controlled pollination study was set up along with the fruit development study. At 6 DAFB, all flower clusters were manually thinned to one king flower per cluster, and randomly assigned for the controlled pollination study. The 'pollinated' group was manually pollinated using pollen from

multiple apple varieties grown on the same location in addition to the open pollination afterward. The flowers of the 'non-pollinated' group were forced to open and the styles were removed using scissors. Fruit from 'pollinated' and 'non-pollinated' group was sampled and total RNA was extracted at 0 and 8 DAFB to investigate the gene expression during fruit set in 'pollinated' and 'non-pollinated' fruit (n=4).

3) Fruit load reduction study

In 'Golden Delicious Smoothee', trees were manually thinned by retaining only one lateral fruit from each fruit cluster at 11 DAFB. The lateral fruit was sampled in 'thinned' and 'control' group at 0, 10, 13 and 17 days after thinning (DAT) and used for total RNA extraction (n=4).

4) Shading study

In 'Golden Delicious Smoothee', the selected branch of each tree was covered with 80% shade black polypropylene during 15 to 25 DAFB. The base of shaded branch was girdled for both 'shaded' and 'control' group (n=4). The fruit was sampled at 0, 3, 7, and 10 days after shading (DAS) and used for total RNA extraction. In this study, only samples at 0 and 3 DAS were used for gene expression analysis.

#### Phylogenetic analysis

The protein sequence of the Arabidopsis *GRF/GIF*, *ARGOS*, and KLUH were retrieved from the Arabidopsis Information Resource (TAIR), and the tomato *FW2.2/CNR* sequence were retrieved from the NCBI database. The homologous genes in the apple genome were identified using BLAST in NCBI and Genome Database for Rosaceae using the *Arabidopsis* and tomato sequences as queries. The protein sequences of all genes within each family were used for multiple alignments using MUSCLE (Multiple Sequence Comparison by Log-Expectation) in MEGA7 (Molecular

Evolutionary Genetics Analysis). Phylogenetic tree of each family was constructed using the neighbor joining distance method in MEGA7.

#### **Primer Design**

Specific primers were designed by aligning genes with similar sequences using ApE (A plasmid Editor v2.0.53c) software. Both forward and reverse primers were selected to have at least 3-4 unique bases from the 3' end to ensure specific binding. Primers specificity was confirmed using NCBI primer blast, and only primers that were specific for the target gene were used. Primer specificity was also confirmed through melt-curve analyses at the end of the qRT-PCR amplification.

#### **Reverse transcription and qRT-PCR**

Reverse transcription was performed using 1  $\mu$ g of total RNA in a total volume of 20  $\mu$ L with ImProm II reverse transcriptase (Promega). Control reactions without template and without reverse transcriptase were included for the downstream analyses. The cDNA was diluted 6-fold, and 1 $\mu$ L of diluted cDNA was used for quantitative RT-PCR on the Stratagene Mx3005P (Agilent Technologies) quantitative real-time PCR instrument. PowerUp SYBR green master mix (ThermoFisher) was used for the analysis (Dash, Johnson et al. 2013). Three reference genes were used for normalization of target gene expression, *ACTIN, GAPDH* and *CACS2*. List of genes and their primer sequences for qRT-PCR analyses are presented in Supplementary Table 3.1. Efficiencies of the qPCR reactions were determined using LinRegPCR (Ruitjers *et al.*, 2009). Relative quantity (RQ) values were determined following efficiency correction and normalized using the geometric mean of RQs of three reference genes to generate normalized RQs (NRQs).

Data analyses were performed on the NRQ values following log<sub>2</sub> transformation. Standard errors were determined as described in Rieu and Powers(Rieu and Powers, 2009). The homologous genes were all quantified in four different experiments including fruit development, fruit set, fruit load reduction and shading studies. However, the homologous genes of KLUH were quantified only in the fruit development study and displayed little expression level during all stages, and thus were not quantified further in the other studies.

#### Statistical analyses

Statistical analyses and graph preparation were performed using RStudio (Version 1.0.143) and Inkscape (Version 0.92.3). Comparison of the transcript levels between thinned and control samples, and shaded and control fruit was performed using *Student's t-test* at each stage.

#### **Results and Discussion**

#### Identification of genes homologous to organ regulator genes

The sequence of FW2.2/*CNR* gene of tomato was used to search the homologous genes from the database of Apple Genome V1.0 Predicted CDS using BLAST in Genome Database for Rosaceae, leading to the identification of 32 candidate genes. These sequences were compared against those obtained from the NCBI database, retaining 24 candidate genes which were represented in both databases. Phylogenetic analysis was performed using these genes and the *CNR* genes from other species including peach, *Arabidopsis*, sweet cherry, grape, rice and soybean (Figure 3.1). A total of 11 *CNR* candidate genes in apple was finalized based on the phylogenetic analysis and their coverage and identity percentage with the tomato *FW2.2/CNR* gene. Among these genes, *CNR1* was the closest one with the tomato *FW2.2/CNR* gene, *CNR5a* and *CNR5b* were duplicated genes

located on chromosome 4 and 12, respectively. Although *CNR5a* and *CNR5b* are relatively distant from the tomato *FW2.2/CNR* gene, they clustered together with PavCNR20 in sweet cherry and PpCNR20 in peach. These genes were previously identified as located within a fruit size QTL (De Franceschi *et al.*, 2013). Another, fruit size QTL-associated CNR from sweet cherry PavCNR12 clustered with MdCNR6.

Nine GRFs have been identified in the Arabidopsis genome, and all of them harbor the conserved QLQ that binds with the GIF proteins, and WRC domains that are responsible for DNA binding and TF targeting. Each cDNA sequence of the 9 Arabidopsis GRF gene was used as a query in BLAST analysis against the database of Apple Genome V1.0 Predicted CDS using BLAST in the Genome Database for Rosaceae. The overlapped candidate genes from all alignments were considered to be the most potential apple GRF genes. These were then further compared to the sequences available from NCBI, leading to the identification of 14 apple GRF genes. A motif search using the motifFinder found that most of the apple GRF genes harbor both QLQ and WRC domains except for the *MdGRF6* and *MdGRF11*, which display only the WRC domain. Phylogenetic analysis indicated that most of the apple GRFs clustered with the Arabidopsis GRFs, especially with AtGRF1, 2 and 3 (Figure 3.2), which regulate leaf and cotyledon development (Kim et al., 2003). The GIFs constitute a relatively smaller gene family whose gene products co-activate the transcription of downstream genes along with GRF through an interaction with the conserved QLQ domains. In Arabidopsis, three GIF members play important roles in cell production during leaf and reproductive organ development (Kim and Kende, 2004). Four apple GIF genes was identified, and confirmed by BLAST analysis with the refseq RNA database in NCBI. As shown in the phylogenetic tree, *MdGIF3* and *MdGIF4* both

clustered within the same branch with *AtGIF1* while *MdGRF1* and *MdGIF2* displayed greater similarity with *AtGIF2* and *AtGIF3*(Figure 3.3).

ARGOS is relative small gene family, and only one ARGOS and one ARGOS-Like (ARL) have been functionally characterized in Arabidopsis. Three candidate ARGOS genes were identified in apple. Owing to their high sequence similarity, the three genes could not be differentiated as ARGOS or ARGOS-Like, thus the three candidates were named as MdARG1, MdARG2 and MdARG3. Phylogenetic analysis revealed that all three candidate genes were clustered together with the two from Arabidopsis (Figure 3.4). The three candidate genes also clustered with the predicted ARGOS candidates from peach, citrus and soybean derived from computational analysis (NCBI). MdARG1 and MdARG2 were used for downstream expression analysis.

Six candidate genes homologous to *AtKLUH* were identified in apple. Phylogenetic analysis indicated that four apple KLUHs were clustered together with AtKLUH (Figure 3.5). Two of them, *MDKLUH5* and *MdKLUH6* were not analyzed further owing to limited sequence similarity with *AtKLUH*.

In summary, a total of 35 genes homologous to the organ regulatory genes *CNRs, GRFs, GIFs, ARGOSs* and *KLUHs* were identified in the apple genome, based on sequence similarity to those functionally characterized in model species. In order to identify the most potential genes within each family in apple, the expression level of all these homologous genes were quantified in different studies. Only primers resulting in amplification of single amplicon as indicated by melt-curve analyses were used for gene expression analysis.

#### Transcript abundance of putative organ growth regulatory genes before bloom

Total RNA extracted from the carpel/floral tube at 26, 17, 7 and 0 days before full bloom (DBFB) in 'Gala' was used to investigate the expression of putative organ growth regulatory genes in apple. Growth in the carpel tissue mediated by cell production occurs between 26 and 7 DBFB and subsequently ceases until after bloom (Malladi and Johnson, 2011).

Three members of the *CNR* family were quantified in this experiment: *CNR5a*, *CNR6* and *CNR8*. All three members displayed elevated expression levels during the period of pre-bloom, reaching the highest level at 0 DBFB (Figure 3.6). The cell number increase by around 56% between 25 and 7 DBFB (Malladi and Johnson, 2011). Transcript abundance of *MdCNR5a*, *MdCNR6* and *MdCNR8* increased by 2.6, 1.2 and 2 fold respectively during this period. During the last week before full bloom, their levels continued to increase by 1.4, 1.2 and 3.2 fold in *MdCNR5a*, 6 and 8, respectively. Due to the cessation of cell production one week before full bloom, it is expected that genes inhibiting cell production are up-regulated between 7 to 0 DBFB or even before 7 DBFB. Taken together, it is likely that the increased transcript abundance of *MdCNR5* and *MdCNR5* and *MdCNR8* one week before full bloom is associated with the decline in cell production in carpel/floral tube. Interestingly, the inhibitor of cell production *MdKRP4* and *MdKRP5* also displayed elevated transcript abundance between 26 to 0 DBFB (Malladi and Johnson, 2011).

All members of apple *GRF* family except *MdGRF3* were analyzed in this study. *MdGRF4*, *MdGRF6* and *MdGRF11* displayed continuously elevated transcript abundance between 26-7 DBFB (Figure 3.6). On the other hand, we observed a declining abundance of *MdGRF1*, 2, 3, 7, 8, 12 transcripts between 26-7 DAFB. No dramatic alteration of the expression was noted with *MdGRF5*, *9*, *10* and *13* during the same period. During the last week before full bloom, *GRF* genes that displayed an increase in the transcript level >2-fold included *MdGRF1*, *2*, *8*, *10*, while the ones that displayed a decreased transcript level > 2-fold included *MdGRF3*, *7a* and *12*. These data suggest that *MdGRF3*, *7a*, and *12* act as positive regulators of cell production during carpel/floral tube development since their transcript levels decreased in parallel with the extent of cell production. Transcript abundance of *MdGIF2* was not substantially altered while *MdGIF1* increased gradually between 26-0 DBFB. Both *MdGIF3*, and *4* displayed a decline in transcript abundance between 26-7 DBFB, however, the expression level was not altered substantially during the last week before full bloom.

Both members of the *ARGOS/ARL* gene family displayed a similar expression pattern during the pre-bloom period. While the transcript levels of both *MdARG1* and *2* decreased by 5-fold from 26 DBFB to 7 DBFB, they were dramatically elevated by 18- and 10-fold, respectively, from 7 DBFB to 0 DBFB (Figure 3.6). It may be speculated that the abundance of these is upregulated dramatically before full bloom in response to changes in phytohormone concentration/signaling.

All four apple KLUH genes were quantified but the expression level was too low to be detected at any stage before full bloom.

#### Transcript abundance of putative organ growth regulatory genes during fruit set

Cell production in the carpel/floral tube ceased one week before full bloom in 'Gala', but such arrest was quickly released upon pollination and fertilization (Malladi and Johnson, 2011). Upon pollination, both cell size and number in the pollinated floral tube increased between 0 to 8 DAFB while little change occurred in the unfertilized flower which abscised around 14 DAFB (Malladi and Johnson, 2011).

