GENOMIC AND PHENOTYPIC ANALYSES OF VEGETABLE DISEASE RESISTANCE AND FRUIT QUALITY

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Lindsay Elizabeth Wyatt

May 2015

GENOMIC AND PHENOTYPIC ANALYSES OF VEGETABLE DISEASE RESISTANCE AND FRUIT QUALITY

Lindsay Elizabeth Wyatt, Ph. D.

Cornell University 2015

This work describes genomics research and applied breeding in pepper and squash towards improvement in fruit quality and disease resistance. A major factor in pepper fruit quality is the presence or absence of pungency, caused by the presence of capsaicinoid molecules and controlled by the major gene *Pun1*. Three mutant alleles of *Pun1* have been identified that cause the production of little to no capsaicinoids. A suite of PCR-based molecular markers are presented that differentiate the four *Pun1* alleles, which will be useful for pepper breeding, seed purity testing, and characterization of new germplasm. Disease resistance is also important for pepper production, and *Phytophthora capsici* is an oomycete pathogen that can cause total yield loss in pepper. Nine bell pepper inbred lines were developed that are resistant to *P. capsici*. They have yields comparable to commercial cultivars, but have smaller fruit. They provide a high level of resistance to F_1 progeny and are therefore suitable for use as both inbred lines and hybrid parents. To study fruit quality in winter squash, an acorn squash fruit and seed transcriptome was sequenced from the cultivar Sweet REBA at five developmental stages. A total of 55,949 unigenes were assembled, and 85% of the unigenes with predicted open reading frames had homology with previously identified genes. Building on this transcriptome,

comparative fruit and seed transcriptome sequencing was conducted to compare ‘Sweet REBA’ with the oilseed pumpkin ‘Lady Godiva’, developed for seed consumption. Pathways for carotenoid, starch, and sucrose biosynthesis in winter squash fruit were assembled from the literature and homologs were identified for all of the structural genes involved. Differential gene expression between the two cultivars was consistent with their observed phenotypes for these three metabolites of interest, indicating that transcriptional regulation of metabolic genes is an important determinant of winter squash fruit quality. Overall, this research supports vegetable breeding through molecular marker development, the evaluation of improved disease-resistant lines, and the investigation of the genetic basis of traits of interest.

BIOGRAPHICAL SKETCH

Lindsay Wyatt was born July 15, 1986 to parents Bill and Betsy Wyatt. She was raised with her siblings Kelsey, Jessi, and Michael in Mentor, Ohio on the shores of Lake Erie. Her childhood spent gardening with her parents and living next door to her family's third-generation retail greenhouse and garden center instilled in her an early love of plants. After graduating from Mentor High School, she pursued a Bachelor of Science degree in Biology at Duke University. At Duke, she conducted research on the biological rhythm of the marine amphipod *Talorchestia longicornis* with Dr. Richard Forward while spending a semester at the Duke University Marine Laboratory. She