Four members of the CNR gene family were investigated in response to pollination: MdCNR1, MdCNR5a, MdCNR6 and MdCNR8. Transcript abundance of MdCNR6 was upregulated by 1.7-fold between 0 and 8 DAFB in both pollinated and non-pollinated flowers. Similarly, the expression level of MdCNR8 was upregulated by 3- and 4-fold between 0 and 8 DAFB in both pollinated and non-pollinated flowers, respectively (Figure 3.7). Interestingly, CNR8 transcript level was >2-fold higher in the non-pollinated flowers at 8 DAFB, suggesting a potential role in negatively regulating cell production in non-pollinated flowers. Transcript abundance of MdCNR5a decreased by 50% in pollinated flowers between 0 to 8 DAFB, but remained unchanged in non-pollinated flowers, suggesting that the increase in cell production at 8 DAFB is facilitated by the down regulation of MdCNR5a transcript abundance in pollinated flowers. MdCNR1 displayed a similar reduction in transcript abundance as in MdCNR5a except that a similar decline was also noted in the non-pollinated flowers. Together, these data suggest that the MdCNR5a and MdCNR8 function as negative regulators of cell production in the developing fruit during fruit set. This is consistent with such a role for these gene products in other species such as tomato, maize and P. floridana (Frary et al., 2000; Guo et al., 2010; Li and He, 2015).

In pollinated flowers, the transcript abundance of *MdGRF1*, *MdGRF7a*, and *MdGRF8* increased by 4-, 4.3- and 1.8-fold between 0 and 8 DAFB (Figure 3.7). A similar increase in the abundance of these genes was also noted in the non-pollinated flowers, indicating little difference in the regulation of these genes during fruit set. *MdGRF2*, *11*, *12*, *13* displayed little change of transcript levels between 0 and 8 DAFB in either pollinated or non-pollinated flowers. Further, the transcript abundance of *MdGRF4*, *5*, *6*, *9*, *10* decreased in both pollinated and non-pollinated flowers. Several *GRFs* including *MdGRF5*, *MdGRF6* and *MdGRF10* displayed >2-fold higher transcript abundance in non-pollinated flowers at 8 DAFB. The transcript abundance of *MdGIF1* 

and 2 was largely unchanged between 0-8 DAFB in both pollinated and non-pollinated flowers (Figure 3.7). During the same period, the expression level of *MdGIF3* increased respectively by 3.2 and 2.9 fold in pollinated and non-pollinated flowers, but the expression level of *MdGIF4* decreased by 4.9 and 8.9 fold in pollinated and non-pollinated flowers, respectively. *MdGIF1* displayed >2-fold higher expression level in non-pollinated flowers at both 0 and 8 DAFB. Considering that similar patterns of transcript abundance were noted in pollinated and non-pollinated flowers, it is unlikely that the gene products of many of the *MdGRFs* and *MdGIFs* have a specific function in regulating growth during fruit set.

The transcript abundance of both *ARGOS* genes increased by over 2-fold between 0 and 8 DAFB in pollinated and non-pollinated flowers (Figure 3.7). But *MdARG1* displayed 1.4-fold higher expression in the pollinated flowers at 8 DAFB. It has been long recognized that auxin plays an important role in the initiation of growth and expansion following fertilization (Gustafson, 1939; Given *et al.*, 1988). It is possible that auxin biosynthesis upon or even before pollination induced the expression of *ARGOS*, and certain amount of auxin may have been established pre bloom based on the dramatically elevated *ARGOS* expression levels one week before bloom. Such high auxin level may be the reason why the non-pollinated flowers also displayed an increased *ARGOS* transcript abundance one week after bloom. In addition, the extra auxin derived from seeds may explain higher level of *MdARG1* in pollinated flowers at 8 DAFB.

### Transcript levels of organ growth regulatory genes during fruit development

Just as before full bloom, all four apple *KLUH* genes displayed little transcript abundance and could not be detected during most stages of fruit development. This was further confirmed through independent RNA-seq analysis in two different apple variety 'Golden Delicious

Smoothee' and 'Empire' (unpublished data, Jing and Malladi). These data suggest that *KLUH* transcript abundance is likely limited or restricted to specific cell types during apple fruit development. Further, these data suggest that KLUH may not have a major role in regulating apple fruit development unlike what was reported in tomato and Arabidopsis (Anastasiou *et al.*, 2010; Chakrabarti *et al.*, 2013).

Only four MdCNR homologous genes showed amplification during apple fruit development: MdCNCR1, MdCNR5a, MdCNR6 and MdCNR8. MdCNR1 and MdCNR8 both displayed their respective highest transcript levels at 0 DAFB, and then decreased between 0-18 DAFB (Figure 3.8). MdCNR1 transcript accumulation was too low to be detected between 18-124 DAFB. MdCNR8 transcript abundance was much lower during later stages compared to early fruit development. The transcript abundance of MdCNR6 was not altered dramatically during fruit development. MdCNR5a transcript accumulation was consistent with its potential role as a negative regulator of cell production. More specifically, the transcript level of MdCNR5a was relatively high at 0 DAFB and declined between 0-8 DAFB, consistent with data from the fruit set experiment, and remained extremely low until 21 DAFB. These data indicate decreased transcript abundance of MdCNR5a during the period of intense cell production. Subsequently, MdCNR5a transcript abundance increased concomitant with exit from the cell production phase of fruit development (Malladi and Johnson, 2011), and was maintained at relatively higher levels until 124 DAFB. Temporal changes in CNR transcript abundance has been proposed to be an important contributor to differences in cell production and organ growth across different tomato genotypes (Cong and Tanksley, 2002).

*MdGRF2*, *4*, *5* and *6* displayed respectively highest transcript abundance at 0 DAFB and decreased gradually to undetectable levels by 14 DAFB (Figure 3.8), indicating a rapid decline in

abundance during early fruit development. The abundance of their transcripts declined to undetectable levels prior to the decline in cell production during early fruit development (Malladi and Johnson, 2011), suggesting limited roles for their gene products in regulating this aspect of fruit growth in apple. Transcript accumulation of *MdGRF1* and *MdGRF7a* was relatively low level at 0 DAFB. MdGRF1 transcript abundance increased by 1.7-fold between 0-8 DAFB and then decreased gradually to undetectable levels by 18 DAFB. Transcript abundance of MdGRF7a increased dramatically by 3.3-fold between 0-11 DAFB before a sudden drop to undetectable levels by 18 DAFB. These data indicate an increase in MdGRF7a transcript abundance during the period of cell production followed by a decline around the period of exit from this phase during early fruit development, consistent with a role in positively regulating this trait during apple fruit development. MdGRF8 - MdGRF13 displayed a gradually decrease in transcript abundance during early fruit development. However, their abundance was also enhanced either during mid or late fruit development. All four *MdGIFs* displayed relatively higher transcript abundance during early fruit development than at later stages (Figure 3.8). MdGIF4 was the only one that did not show an initial increase of transcript levels between 0-8 DAFB. Instead, MdGIF4 displayed highest abundance at 0 DAFB, and declined dramatically between 0 to 8 DAFB and remained extremely low level afterwards. MdGIF3 displayed a similar expression pattern as in MdGRF1 and *MdGRF7a*: a sharp increase in transcript abundance around 4.4-fold between 0-8 DAFB before gradually decreasing until the end of study. The transcript levels of *MdGIF1* and 2 increased slightly between 0-8 DAFB, but decreased gradually by 21 DAFB and remained unchanged largely until the end of experiment. Hence, it is possible that MdGRF1, 7a and MdGIF3 gene products may play a role in promoting cell production similar to the *GRF/GIF* gene products in Arabidopsis. Functional analysis of these genes is essential to determine their specific roles in regulating cell

production during fruit growth. Further, it may be of interest to determine the specific interactions among the GRFs and GIFs, especially between MdGRF7a and MdGIF3.

Both *MdARG* genes displayed similar expression patterns during fruit development. They displayed two major increases (>3-fold) in transcript abundance: one between 8-14 DAFB, and another between 21-24 DAFB peaking at 24 DAFB (Figure 3.8). These peaks in abundance are coincident with period of high cell production and the subsequent exit from it into post-mitotic cell expansion. Such increases in ARGOS transcript abundance are potentially a reflection of auxin concentration during fruit development, partly because ARGOS is considered inducible by auxin in Arabidopsis and rice. Transcript abundance of MdARF106, a putative ARF associated with mediating auxin responses and a gene co-localized to a fruit size QTL, similarly displayed peaks in abundance during cell production and expansion phases of apple fruit development (Devoghalaere et al., 2012; Dash et al., 2013). Since auxins are known to play an important role along with gibberellins in controlling cell division and cell expansion, it is possible that the two peaks of ARGOS transcript abundance during fruit development are also associated with regulation of cell production and expansion (Csukasi et al., 2011, McAtee et al., 2013). It has been proposed that ARGOS regulates the activity ANT, which is a positive regulator of cell production in apple (Hu et al., 2003, Dash and Malladi, 2012). Hence, it is likely that a cell production promoting role for *MdARG* may be facilitated through the apple ANTs.

#### Transcript abundance of organ growth regulatory genes in response to reduction in fruit load

Based on the changes in transcript abundance of the potential organ growth regulatory genes during pre-bloom, fruit set, and fruit development, the following sub-set of genes were identified as potentially involved in regulating organ growth: *MdCNR5a, MdCNR6, MdCNR8, MdGRF1, MdGRF2, MdGRF7a, MdGRF8, MdGIF1, MdGIF3, MdARG1, MdARG2*. In order to investigate

the potential roles of these genes further, the transcript abundance of these genes was evaluated in a fruit load reduction experiment where all lateral fruits within each fruit cluster were removed at 11 DAFB (Dash et al., 2013). Fruit growth was increased due to elevated cell production in response to reduction in fruit load (Dash *et al.*, 2013). The transcript levels of the above 11 genes were quantified at 0 (11 DAFB), 10, 13 and 17 days after treatment (DAT) in both control and thinned fruits. At 0 DAT, the transcript abundance of MdCNR6 was significantly higher by 1.5fold in the thinned fruits compared to the control, but no significant difference was observed for MdCNR5a or MdCNR8 (Figure 3.9). The transcript levels of MdCNR5a and MdCNR8 were reduced significantly by 1.5- and 2.2-fold due to thinning at 10 DAT, and by 2.4- and 3.7- fold at 17 DAT, respectively. The declined transcript abundance due to thinning is consistent with their potential roles as negative regulators of cell production owing to the increase in this process in response to fruit load reduction (Malladi and Johnson, 2011). MdGRF1 displayed significantly elevated transcript abundance in response to fruit load reduction by 2-, 10- and 10- fold at 0, 13 and 17 DAT, respectively (Figure 3.9). Transcript abundance of the other three GRF members was increased in response to fruit load reduction by >2-fold though not significantly, especially at 13 and 17 DAT. MdGIF1 transcript abundance was not largely changed due to fruit load reduction (Figure 3.9). *MdGIF*3 displayed significantly higher transcript accumulation by 1.3- and 1.8- fold in the reduced fruit load treatment at 0 and 10 DAT, respectively. These data suggest that increased growth of thinned fruits through greater cell production was mediated by increased GRF and GIF transcript abundance. The transcript abundance of both MdARG1 and MdARG2 was downregulated significantly by 60% at 10 DAT in response to fruit load reduction (Figure 3.9). The effects of fruit load reduction on the auxin levels are not clear. The downregulation of both apple ARGOS may be associated with their potential role of promoting cell expansion. It's possible

that the cell expansion may be temporarily inhibited when cell production is promoted under reduced fruit load conditions, and the genes promoting cell expansion may be downregulated to sustain the enhanced cell production.

#### Transcript abundance of organ growth regulatory genes in response to shading

Severe shading in 'Gala' was performed around 17 DAFB, resulting in reduction of fruit growth due to the reduced photosynthate availability (Dash *et al.*, 2012). Cell number was significantly reduced in shaded fruit 3 days after shading (DAS). The transcript levels of the 11 genes potentially regulating cell production were analyzed in response to shading (Figure 3.10). Transcript levels in shaded fruit were increased significantly by 4-, 1.7-, and 4- fold at 3 DAS for *CNR5a*, *CNR6* and *CNR8*, respectively. *MdGRF7a* and *MdGRF8* displayed downregulated transcript abundance at 3 DAS in shaded fruit by around 60%. Transcript abundance of *MdGRF2* in shaded fruit increased by 4-fold at 3 DAS, while little effect of shading on *MdGRF1* transcript abundance was noted. A 3-fold increase in *MdGIF1* and a 40% reduction in *MdGIF3* transcript abundance was noted in shaded fruit at 3 DAT. The transcript abundance of both *MdARG* genes was elevated dramatically by around 5.8- and 4.3-fold at 3DAS.

#### Conclusions

The transcript accumulation patterns studied here during various stages of fruit development suggest that *MdCNR5a*, *MdCNR6* and *MdCNR8* act as cell production inhibitors owing to the negative association between their expression and cell production. *MdCNR5a* was the only one that was down regulated in pollinated flowers at 8 DAFB. In response to fruit load reduction, *MdCNR5a* and *MdCNR8* expression was reduced at multiple stages in fruit, further suggesting that

the increased fruit size in response to fruit load reduction may be due to the reduced inhibition of cell production. Under shading treatment, the expression level of all three apple *CNRs* was elevated, further supporting such a role. Most of the *GRF* and *GIF* genes displayed high transcript abundance during early fruit development, declined subsequently and became undetectable during later fruit development, consistent with the role of a positive regulator of cell production. Particularly, the transcript abundance of *MdGRF1*, *MdGRF7a*, *MdGRF8*, and *MdGIF3* was elevated by fruit load reduction but downregulated by shading. *MdARG1* and *MdARG2* shared high sequence similarity with each other, and displayed almost identical expression patterns in all experiments tested in this study. Both *MdARG* genes displayed high transcript abundance whenever cell production is low during fruit development, their abundance was downregulated under fruit load reduction, and elevated in shaded fruit. These data are consistent with potential negative regulation of cell production. Hence, further functional characterization is necessary to determine the potential role of the ARGs in regulating fruit development.

## References

Adamski, N. M., E. Anastasiou, S. Eriksson, C. M. O'Neill and M. Lenhard. 2009. Local maternal control of seed size by KLUH/CYP78A5-dependent growth signaling. <u>Proceedings of the National Academy of Sciences</u> **106**: 20115-20120.

Anastasiou, E., S. Kenz, M. Gerstung, D. MacLean, J. Timmer, C. Fleck and M. Lenhard. 2007. Control of plant organ size by KLUH/CYP78A5-dependent intercellular signaling. Developmental cell 13: 843-856.

**Bangerth, F.** 2000. Abscission and thinning of young fruit and thier regulation by plant hormones and bioregulators. <u>Plant Growth Regulation</u> **31**: 43-59.

Breuninger, H. and M. Lenhard. 2010. Control of tissue and organ growth in plants. <u>Current</u> topics in developmental biology, Elsevier. **91:** 185-220.

Chakrabarti, M., N. Zhang, C. Sauvage, S. Muños, J. Blanca, J. Cañizares, M. J. Diez, R. Schneider, M. Mazourek and J. McClead. 2013. A cytochrome P450 regulates a domestication trait in cultivated tomato. <u>Proceedings of the National Academy of Sciences</u> 110: 17125-17130.

**Cong, B. and S. D. Tanksley.** 2006. FW2.2 and cell cycle control in developing tomato fruit: a possible example of gene co-option in the evolution of a novel organ. <u>Plant Molecular Biology</u> **62**: 867-880.

Csukasi, F., S. Osorio, J. R. Gutierrez, J. Kitamura, P. Giavalisco, M. Nakajima, A. R. Fernie, J. P. Rathjen, M. A. Botella and V. Valpuesta. 2011. Gibberellin biosynthesis and signalling during development of the strawberry receptacle. <u>New Phytologist</u> **191**: 376-390.

**Dash, M., L. K. Johnson and A. Malladi.** 2013. Reduction of fruit load affects early fruit growth in apple by enhancing carbohydrate availability, altering the expression of cell production-related genes, and increasing cell production. Journal of the American Society for Horticultural Science 138: 253-262.

**Dash, M. and A. Malladi.** 2012. The AINTEGUMENTA genes, MdANT1 and MdANT2, are associated with the regulation of cell production during fruit growth in apple (Malus× domestica Borkh.). <u>BMC plant biology</u> **12**: 98.

**De Franceschi, P., T. Stegmeir, A. Cabrera, E. Van Der Knaap, U. Rosyara, A. Sebolt, L. Dondini, E. Dirlewanger, J. Quero-Garcia and J. Campoy.** 2013. Cell number regulator genes in Prunus provide candidate genes for the control of fruit size in sweet and sour cherry. <u>Molecular Breeding</u> **32**: 311-326.

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

Given, N., M. Venis and D. Gierson. 1988. Hormonal regulation of ripening in the strawberry, a non-climacteric fruit. <u>Planta</u> 174: 402-406.

Guo, M., M. A. Rupe, J. A. Dieter, J. Zou, D. Spielbauer, K. E. Duncan, R. J. Howard, Z. Hou and C. R. Simmons. 2010. Cell Number Regulator1 affects plant and organ size in maize: implications for crop yield enhancement and heterosis. <u>The Plant Cell</u> **22**: 1057-1073.

**Guo, M. and C. R. Simmons.** 2011. Cell number counts–the fw2. 2 and CNR genes and implications for controlling plant fruit and organ size. <u>Plant Science</u> **181**: 1-7.

Gustafson, F. G. 1939. Auxin distribution in fruits and its significance in fruit development. <u>American Journal of Botany</u> 26: 189-194.

**Horiguchi, G., G. T. Kim and H. Tsukaya.** 2005. The transcription factor AtGRF5 and the transcription coactivator AN3 regulate cell proliferation in leaf primordia of Arabidopsis thaliana. <u>The Plant Journal</u> **43**: 68-78.

Hu, Y., H. M. Poh and N. H. Chua. 2006. The Arabidopsis ARGOS-LIKE gene regulates cell expansion during organ growth. <u>The Plant Journal</u> **47**: 1-9.

Hu, Y., Q. Xie and N.-H. Chua. 2003. The Arabidopsis auxin-inducible gene ARGOS controls lateral organ size. <u>The Plant Cell</u> 15: 1951-1961.

Kalve, S., D. De Vos and G. T. Beemster. 2014. Leaf development: a cellular perspective. <u>Frontiers in plant science</u> **5**: 362.

Kim, J. H., D. Choi and H. Kende. 2003. The AtGRF family of putative transcription factors is involved in leaf and cotyledon growth in Arabidopsis. <u>The Plant Journal</u> **36**: 94-104.

**Kim, J. H. and H. Kende**. 2004. A transcriptional coactivator, AtGIF1, is involved in regulating leaf growth and morphology in Arabidopsis. <u>Proceedings of the National Academy of Sciences of the United States of America</u> **101**: 13374-13379.

**Kim, J. H. and B. H. Lee.** 2006. GROWTH-REGULATING FACTOR4 of Arabidopsis thaliana is required for development of leaves, cotyledons, and shoot apical meristem. <u>Journal of Plant</u> <u>Biology</u> **49**: 463-468.

Lee, B. H., J.-H. Ko, S. Lee, Y. Lee, J.-H. Pak and J. H. Kim. 2009. The Arabidopsis GRF-INTERACTING FACTOR gene family performs an overlapping function in determining organ size as well as multiple developmental properties. <u>Plant physiology</u> 151: 655-668. Li, Z. and C. He. 2014. Physalis floridana Cell Number Regulator1 encodes a cell membraneanchored modulator of cell cycle and negatively controls fruit size. <u>Journal of experimental</u> <u>botany</u> 66: 257-270. Liu, H., S. Guo, Y. Xu, C. Li, Z. Zhang, D. Zhang, S. Xu, C. Zhang and K. Chong. 2014. OsmiR396d-regulated OsGRFs function in floral organogenesis in rice through binding to their targets OsJMJ706 and OsCR4. <u>Plant physiology</u> **165**: 160-174.

**Malladi, A. and P. M. Hirst.** 2010. Increase in fruit size of a spontaneous mutant of 'Gala'apple (Malus× domestica Borkh.) is facilitated by altered cell production and enhanced cell size. Journal of experimental botany **61**(: 3003-3013.

**Malladi, A. and L. K. Johnson.** 2011. Expression profiling of cell cycle genes reveals key facilitators of cell production during carpel development, fruit set, and fruit growth in apple (Malus× domestica Borkh.). Journal of experimental botany **62**: 205-219.

McAtee, P., S. Karim, R. J. Schaffer and K. David. 2013. A dynamic interplay between phytohormones is required for fruit development, maturation, and ripening. <u>Frontiers in Plant</u> <u>Science</u> **4**: 79.

**Omidbakhshfard, M. A., S. Proost, U. Fujikura and B. Mueller-Roeber.** 2015. Growthregulating factors (GRFs): a small transcription factor family with important functions in plant biology. <u>Molecular plant</u> **8**: 998-1010.

**Rieu, I. and S. J. Powers.** 2009. Real-time quantitative RT-PCR: design, calculations, and statistics. <u>The Plant Cell</u> **21**: 1031-1033.

Rodriguez, R. E., M. A. Mecchia, J. M. Debernardi, C. Schommer, D. Weigel and J. F. Palatnik. 2010. Control of cell proliferation in Arabidopsis thaliana by microRNA miR396. Development 137: 103-112.

Wang, B., Y. Sang, J. Song, X.-Q. Gao and X. Zhang. 2009. Expression of a rice OsARGOS gene in Arabidopsis promotes cell division and expansion and increases organ size. Journal of <u>Genetics and Genomics</u> 36: 31-40.

**Table 3.1.** List of the homologous genes of *CNR*, *GRF/GIF*, *ARGOS* and *KLUH* in apple and the sequence of primers used in quantitative RT-PCR analyses.

Gene	Accession Number	Primer Orientation	Primer Sequence (5'-3')
MdACTIN	EB127077	Forward	ACCATCTGCAACTCATCCGAACCT
		Reverse	ACAATGCTAGGGAACACGGCTCTT
MdGAPDH	EB146750	Forward	TGAGGGCAAGCTGAAGGGTATCTT
		Reverse	TCAAGTCAACCACACGGGTACTGT
MdGIF1	MDP0000297641	Forward	CCATGCATCATACTGAGGGCTCATT
		Reverse	CCTCTCCATCACCGTTGCCAT
MdGIF2	NM_001293869.1	Forward	CATGCTTCATACTGAGGCCACACA
		Reverse	CGTCCTCTCCATCACCATTTCCAC
MdGIF3	xm_008361670.1	Forward	ACCGGCATGATGCAGCCA
		Reverse	CTCATTCCCGAGCATGTGCAGA
MdGIF4	xm_008372504.1	Forward	TCCACCGGTATGATGCAACCG
		Reverse	CCTCATTCCCGAGCATGTGAAGT
MdARG1	MDP0000744273	Forward	CGACTCTGCTCATCCTTCCTTTGATA G
		Reverse	CCAAGATCATAAGCACGCCCAG
MdARG2	MDP0000255770	Forward	GACTCTGCTGATCCTTCCTTTGGTAC
		Reverse	AAGCCAAGATCATAAGCACGGCTAA
MdCNR1	xm_008374937.1	Forward	AGCCAAGGTTCCACACCTTGTG
		Reverse	AGCAGAAGTGCACTAGGCAATCC
MdCNR5a	XM_008371558.1	Forward	ATGGCGACCAACAACAGGGAGAGC
		Reverse	CTCCATTCACATGCCCCACATCCGTC
MdCNR6	XM_008377817.1	Forward	GCGGTATGTGAAGCTGACGAAAGAA
		Reverse	GGGTAAAGGCTGTCCACATTCGTT
MdCNR8	xm_008361807.1	Forward	TGTCCATGCATCACCTTTGGCC
		Reverse	TCATTTTGGAGCGGTAGAAGCAGG
MdGRF1	XM_008338802_1	Forward	GGCAAAGGACCTTTCACTCCATCT
		Reverse	GGGCCTTCTTGATAGGGACGAGT
MdGRF2	XM_008343878_1	Forward	GCAGTGGTGTATCCAGCAGCC
		Reverse	GCAGACATCACGGAAACTCCCTG
MdGRF4	XM_008345814_1	Forward	AGAAACAGCCGGTTTCCTTTCACTC
		Reverse	AGGAGGGTGGTGAGAGAAGAGC
MdGRF5	XM_008358341_1	Forward	AGGTGTTCCAGAGATGTAGCTCCG
		Reverse	GACTTTGCTGGTTGGTGATGTTGC

Gene	Accession Number	Primer Orientation	Primer Sequence (5'-3')
MdGRF6	XM_008358803_1	Forward	AGTGAACCTCGACAGCAACAAAATG AC
		Reverse	ACCTCCATCCTCTTCCGTTAACTCT
MdGRF7a	XM_008362323_1	Forward	ACGAGGTGCAGGAGGACAGATG
		Reverse	TGGTTGTGGTTGTCACTTCCCA
MdGRF8	XM_008367363_1	Forward	CAGAGGACCTTTCACGCCATCC
		Reverse	AAGAGCCTTCTTGATTGGGAAGAGC
MdGRF9	XM_008375233_1	Forward	ATTCAAGTATCTCAAGGCAGGAGTCC
		Reverse	GTCTATCTTCTTCCCGCAATAGGAAG C
MdGRF10	XM_008378582_1	Forward	TGATGGTCCATCATGATAATCACCGC
		Reverse	CGCCACTCGAAGAACCAGTACC
<i>MdGRF11</i>	XM_008389001_1	Forward	AAAATGTACCTCGACAGCAACAAGG
		Reverse	CCTCCATCCTCTTCCATTAACGCG
MdGRF12	XM_008393169_1	Forward	AGCAACTCAGTGGCAAGAGCTAG
		Reverse	TGCCAAATCCCATCTCATAACACCC
MdGRF13	XM_008394413_1	Forward	CGATGCATGGTCCACGTTAGCA
		Reverse	ACCCAAGACATGGGGTTCGC
<i>MdKLUH1</i>	MDP0000270602	Forward	CCTTCCGGCTTTCCTATTCTTGGT
		Reverse	GCTGTTCAAGAGCTCTTTTGCAGTAT TG
MdKLUH2	xm_008352667.1	Forward	GCGTATGATGTTGAAGAACGGCCAA
		Reverse	AGCTCATACCCTTCAGTCACCAGTT
MdKLUH3	xm_008354732.1	Forward	CGTTTGATGACGGAGAACGGTGAG
		Reverse	AGTAGTTCATACCCTTCAGTCACCAG AC
MdKLUH4	xm_008376432.1	Forward	TCCGGCCTTCCTCTTCTAGGG
		Reverse	CTGTTCAAGAGCTCTTTCGCTGTATC A



Figure 3.1. *Phylogenetic tree of CNR gene family derived from protein sequences.* 

The relationships among MdCNR in tomato (Sl), soybean (Gm), Arabidopsis (At), peach (Pp), sweetcherry (Pav), grape (Vv), rice (Os), maize (Zm) and medicago (Mt) are displayed. The original gene used for blast against is labeled as green, and the apple CNRs are labeled as red.



Figure 3.2. Phylogenetic tree of GRF gene family derived from protein sequences.

The relationships among MdGRFs in woodland strawberry (Fv), tomato (Sl), soybean (Gm), Arabidopsis (At), peach (Pp), grape (Vv), rice (Os) and maize (Zm) are displayed. The original genes used for blast against are labeled as green, and the apple GRFs are labeled as red.



Figure 3.3. Phylogenetic tree of GIF gene family derived from protein sequences.

The relationships among MdGIF genes in Arabidopsis (At), peach (Pp), grape (Vv), pepper (Ca) and maize (Zm) are displayed. The original genes used for blast against are labeled as green, and the apple GIF genes are labeled as red.


Figure 3.4. *Phylogenetic tree of ARG gene family derived from protein sequences.* 

The relationships among MdARGs in Arabidopsis (At), peach (Pp), soybean (Gm), citrus (Cs), wild rice (Ob), grape (Vv) and maize (Zm) are displayed. The original genes used for blast against are labeled as green, and the apple ARGOS are labeled as red.



Figure 3.5. *Phylogenetic tree of KLUH gene family derived from protein sequences.* 

The relationships among MdKLUHs in Arabidopsis (At), soybean (Gm), peach (Pp), grape (Vv), tomato (Sl) and woodland strawberry (Fv) are displayed. The original genes used for blast against are labeled as green, and the apple KLUHs are labeled as red.



**Figure 3.6.** *Expression profiles of the homologous genes of CNR, GRF/GIF and ARGOS between 26 and 0 days before full bloom.* 

Relative expression levels to its expression at 26 DBFB is shown for each gene. The raw transcription levels are scaled in R for heatmap analysis.



**Figure 3.7.** *Expression profiles of the homologous genes of CNR, GRF/GIF and ARGOS in pollinated (P) and non-pollinated (NP) fruits at 0 and 8 days after full bloom.* 

Relative expression levels to its expression in pollinated flower at 0 DAFB is shown for each gene. The raw transcription levels are scaled in R for heatmap analysis.



**Figure 3.8.** *Expression profiles of the homologous genes of CNR, GRF/GIF and ARGOS during fruit development.* 

Relative expression levels to its expression at 0 DAFB is shown for each gene. The raw transcription levels are scaled in R for heatmap analysis.



**Figure 3.9.** *Expression profiles of the homologous genes of CNR, GRF/GIF and ARGOS in thinned and control fruits at 0, 10, 13 and 17 days after thinning.* 

Relative expression levels to its expression in thinned fruit at 0 DAFB is shown for each gene. White bars represent 'Control' fruits and black bars represent 'Thinned' fruits. Error bars indicate the standard error of the means (n=4). All the difference of expression levels of genes between thinned and control fruit is indicated by 1, 2 or 3, asterisks representing significant level at  $\alpha = 0.05$ , 0.01, 0.001, respectively.



**Figure 3.10.** *Expression profiles of the homologous genes of CNR, GRF/GIF and ARGOS in shaded and control fruits at 0, and 3 days after shading.* 

Relative expression levels to its expression in control fruit at 0 DAS is shown for each gene. White bars represent 'Control' fruits and black bars represent 'Shaded' fruits. Error bars indicate the standard error of the means (n=4). All the difference of expression levels of genes between shaded and control fruit is indicated by 1 or 2 asterisks representing significant level at  $\alpha = 0.05$ , 0.01, respectively.

## CHAPTER 4

# TRANSCRIPTIONAL PROFILING REVEALS GENE NETWORKS REGULATING APPLE

## FRUIT DEVELOPMENT

Jing, S., Wang YW., Ji, LX., Nambeesan, S., Loraine, A. and Malladi, A. To be submitted to

BMC Plant Biology.

## Abstract

Apple fruit has a unique structure which includes the edible accessory fruit (cortex) and the true fruit (pith). These tissues of diverse origins display different growth characteristics and likely involves spatially different molecular mechanisms. Growth during different stages of fruit development is supported by diverse processes. Progression through these stages and transitions between them requires temporal mechanisms for its regulation. In order to understand the potential molecular mechanisms regulating spatio-temporal aspects of apple fruit growth, we performed transcriptome analysis of the cortex and pith tissues at different phases, and in response to fruit load reduction. In 'Golden Delicious Smoothee', fruit load reduction was performed at 11 days after full bloom, which resulted to increased fruit size. The cortex displayed significantly higher cell number and area compared to the pith during most stages of fruit development. The pith and cortex tissues at 8, 19 and 47 days after treatment (DAT) was used for pair-end RNA-seq analysis. The average sequencing depth of all samples was around 20 million reads. The largest variation in the transcriptomes observed was between 8, and 19 & 47 DAT, indicating distinct molecular mechanisms involved in regulating early fruit growth. Within each stage, there was a distinct difference of transcriptome profile between the pith and cortex tissues, suggesting tissue specific regulatory networks. Little alteration of the transcriptome was observed in response to fruit load reduction, especially at 8 and 47 DAT. A weighted gene co-expression network analysis (WGCNA) revealed a module containing a group of genes that are significantly correlated with Relative Tissue Growth Rate (RTGR). A putative TCP, (MdTCP2b MD05G1281100) was identified as one of the genes within this module with high correlation to RTGR. TCPs within the clade to which MdTCP2b was clustered are known to be involved in a regulatory cascade involving miR319 and miR396, GRFs and KRPs, which negatively regulates cell production during organ growth. The

transcript accumulation patterns of *MdTCP2b*, *MdGRF7a*, *MdGIF3*, and *MdKRP4* during fruit development are consistent with the presence of a similar regulatory pathway during apple fruit growth. Further, the presence of miRNA target sites in *MdTCP2b*, *GRF7a* and a potential TCP binding site in the promoter of *KRP4* are consistent with the operation of this pathway in apple fruit growth. Hence, it is proposed that this regulatory pathway contributes to the spatio-temporal regulation of growth during apple fruit development.

Keywords: Malus × domestica, fruit development, RNA-seq, tissue-specific, WGCNA, TCP

## Introduction

Apple (Malus × domestica Borkh) fruit is a typical pome which includes the true fruit derived from ovary (pith) and the accessory fruit (cortex) derived from the fused regions of the basal part of floral organs including sepals, petals and stamen (MacDaniels 1940). The cortex region in domesticated apple has increased greatly in size compared to their wild counterparts after long history of domestication, due to more extensive growth in the cortex driven by a combination of cell production and expansion. During fruit development, the cortex displays preferential growth in comparison to the pith and occupies greater than 85% of the fruit volume at maturity (Tukey and Young, 1942; Goffinet et al., 1995; Malladi, unpublished data). Further, tissue porosity and void space development were substantially different between the cortex and pith (Herremans et al., 2015), indicating differential growth characteristics across these tissues. Despite different origins and growth patterns of the cortex and pith, fruit development involves a coordination of enlargement in both types of tissues, which is regulated by unknown molecular mechanisms. Distinctive metabolic characteristics of the cortex and pith during apple fruit development have been identified. For example, the concentration of the translocated form of C and N, sorbitol and asparagine, is largely different between the true and accessory fruit, especially during early and mid-fruit development. In terms of the total content of the metabolites, more sugars, organic acids and starch accumulated in the cortex than the pith. In addition, clear shifts in metabolic profiles have been observed in fruit during different developmental stages in apple, similar to that reported in many other species such as tomato, peach, strawberry, and grape (Carrari et al., 2006; Zhang et al., 2010; Lombardo et al., 2011; Dai et al., 2013). Such spatial and temporal metabolic programs may facilitate the different growth patterns observed across these tissue types. The molecular

mechanisms mediating these spatial and temporal characteristics of growth have not yet been investigated in apple.

Apple fruit growth is initially driven by cell production during the first 3-6 weeks after fertilization, and subsequently mediated by post-mitotic cell expansion later until growth ceases (Malladi and Hirst, 2010). During the first 3-6 weeks after pollination, the number of cells in the cortex increases by 8- to10-fold, and the total cell volume increases by around 20-fold before cell production ceases (Dash et al., 2012). During later stages of development, the total cell volume increases by up to 1500-fold before harvest. A sub set of cell cycle genes, cell expansion related genes and sugar metabolism related genes have been identified in apple, and the functional characterization of some genes has been performed by expressing some of these candidate genes in Arabidopsis. The transcription factor, AINTEGUMENTA, may contribute to the regulation of apple fruit development by regulating cell production, potentially in association with changes in CDK-CYC complexes including MdCYCA2;3, MdCYCB1;1, MdCDKB1;2, MdCDKB2;1 and MdDEL1 (Dash and Malladi, 2012). Two negative regulators of cell production, MdKRP4 and MdKRP5 displayed negative regulation of cell production (Malladi and Johnson, 2011). Expression of these negative regulators in Arabidopsis under a constitutive promoter reduced leaf size and increased leaf serration (Johnson, 2013), consistent with a phenotype displayed by AtKRP2 overexpressing lines (De Veylder et al., 2001). Several a-type Expansin (MdEXPA) and COBRA(MdCOB1) genes displayed potential association with cell expansion during apple fruit development (Dash, Johnson et al. 2012).

Although many genes regulating fruit development have been identified and characterized in apple, a comprehensive understanding of regulatory molecular mechanisms is still lacking. Multiple transcriptomic studies on fruit development have been performed in species belonging to Rosaceae family including strawberry, peach, pear and etc.(Kang *et al.*, 2013; Xie *et al.*, 2013; Hollender *et al.*, 2014; Sánchez-Sevilla *et al.*, 2017; Gu *et al.*, 2018; Hu *et al.*, 2018). The woodland strawberry has a unique fruit structure that the true fruit or the achenes dot the surface of accessory fruit developed from the receptacle. Spatial and temporal RNA-seq analysis was performed in both flower and fruits of woodland strawberry pre- and post- fertilization, leading to the identification of hub genes specifically regulating receptacle development through Weighted Gene Correlation Network Analysis. The endosperm and seed coat have been shown to play important role in receptacle development by regulating auxin and gibberellin biosynthesis (Kang *et al.*, 2013). In peach, differentially expressed genes were identified during the transition pre and post the first and second fruit swelling stages of fruit, and the top of these DEGs were found to be regulators of both cell proliferation and expansion. A crucial gene network mediating the swelling of fruit was also suggested in peach through RNA-seq analysis(Gu *et al.*, 2018).

Owing to the nature of fruit tissue development in apple, it may be hypothesized that distinct molecular mechanisms mediate differential growth in the cortex and pith tissues. The objective of this study was to determine the overall transcriptomic variations in the cortex and pith across different developmental stages. In this study, fruit undergoing different growth phases (cell production, transition away from cell production, and cell expansion) were analyzed. In addition, fruit load reduction, a common horticultural practice to reduce fruit load and increase size in apple, was performed to determine transcriptomic alterations mediating enhanced fruit growth.

## **Material and Methods**

## **Plant materials**

The study was performed in the apple cultivar 'Golden Delicious Smoothee' grown on M7 rootstocks. Trees were in their adult phase, and maintained at the Mountain Research and Education Center, University of Georgia in Blairsville, Georgia. Trees used in this study were maintained according to commercial apple production practices, except that no chemical fruit load reduction was performed in the trees. Eight trees were selected and four of them were assigned for the fruit load reduction treatment and another four were assigned as control. For fruit load reduction treatment, all lateral fruits of each cluster were removed at 11 days after full bloom (DAFB), leaving only the king fruit. For each tree, ten king fruit were tagged and the diameter and length of these fruit were recorded at the following stages: 0, 8, 12, 19, 26, 33, 47, 77, and 118 days after fruit load reduction treatment (DAT). In addition, four king fruit were sampled from each tree at each stage, which were peeled and cut in half longitudinally with the seeds removed. One half of the fruit was fixed in CRAF fixative (Chromic acid: Acetic acid: formalin) for cytological analysis, the other half was used for molecular analysis. For the latter, the pith and cortex tissue was separated using a biopsy punch and frozen in liquid nitrogen immediately. Henceforth, the control and reduced fruit load group was referred as CL and RL, respectively. Abbreviations for the pith and cortex of the control and reduced fruit load group are used as follows: CP (control pith), RP (reduced fruit load pith), CC (control cortex), RC (reduced fruit load cortex).

## Measurement of cell size and number in pith and cortex

Longitudinally sliced fruit fixed in CRAF was scanned using a flatbed scanner (V600, Epson), and ImageJ (V1.52a; National Institutes of Health) was used to measure the pith and cortex area. The

cortex area was calculated by subtracting the core from the fruit, and the pith area was obtained by subtracting the seed locules from the core. The central part of the pith and cortex was cut to a around 1 cm<sup>2</sup> and embedded in the Optimum cutting temperature (OCT; Ted Pella, Inc.) compound for sectioning using a cryostat (Leica Jung Frigocut 2800N, Germany). Section thickness ranged from 10 µm to around 50 µm depending on the stage of fruit development. The sections were stained with 0.1% toluidine blue, and images were obtained with a microscope (BX51, Olympus equipped with DP70 camera). For each fruit, one section each was obtained for the pith and the cortex tissues, respectively. The cell number and area of the pith and cortex was measured as follows. In the cortex, around 50 cells were selected randomly and the area of each cell was calculated using particle analysis in ImageJ. The average cell area was divided by the cortex tissue area to obtain the cell number per tissue section in the cortex (referred to henceforth as the cell number). As fruit growth progressed, voids development was initiated as reported previously (Herremans et al., 2015). The average voids percentage were calculated from 6 random sections from each stage, assuming that the percent voids were similar between the CL and RL fruit. The void space area was removed from the tissue area for cell number calculation in the cortex. Little void development was observed in the pith within the period of development reported here. However, the pith displayed large variations in cell area among parts of the tissue. Cell area in the mid and peripheral regions of the pith were measured separately, and the proportion of these regions in relation to the total pith area was calculated. The cell area in the pith was calculated by (Average cell area of region 1 x proportional area of region 1) + (Average cell area of region 2 x proportional area of region 2). The cell area in the pith was divided by the pith area to obtain the pith cell number. After obtaining the tissue area, cell number and cell area in the pith and cortex,

relative tissue growth rate (RTGR), relative cell production rate (RCPR) and relative cell expansion rate (RCER) was calculated by the following formula:

 $RTGR = [Ln (Tissue Area_2)-Ln (Tissue Area_1)]/(T_2-T_1)$  $RCPR = [Ln (Cell Number_2)-Ln (Cell Number_1)]/(T_2-T_1)$  $RCER = [Ln (Cell Area_2) - Ln (Cell Area_1)]/(T_2-T_1)$ 

## RNA extraction and RNA-seq library construction

Total RNA was extracted from the pith and cortex tissue from 9 stages using the CTAB based extraction method (Vashisth *et al.*, 2011). RNA-seq analysis was performed for CP, RP, CC, RC at three stages: 8, 19, and 47 DAT. For each stage, three biological replications for each condition were selected for library preparation. A total of 36 RNA-seq libraries were constructed using KAPA stranded mRNA-seq kit, and the quality of libraries were examined using Bio-analyzer (Agilent). Sequencing was performed using Illumina NextSeq500 with pair-end reads of 75 bp at the Georgia Genomics and Bioinformatics Core, University of Georgia.

#### *RNA-seq data analysis*

After obtaining the 36 FASTQ files, the adapters of the raw reads were initially trimmed using 'Trimmomatic' with the following parameters: LEADING:3, TRAILING:3, SLIDINWINDOW:4:15, MINLEN:30. The trimmed reads were aligned to the apple reference genome obtained from the GDDH13 Version 1.1 database (Genome Database for Rosaceae) using Tophat2 with the following parameters: Minimum intron length: 30, Maximum intron length: 30000. The FeatureCounts program was used to count the number of reads aligned to the apple reference genome, which generated three types of read counts: single-mapping without multi-

mapping ("sm"), single-mapping with multi-mapping ("mm"), and primary alignment reads only ("pm")(Liao, Smyth et al. 2013). In this study, only single-mapping ("sm") reads that do not include multi-mapping reads were used for subsequent data analysis. Expression levels in transcript per million (TPM) were calculated from the single-mapping reads(Wagner *et al.*, 2012). Genes with no count number or TPM value were removed from further data analysis. TPM value was transformed using "scale" in R. Hierarchical clustering analysis was performed for the average scaled TPM values using heatmap.2 function in R.

EdgeR was used to identify the differentially expressed genes (DEGs)(Robinson *et al.*, 2010). Likelihood ratio test was performed for a pairwise comparison using glmLRT function to identify the DEGs between the pith and cortex, and between RL and CL. The raw counts ("sm") were used for DEG analysis, and transcripts that have 0 total count in 36 conditions were removed for DEG analysis. The adjusted p value or false discovery rate (FDR) were calculated using the method of Benjamini and Hochberg (1995). The cutoff FDR of 0.005 was applied for all comparisons. Furthermore, counts per million (CPM) were calculated and converted using log2 transformation. A cutoff of 1 (2-fold) was applied for all comparisons. The DEG identified using EdgeR were rendered for gene ontology enrichment analysis using GoMapMan analysis. The DEGs were classified into GoMapMan functional plant categories/bins by annotation of their protein sequences using the tool Mercator. The database used for BLAST analysis includes the TAIR Release 10 (TAIR), SWissProt/UniProt Plant Proteins (PPAP), and Clusters of orthologous eukaryotic genes database (KOG) provided by Mercator, and the blast cutoff of 80 was used.

Weighted gene co-expression network analysis (WGCNA) was used to investigate the coexpression among genes and identify genes associated with tissue area and Relative Tissue Growth Rate(Langfelder and Horvath, 2008). Briefly, all genes with high correlation among each other were clustered into the same module, and the correlation of each module to the trait of interest is quantified. The module displaying high correlation with the trait of interest is analyzed further to narrow down a list of potential genes controlling the trait of interest. In this study, only genes with an average TPM value among all conditions greater than 1 were applied for the WGCNA analysis. The log2 transformation for the TPM plus one value was used for WGCNA analysis. A step-by-step network constructions and module detection was performed with the following parameters: a power of 14, a minimal module size of 100, and a branch merge cut off of 0.2. After all genes were classified into different modules, the correlation between the module eigen value and RTGR was evaluated to identify module associated with this trait. Among the genes within the interested module, genes were ordered based on their gene significance and module membership values to select the hub genes associated with RTGR.

#### Real-time quantitative RT-PCR analysis

Quantitative RT-PCR analysis was performed on genes derived from WGCNA analysis, and other candidate genes that were potentially involved in the regulation of cell production. For all genes, the expression level in the pith and cortex of CL and RL fruit was quantified during 6 developmental stages: 0, 8, 19, 26, 47 and 118 DAT. For each sample, 4 biological replications were included. The primers were designed using the plasmid editor ApE (v2.9.55), and the specificity of primers were checked using NCBI Primer-Blast. Reverse transcription and gene expression analysis was done as described previously (Dash and Malladi, 2012) using Stratagene Mx3005P with the exception of the usage of PowerUp SYBR green master mix (ThermoFisher). At the end of PCR amplification, melt-curve analysis was performed to validate the specificity of

primers. Control reactions without the template and without reverse transcriptase were included in the analyses. *MdACTIN*, *MdGAPDH* and *MdCACS2* were used as the reference genes for normalization of target gene expression. The efficiencies of the reactions were determined using LinRegPCR (Ruitjers *et al.*, 2009). Relative quantity (RQ) values were determined and normalized to those of the reference genes to generate the normalized relative quantities (NRQs). The geometric mean of the RQ values of the three reference genes was used for normalization. Data analysis were performed on the NRQ values after log<sub>2</sub> transformation. Expression of all genes is presented as fold change in relation to the mean expression of the target gene in the cortex of fruit from the reduced fruit load treatment at 0 DAT. Standard errors were determined as described in Rieu and Powers (2009).

#### **Results and Discussion**

#### Greater cell production and expansion facilitate higher RTGR in the cortex

Relative tissue growth rate (RTGR) was consistently higher in the cortex than the pith during early fruit development, indicating that greater fruit growth occurred in the cortex (Figure 4.1). RTGR in the cortex increased by 1.8-fold in RL fruit while it remained unchanged in CL fruit from 8 DAT to 12 DAT (Figure 4.1B). The RTGR in the pith decreased by 50% in the control fruit while remained unchanged in the thinned fruit during the same period (Figure 4.1C). RTGR in the cortex decreased continuously after 12 DAT. The pith RTGR increased abruptly between 12 DAT to 19 DAT and then declined to almost zero by 77 DAT. High RTGR in the cortex during early fruit development suggests that the fruit growth is more rapid and intensive during early fruit development specifically in the cortex, although this period was much shorter compared to late

fruit development. However, fruit load reduction did not affect the RTGR in general except for 20% lower RTGR in the cortex at 8 DAT, and 31% higher in the cortex at 26 DAT.

Cell number in the pith did not change dramatically between 0 and 47 DAT, potentially associated with the limited growth of the pith (Figure 4.2A). Although cell number in the cortex was only slightly higher than that in the pith at 0 DAT, the difference increased continuously between 0 and 12 DAT reaching 3-fold higher levels in the cortex. Subsequently, the difference in cell number was relatively stable until 33 DAT after which a slight increase in cell number occurred in the cortex. Cell number in the pith was not altered due to fruit load reduction while it increased significantly in the cortex at 12 and 47 DAT. Cell area remained similar between the pith and cortex until 12 DAT (Figure 4.2B). Between 12 and 19 DAT, cell area increased in the cortex and pith by 3- and 2- fold, respectively. Cell area continued to increase in both types of tissue until 47 DAT. Fruit load reduction decreased cell area by 13% in the cortex at 8 DAT and by 10% in the pith at 47 DAT. The RCPR was relatively low in the pith at all stages, consistent with relatively smaller changes in pith cell number (Figure 4.2C). RCPR in the cortex increased by around 8% between 0 and 8 DAT, and by around 10% between 8 and 12 DAT, suggesting a rapid cell production between 0 and 12 DAT in the cortex. The RCER was relatively low in both cortex and pith until 12 DAT and increased greatly by 19 DAT (Figure 4.2D). The RCER slowed down gradually in both tissues after 19 DAT, reaching basal levels by 47 DAT. Overall, little growth occurred in the pith during fruit development due to relatively low extent of cell production. Increase in tissue growth and RTGR was evident at the onset of post-mitotic cell expansion in the pith (between 12-19 DAT). However, considering the large variation in cell size across different parts of the pith it may be likely that further sub-spatial differences in cell production and expansion occur within this tissue. Intensive growth of the cortex during early fruit development is associated extensively with cell production up to around 12 DAT. These data indicate clear spatial differences in growth within the apple fruit. These data also suggest molecular mechanisms specific to the cortex that allow for rapid cell production during early fruit development. Subsequently, growth was supported by increase in cell size from around 19 DAT. Fruit load reduction did not significantly alter the RTGR, but increased cell production specifically in the cortex. These data suggest that enhanced availability of resources was specifically associated with increased cell production as has been reported previously (Dash and Malladi, 2102; Dash *et al.*, 2013). The data from this study demonstrate that this occurs specifically in the cortex.

## General features of the fruit tissue transcriptomes

Three stages were selected for RNA-seq analysis: 8 DAT, 19 DAT, and 47 DAT, to focus on the genes regulating the following stages: cell production, the transition from cell production to cell expansion, and cell expansion, respectively. A total of 781.7 million reads were retained after adapter trimming of 808.5 million RNA-seq raw reads. An average reads number of 22.4 million was obtained for all samples at 19 and 47 DAT, while around 11.1 million reads were obtained averagely for samples at 8 DAT (Figure 4.3A). Greater than 91% of the trimmed reads were mapped to the apple reference genome uniquely (Table S4.1). A total of 44450 genes of them were identified with the TPM values greater than zero, and 30604 of them were identified with an average TPM greater than 1. Multidimensional scaling was performed for all genes with TPM value greater than zero to explore global similarities among the 36 transcriptomes. All biological replications for each condition clustered closely, with the exception of 47CP and 19CP (Figure 4.3C) indicating generally lesser variation across the replications. Further, this analysis indicated clear temporal and spatial separation of the transcriptomes, while fruit load reduction did not allow

for such separation of the transcriptomes (Figure 4.3C). For CC, 26312 genes were commonly expressed across 8, 19 and 47 DAT. Around 1159, 376 and 891 genes were specifically expressed in the CC at 8, 19 and 47 DAT, respectively, suggesting distinctive characteristics of transcriptomic profiles during each developmental stage (Figure 4.3B). The number of genes with an average TPM greater than 1 was also analyzed among CC, CP, RC and RP during the three stages. For 8 DAT, a total of 27581 genes were commonly expressed among all four conditions (Figure 4.3D). The cortex showed lower number of uniquely expressed genes compared to the pith. While 190 and 163 unique genes were present in CC and RC respectively, 516 and 302 unique genes were present in CP and RP, respectively (Figure 4.3D).

#### Transcriptome profiles are distinctly different between the pith and cortex

Hierarchical cluster analysis was performed to explore the relationship of the transcriptome between the pith and cortex in response to fruit load reduction at 8, 19 and 47 DAT. The transcriptome profiles of the 12 conditions were classified into two major clusters: 8 DAT, and 19 and 47 DAT, suggesting that the overall transcriptome of early fruit development involving intensive cell production is distinctively different from that of the mid and late stages of fruit development. Relatively lesser differences in the transcriptome profiles were observed between the stages of transition and that of post-mitotic cell expansion (Figure 4.4A). Within each stage, the pith and cortex tissues were further divided into two subgroups, suggesting distinct different transcriptomic profiles between the two types of tissue during all stages, supporting the multi-dimensional scaling analysis. This is consistent with the large number of differentially expressed genes showing distinct stage or tissue specific expression patterns in many other fruits such as strawberry and tomato (Kang *et al.*, 2013; Pattison *et al.*, 2015). For both pith and cortex tissue

during each stage, the fruit load reduction group (RL) was clustered with the control group (CL), indicating little transcriptomic profile variation in response to fruit load reduction.

Differentially expressed gene (DEG) analysis was performed for four pairwise comparisons at each stage, including pith vs cortex in CL and RL group (CP vs CC; RP vs RC), and CL vs RL in cortex and pith (CC vs RC; CP vs RP). A large number of DEGs were identified between the pith and cortex in CL and RL fruit during all three stages. More genes were up regulated rather than down regulated in the pith than in the cortex at each stage. For RL fruit, 4635 (Down 1547, UP 3088), 3036 (Down 642, UP 2394), and 4623 (Down 1274, UP 3349) genes was differentially expressed in the pith than the cortex at 8, 19, and 47 DAT, respectively (Figure 4.4B). The CL fruit displayed similar number of DEGs between pith and cortex at three stages: 4229(Down 1464, UP 2828), 2298(Down 557, UP 1741), and 4428(Down 1406, UP 3022) DEGs between the pith and the cortex at 8, 19 and 47 DAT, respectively. Most of the DEGs in CL and RL group overlapped suggesting further that little change in the transcriptome occurred in response to fruit load reduction. For example, 2350 and 1127 DEGs were respectively up and down regulated in the pith than the cortex in both CL and RL at 8 DAT. The distinct transcriptomic profiles between the pith and cortex was consistent with the different growth patterns between the two types of tissues.

In contrast to the large number of DEGs between pith and cortex, very few genes were differentially expressed between CL and RL group, especially at 8 and 47 DAT. Additionally, more DEGs were down-regulated than up-regulated due to fruit load reduction. In the pith, there were 27 (Down14, UP 13), 710 (Down 582, UP128), and 3 (Down 2, UP1) DEGs between RL and CL at 8, 19, and 47 DAT, respectively (Figure 4.4C). This is consistent with the limited effect on cell number and area in the pith due to fruit load reduction. In the cortex, however, there were

relatively more DEGs between CL and RL groups: 56 (Down41, UP15), 911 (Down764, UP147), and 62 (Down24, UP38) at 8, 19 and 47 DAT, respectively. Together, these data suggest that fruit load reduction resulted to little variation of transcriptome profiles during early and mid fruit development, especially in the pith, this may explain why cell area and number in the pith was not significantly altered by fruit load reduction. It is worth noting that both cortex and pith displayed relatively more DEGs at 19 DAT during the period of transition from cell production to expansion. Previous studies indicated that cell area in the cortex was largely unaltered due to fruit load reduction in apple while the cell production responded quickly and increased since around 10 days after fruit load reduction treatment (Dash *et al.*, 2013). However, transcriptomic profile was not greatly affected in either tissue at 8 DAT in this study.

## Cluster analysis reveals specific genes expression characteristics for different stages

Due to the distinctive transcriptome profiles across three different stages, ANOVA-like test (QL F test) was performed in edgR to identify the DEGs across three different stages for all four conditions. It was revealed that a total of 5926 (CC), 1366 (CP), 8247 (TC), and 6678 (TP) genes were differentially expressed during the time course using an FDR value of 0.0001. All DEG genes for each condition were used for cluster analysis using Mfuzz in R. The cluster number for mfuzz was determined as 4 using factoextra package in R. The 1366 DEGs across three stages in CP were grouped into four clusters: cluster 1 (473 DEGs) and 4 (150 DEGs) representing genes that have the highest expression at 47 DAT, cluster 3 (173 DEGs) representing genes that have the lowest expression at 47 DAT, and cluster 2 (570 DEGs) representing genes with the highest expression at 8 DAT (Figure 4.5A). Cluster analysis for the other 3 conditions (CC, TC, TP) also revealed that clusters representing genes with the highest expression at 8 DAT is the most abundant one

among all clusters, indicating that large number of DEGs were actively expressed during early fruit development.

Gene ontology analysis was performed for all DEGs during three developmental stages in CP within cluster 2, which displayed highest expression levels at 8 DAT. Among the 570 DEGs, 346 genes were assigned to 28 different functional bins. The top bins with the highest number of DEGs were cell cycle (50 DEGs), chromatin organization (32 DEGs), RNA biosynthesis (31 DEGs), solute transport and enzyme classification (both 28 DEGs), and Protein modification and degradation (both 20 DEGs) (Figure 4.5B). Hierarchical clustering analysis was performed for the 50 DEGs within the BIN 'cell cycle' (Figure 4.5C). Unsurprisingly, all 50 DEGs were upregulated at 8 DAT in both cortex and pith. While some genes from this BIN displayed higher expression in the pith than the cortex at 8 DAT such as MD09G1068400 and MD17G1061300, there are multiple others displaying higher expression in the cortex than the pith, suggesting that regulation of cell production in different tissues may be mediated by specific cell cycle genes.

## Identification of WGCNA modules associated with RTGR

WGCNA was performed for all genes with an average TPM greater than 1, resulting in the identification of 19 modules. The module-trait relationship was investigated to explore modules associated with traits of interest including tissue area and RTGR. The module 'Ivory' containing 1258 genes was found to be negatively associated with the RTGR (r = -0.75, P = 0.0000001) (Figure 4.6), potentially including genes that have an important role in controlling processes associated with tissue growth. Genes within the module 'Ivory' were ordered based on their gene significance to explore potential candidate genes closely associated with RTGR. Among the top five genes, the 2<sup>nd</sup> most negatively correlated gene with RTGR was a transcription factor,

MD05G1281100 (Gene Significance = -0.8: *P*=0.000000005; Figure 4.7). MD05G1281100 is a gene belonging to the *TEOSINTE BRANCHED 1/CYCLOIDEA/PROLIFERATING CELL NUCLEAR ANTIGEN FACTOR1* (TCP) gene family with multiple roles in organ development in Arabidopsis (Nicolas and Cubas, 2016). In addition to the high negative gene significance value indicating the biological significance for relative tissue growth rate, MD05G1281100 also displayed a high module membership value (0.94) which suggests that it is an important hub gene associated with RTGR. The other top genes within the module 'Ivory' did not display a function related with cell production or expansion and thus were not investigated further in this study.

## A potential TCP dependent regulatory cascade may regulate cell production in apple

TCP family proteins are broadly grouped into two classes based on their TCP domains (Martin-Trillo and Cubas, 2010): class I and class II. While TCP members in class I promote plant growth and cell production, members in class II negatively regulate plant growth by repression of cell production (Kosugi and Ohashi, 1997; Li *et al.*, 2005; Ma *et al.*, 2014). The class II members are further sub-divided into two clades: CIN clade and the CYC/TB1 clade.

A regulatory cascade involving miRNAs, TCPs, KRPs and GROWTH REGULATING FACTORS (GRFs) has been proposed to control cell proliferation during leaf development in Arabidopsis (Rodriguez *et al.*, 2016). In this cascade, cell proliferation is positively regulated by GRFs and their interacting partners, GIFs (GRF INTERACTING FACTORs). *GRFs* are post-transcriptionally regulated by miR396. Many of the Arabidopsis *GRFs* contain a binding site for *miR396* (Debernardi *et al.*, 2012). The miR396 are in turn directly activated by TCPs (Schommer *et al.*, 2014). Further, multiple *TCPs* in Arabidopsis (*TCP2*, *TCP3*, *TCP4*, *TCP10* and *TCP24*) contain a binding site and are post transcriptionally targeted by *miR319* (Palatnik *et al.*, 2003). The

Arabidopsis TCP4 was also demonstrated to directly bind to the promoter region of *AtKRP1* and activate its expression. KRPs are key components of the cell cycle involved in negatively regulating the progression of the cell cycle by binding to the CDK/CYC complexes. Together, these processes allow for multiple routes for a TCP mediated negative regulation of cell production during organ development.

Considering that a potential *TCP* (*MdTCP2b*) was identified as a top gene negatively associated with negative regulation of tissue growth, it was hypothesized that a similar regulatory cascade involving *TCP*, *GRF/GIFs* and *KRPs* might be functional in negatively regulating cell production during apple fruit growth. To understand the role of such pathway in apple, the orthologous genes for the *TCP* gene family were identified using BLAST analysis from the apple genome. Overall, 28 *TCP* candidate genes were identified in the apple genome, and 17 among them belong to Class I and 11 to class II of the *TCP* family. MD05G1281100 and its closest paralog in apple MD10G1259500 are closest to Arabidopsis TCP2 that belongs to the CIN clade of TCP proteins which are involved in later organ development (Martin-Trillo and Cubas, 2010). Hence, MD10G1259500 and MD05G1281100 are designated as *MdTCP2a* and *MdTCP2b* respectively (Figure 4.8A).

Transcript abundance of genes involved in the regulatory cascade were quantified in the pith and cortex tissues during six developmental stages. These genes included *MdTCP2b*, *MdGRF7a* and *MdGIF3*, and *MdKRP4*. The transcript abundance of *MdTCP2b* decreased sharply in both cortex and pith during early fruit development (0-8DAT), a period associated with intensive cell production. *MdTCP2b* transcript abundance increased between 12 and 19 DAT, the period of transition from cell production to expansion and remained high during the rest of fruit development (Figure 4.8B). Noticeably, the expression level of *MdTCP2b* was consistently higher in the pith

than the cortex during most stages of fruit development. A similar expression pattern was observed for the negative regulator of cell production in apple, MdKRP4. The pith displayed up to two-fold higher transcript level compared to the cortex during most of fruit development. Relatively lower *MdKRP4* transcript abundance was noted during early fruit development. It slowly increased from 8 DAT in the cortex and 19 DAT in the pith, and remained at a relatively high level in both types of tissue during mid-late fruit development. The relatively low expression of MdTCP2b and MdKRP4 during early fruit development, especially in the cortex suggest that both genes might be repressed in the cortex during the intensive cell production phase, releasing the inhibition of cell production and contributing to greater growth in the cortex. The KRP4 promoter region contained a potential TCP binding site suggesting that its expression may be regulated by this transcription factor. In Arabidopsis, GRF and GIF form a transcription factor complex which is involved in leaf, stem and floral organ development. For both MdGRF7a and MdGIF3, highest transcript abundance was observed at 0 DAT. MdGRF7a transcript abundance decreased sharply between 0 and 8 DAT, especially in the cortex, and declined to almost zero by 26 DAT. Similarly, MdGIF3 transcript abundance declined continuously in both pith and cortex tissues to almost undetectable levels by 47 DAT. The transcript level of both *MdGRF7a* and particularly *MdGIF3* was generally higher in the cortex than in the pith during early fruit development. Together, the transcript abundance patterns of the genes were consistent with the previously proposed regulatory pathway associated with negative regulation of cell production.

The sequence of *MdTCP2b* was used as the target to search for potential miRNA binding sites using psRNATarget. The apple *miRNA319a* and *miRNA319b* were identified as having a potential targeting site within *MdTCP2b*. Similarly, the sequence of *MdGRF7a* was used to search for potential *miRNA* binding sites. Seven *miR396* were identified as potentially targeting

*MdGRF7a* at the same site (Figure 4.8C). These data are consistent with the involvement of the miR319 and miR396 in regulating *MdTCP2b* and *MdGRF7a*, respectively. Further, they support the potential role of a similar regulatory cascade in controlling cell production during apple fruit growth as proposed in Figure 4.8D (adapted from Schommer *et al.*, 2014).

## References

**Bain, J. M. and R. N. Robertson.** 1950. The physiology of growth in apple fruits I. Cell size, cell number, and fruit development. <u>Australian Journal of Biological Sciences</u> **4**: 75-91.

Carrari, F., C. Baxter, B. Usadel, E. Urbanczyk-Wochniak, M.-I. Zanor, A. Nunes-Nesi, V. Nikiforova, D. Centero, A. Ratzka and M. Pauly. 2006. Integrated analysis of metabolite and transcript levels reveals the metabolic shifts that underlie tomato fruit development and highlight regulatory aspects of metabolic network behavior. <u>Plant Physiology</u> **142**: 1380-1396.

**Dai, Z. W., C. Léon, R. Feil, J. E. Lunn, S. Delrot and E. Gomès.** 2013. Metabolic profiling reveals coordinated switches in primary carbohydrate metabolism in grape berry (Vitis vinifera L.), a non-climacteric fleshy fruit. Journal of Experimental Botany **64**: 1345-1355.

**Dash, M., L. K. Johnson and A. Malladi.** 2012. Severe shading reduces early fruit growth in apple by decreasing cell production and expansion. Journal of the American Society for <u>Horticultural Science</u> **137**: 275-282.

**Dash, M., L. K. Johnson and A. Malladi.** 2013. Reduction of fruit load affects early fruit growth in apple by enhancing carbohydrate availability, altering the expression of cell production-related genes, and increasing cell production. Journal of the American Society for Horticultural Science 138: 253-262.

**Dash, M. and A. Malladi.** 2012. The AINTEGUMENTA genes, MdANT1 and MdANT2, are associated with the regulation of cell production during fruit growth in apple (Malus× domestica Borkh.). <u>BMC plant biology</u> **12**: 98.

Gu, C., Y. H. Zhou, W. S. Shu, H. Y. Cheng, L. Wang, Y. P. Han, Y. Y. Zhang, M. L. Yu, D. Joldersma and S. L. Zhang. 2018. RNA-Seq analysis unveils gene regulation of fruit size cooperatively determined by velocity and duration of fruit swelling in peach. <u>Physiologia</u> <u>plantarum</u> 164: 320-336.

Hollender, C. A., C. Kang, O. Darwish, A. Geretz, B. F. Matthews, J. Slovin, N. Alkharouf and Z. Liu. 2014. Floral transcriptomes in woodland strawberry uncover developing receptacle and anther gene networks. <u>Plant physiology</u> **165**: 1062-1075.

Hu, P., G. Li, X. Zhao, F. Zhao, L. Li and H. Zhou. 2018. Transcriptome profiling by RNA-Seq reveals differentially expressed genes related to fruit development and ripening characteristics in strawberries (Fragaria× ananassa). <u>PeerJ</u> 6: e4976.

Kang, C., O. Darwish, A. Geretz, R. Shahan, N. Alkharouf and Z. Liu. 2013. Genome-scale transcriptomic insights into early-stage fruit development in woodland strawberry Fragaria vesca. <u>The Plant Cell</u> **25**: 1960-1978.

Kosugi, S. and Y. Ohashi. 1997. PCF1 and PCF2 specifically bind to cis elements in the rice proliferating cell nuclear antigen gene. <u>The Plant Cell</u> **9**: 1607-1619.

Li, C., T. Potuschak, A. Colón-Carmona, R. A. Gutiérrez and P. Doerner. 2005. Arabidopsis TCP20 links regulation of growth and cell division control pathways. <u>Proceedings of the National Academy of Sciences</u> **102**: 12978-12983.

Li, M., F. Feng and L. Cheng. 2012. Expression patterns of genes involved in sugar metabolism and accumulation during apple fruit development. <u>PloS one</u> 7: e33055.

Lombardo, V. A., S. Osorio, J. Borsani, M. A. Lauxmann, C. A. Bustamante, C. O. Budde, C. S. Andreo, M. V. Lara, A. R. Fernie and M. F. Drincovich. 2011. Metabolic profiling during peach fruit development and ripening reveals the metabolic networks which underpin each developmental stage. <u>Plant Physiology</u>: pp. 111.186064.

Ma, J., Q. Wang, R. Sun, F. Xie, D. C. Jones and B. Zhang. 2014. Genome-wide identification and expression analysis of TCP transcription factors in Gossypium raimondii. <u>Scientific reports</u> **4**: 6645.

**MacDaniels, L. H.** 1940. <u>Morphology Of The Apple And Others Pome Fruits</u>, Cornell University, Agricultural Experiment Station; Ithaca, New York.

Malladi, A. and P. M. Hirst. 2010. Increase in fruit size of a spontaneous mutant of 'Gala'apple (Malus× domestica Borkh.) is facilitated by altered cell production and enhanced cell size. Journal of experimental botany 61: 3003-3013.

Pattison, R. J., F. Csukasi, Y. Zheng, Z. Fei, E. van der Knaap and C. Catalá. 2015. Comprehensive tissue-specific transcriptome analysis reveals distinct regulatory programs during early tomato fruit development. <u>Plant physiology</u> **168**: 1684-1701.

Rodriguez, R. E., C. Schommer and J. F. Palatnik. 2016. Control of cell proliferation by microRNAs in plants. <u>Current opinion in plant biology</u> **34**: 68-76.

Sánchez-Sevilla, J. F., J. G. Vallarino, S. Osorio, A. Bombarely, D. Posé, C. Merchante, M. A. Botella, I. Amaya and V. Valpuesta. 2017. Gene expression atlas of fruit ripening and transcriptome assembly from RNA-seq data in octoploid strawberry (Fragaria× ananassa). Scientific reports 7: 13737.

Vashisth, T., L. K. Johnson and A. Malladi. 2011. An efficient RNA isolation procedure and identification of reference genes for normalization of gene expression in blueberry. <u>Plant cell</u> reports **30**: 2167-2176.

Xie, M., Y. Huang, Y. Zhang, X. Wang, H. Yang, O. Yu, W. Dai and C. Fang. 2013. Transcriptome profiling of fruit development and maturation in Chinese white pear (Pyrus bretschneideri Rehd). <u>BMC genomics</u> 14: 823.

Zhang, J., X. Wang, O. Yu, J. Tang, X. Gu, X. Wan and C. Fang. 2010. Metabolic profiling of strawberry (Fragaria× ananassa Duch.) during fruit development and maturation. Journal of experimental botany 62: 1103-1118.



**Figure 4.1.** Effects of reduction in fruit load on relative tissue growth rate (RTGR) in the pith and cortex of apple fruit.

Mean  $\pm$  S.E (n = 4) are displayed here. Asterisk indicates significant difference between the control (CL) and reduced fruit load (RL) treatments at  $\alpha = 0.05$ .



Figure 4.2. Effects of reduction in fruit load on cell growth.

Cell number, cell area, relative cell production rate (RCPR), and relative cell expansion rate (RCER) of the pith and cortex apple fruit are presented. The mean and standard error of the mean (n = 4) are displayed. CC: Control fruit load cortex; CP: Control fruit load pith; RC: Reduced fruit load cortex; RP: Reduced fruit load pith. Asterisks and † indicates the significant difference between control and reduced fruit load treatments in the cortex and pith, respectively ( $\alpha = 0.05$ ).



**Figure 4.3.** *Global aspects of the transcriptome in the cortex and pith during three stages of fruit development.* 

CC: Control fruit load cortex; CP: Control fruit load pith; RC: Reduced fruit load cortex; RP: Reduced fruit load pith.

(A) Number of counts from sequencing in each sample.

(B) A Venn diagram showing the number of commonly and uniquely expressed genes among three stages in CC.

(C) Multiple dimensional scaling(MDS) plot showing the relationship among the transcriptomes in all samples.

(D) A Venn diagram showing the number of commonly and uniquely expressed genes among CC,CP, RC and RP at 8 DAT.



**Figure 4.4.** *Hierarchical clustering and differentially expressed gene analysis of the transcriptomes in the cortex and pith during three stages of fruit development.* 

(A) Heat map of hierarchical clustering indicate different transcriptome profiles among three developmental stages, as well as between cortex and pith. CC: Control fruit load cortex; CP: Control fruit load pith; RC: Reduced fruit load cortex; RP: Reduced fruit load pith.

(B) Differential expressed genes between the cortex and pith during three developmental stages. The blue color indicate DEGs downregulated in the pith, and the pink indicate DEGs upregulated in the pith compared to the cortex. CL: control, RL: reduced fruit load.

(C) Differential expressed genes between the control (CL) and reduced fruit load (RL) treatments during three developmental stages. The blue color indicate DEGs downregulated in the RL, and the pink indicate DEGs upregulated in the RL compared to the CL.


**Figure 4.5.** *Clustering and gene ontology analysis of the differentially expressed genes among three developmental stages in CP (control fruit load pith).* 

(A) Clustering of 1366 DEGs in CP using Mfuzz.

(B) Gene ontology analysis for the 570 DEGs (cluster 2) displaying the highest expression at 8

DAT in control fruit load pith.

(C) Heat map showing the expression of 50 DEGs from the BIN "cell cycle" in the cortex and pith

of thinned and control fruit during three developmental stages. CC: Control fruit load cortex; CP:

Control fruit load pith; RC: Reduced fruit load cortex; RP: Reduced fruit load pith.



**Figure 4.6.** Weighted gene co-expression network analysis (WGCNA) of all genes with TPM greater than 1 in all sequenced samples.

Module-trait correlations and corresponding P values were presented. The color scale shows module trait correlation from -1(blue) to 1 (red).



Module Membership of genes in ivory module

**Figure 4.7.** *Module membership (MM) of all genes in module "Ivory", and their gene significance (GS) with the relative tissue growth rate (RTGR).* 

The 2<sup>nd</sup> most important hub gene based on the GS, MD05G1281100 was labeled with star in red.



**Figure 4.8.** *A potential regulatory cascade involved in regulation of cell production during apple fruit growth.* 

(A) Phylogenetic tree displaying the relationships among TCP gene family in Arabidopsis (At), rice (Os) and apple. TCP genes were clustered into three classes. Class I: blue colored branch, Class II green colored branch and CYC/TB1: red colored branch.

(B) Expression level of *MdTCP2b*, *MdKRP4*, *MdGRF7a* and *MdGIF3* in the cortex and pith during fruit development. CC: Control fruit load cortex; CP: Control fruit load pith; RC: Reduced fruit load cortex; RP: Reduced fruit load pith.

(C) Targeting site between *MdTCP2b* and *miR319a*, and *MdGRF7a* and *miR396a*.

(D) A potential regulatory cascade regulating cell production during apple fruit development.(Adapted from Schommer et al., 2014)

# Supplement Data

Samples	NO. of raw read	Adapter trimming		Data mapping	
		Remained	%	Aligned pairs	%
8CC1	11,687,360	11,394,679	97.50%	10,470,411	91.90%
8CC2	12,700,721	12,386,489	97.50%	11,296,442	91.20%
8CC3	13,729,786	13,383,447	97.50%	12,299,652	91.90%
8CP1	11,523,096	11,274,317	97.80%	10,320,718	91.50%
8CP2	15,085,526	14,741,277	97.70%	13,399,888	90.90%
8CP3	15,600,802	15,265,558	97.90%	13,947,468	91.40%
8RC1	14,010,049	13,667,115	97.60%	12,577,633	92.00%
8RC2	13,179,569	12,856,108	97.50%	11,831,137	92.00%
8RC3	11,795,408	11,491,497	97.40%	10,520,521	91.60%
8RP1	12,596,828	12,330,057	97.90%	11,354,416	92.10%
8RP2	11,963,151	11,671,777	97.60%	10,690,681	91.60%
8RP3	18,245,122	17,809,991	97.60%	16,370,215	91.90%
19CC1	32,087,827	30,894,033	96.30%	28,235,458	91.40%
19CC2	21,753,014	20,865,340	95.90%	19,026,835	91.20%
19CC3	18,603,496	17,820,439	95.80%	16,175,257	90.80%
19CP1	35,097,932	33,239,351	94.70%	29,977,260	90.20%
19CP2	20,762,077	19,912,168	95.90%	18,008,660	90.40%
19CP3	20,134,457	19,262,309	95.70%	17,428,658	90.50%
19RC1	32,646,917	31,348,086	96.00%	28,636,484	91.40%
19RC2	28,071,480	26,964,807	96.10%	24,594,038	91.20%
19RC3	19,272,074	18,503,206	96.00%	16,901,836	91.30%
19RP1	27,954,740	26,766,496	95.70%	24,206,448	90.40%
19RP2	42,065,586	40,252,877	95.70%	36,469,882	90.60%
19RP3	24,102,805	23,025,674	95.50%	20,817,284	90.40%
47CC1	20,903,667	20,231,806	96.80%	18,551,067	91.70%
47CC2	28,263,226	27,313,183	96.60%	24,890,128	91.10%
47CC3	31,751,980	30,703,945	96.70%	28,212,353	91.90%
47CP1	22,454,560	21,904,345	97.50%	20,157,070	92.00%
47CP2	25,546,594	24,900,729	97.50%	22,836,992	91.70%
47CP3	27,516,973	26,852,989	97.60%	24,661,297	91.80%
47RC1	20,231,006	19,557,629	96.70%	17,957,660	91.80%
47RC2	20,476,428	19,746,284	96.40%	18,027,190	91.30%
47RC3	23,230,300	22,414,952	96.50%	20,515,450	91.50%
47RP1	41,475,910	40,476,659	97.60%	37,154,101	91.80%
47RP2	36,466,690	35,523,320	97.40%	32,600,504	91.80%
47RP3	25,562,737	24,937,091	97.60%	22,898,427	91.80%

**Table S4.1.** Raw read and read mapping statistics for all samples.

#### CHAPTER 5

#### CONCLUSIONS

Since apple fruit is a classic accessory fruit, the unique fruit structure presents apple as a potential model species to study the mechanisms regulating fruit development. In this study, mechanisms regulating apple fruit growth in a spatial and temporal manner were investigated. We applied different methods to understand the metabolic and genetic regulation of fruit growth. The cortex and pith tissue, derived from different origins displayed dramatic differences in tissue area due to enhanced cell production and expansion in the cortex. Fruit load reduction improved fruit size by improving cell production during multiple stages, and hence the growth in the cortex specifically. Targeted metabolic profiling revealed significant difference between the cortex and pith, potentially associated with the dramatic different growth patterns between the two types of tissue. Apple fruit growth is dependent on the translocated C, N and other reserves, especially during early fruit development. Although little research has been performed for the N metabolism regulating apple fruit development, a high association between the major C and N resources, sorbitol and asparagine was revealed in this study. Expression profiling of C and N metabolisms related genes during fruit growth indicated spatial characteristics of C and N metabolisms. In addition, distinct temporal metabolites accumulation patterns were observed during different developmental stages. Although early fruit development lasts for a much shorter period (0-26 DAT) compared to mid (33-47 DAT) and late fruit development (77-118 DAT), significant variation of

metabolic profile actually occurred during early fruit development, when fruit growth was mainly mediated by cell production.

The homologous genes in the families including in the FW2.2/CNR, GRF/GIF, ARGOS/ARGOS-Like and KLUH gene family were identified in the apple genome. Expression profiling of the putative gene organ size regulatory genes was performed during pre-bloom, fruit set and fruit development in apple, leading to the identification of potential positive regulators of cell production: *MdGRF1*, *MdGRF7a*, *MdGRF8*, *MdGIF3*, negative regulators of cell production: MdCNR5a, MdCNR6, MdCNR8, MdARG1 and MdARG2. RNA-seq analysis was performed to explore genes and gene networks regulating fruit development. A significant number of DEGs were revealed between the pith and cortex during cell production, transition from cell production, and cell expansion phase. However, the transcriptome profile was not affected greatly due to fruit load reduction. Cluster analysis and the gene ontology analysis yielded significant information relating genes that were expressed during a specific stage, and proven to be effective to identify potential groups of cell production or expansion related genes in the future. Based on the WGCNA and expression profiling analysis, a potential regulatory network involving miR319-MdTCP*miR396-MdGRF/MdGIF* was proposed that may play a similar role regulating cell production as in Arabidopsis.

### **Future Research**

Further research needs to be performed to understand the functions of genes and gene network identified in this study. The apple genome is duplicated but the expression profiling was not performed for all duplicated genes in this study(Velasco *et al.*, 2010). The duplicated genes, *MdCNR5a* and *MdCNR5b*, and *MdGRF7a* and *MdGRF7b*, displayed very similar expression

pattern during fruit development. However, this was not confirmed for all duplicated genes. Another limitation of this study was that the focus was on the individual gene that may play an important role regulating fruit development, however, the regulation of such complicated process would involve groups of genes coordinating with each other. The miR319-AtTCP-miR396-AtGRF/GIF regulatory cascade has been proposed to play crucial roles regulating organ growth in Arabidopsis(Schommer et al., 2014). This study presented evidence of a similar cascade regulating apple fruit development, but more research needs to be performed to investigate the coordination of members within the cascade. For example, the quantification of miR319 and miR396 during different developmental stages, or the evaluation of interaction between miRNA and their target, and the identification of downstream genes regulated by GRF/GIF transcription factor complex using Chip-Seq analysis(Lu et al., 2011). The separation of pith and cortex presented initial evidence for a spatial feature of C and N metabolism during fruit development, however enzymatic activities of crucial proteins related with C and N metabolism were not investigated in this study(Baud and Graham, 2006). In addition, a finer separation of tissue using laser micro dissection would provide more information relating metabolic and molecular regulation of fruit development(Matas et al., 2010).

## References

**Baud, S. and I. A. Graham.** 2006. A spatiotemporal analysis of enzymatic activities associated with carbon metabolism in wild-type and mutant embryos of Arabidopsis using in situ histochemistry. The Plant Journal **46**: 155-169.

Lu, Y., Y. Zhou, W. Qu, M. Deng and C. Zhang. 2011. A Lasso regression model for the construction of microRNA-target regulatory networks. <u>Bioinformatics</u> 27(: 2406-2413.

Matas, A. J., J. Agustí, F. R. Tadeo, M. Talón and J. K. Rose. 2010. Tissue-specific transcriptome profiling of the citrus fruit epidermis and subepidermis using laser capture microdissection. Journal of experimental botany 61: 3321-3330.

Schommer, C., J. M. Debernardi, E. G. Bresso, R. E. Rodriguez and J. F. Palatnik. 2014. Repression of cell proliferation by miR319-regulated TCP4. <u>Molecular plant</u> 7: 1533-1544.

Velasco, R., A. Zharkikh, J. Affourtit, A. Dhingra, A. Cestaro, A. Kalyanaraman, P. Fontana, S. K. Bhatnagar, M. Troggio and D. Pruss. 2010. The genome of the domesticated apple (Malus× domestica Borkh.). <u>Nature genetics</u> **42**: 833.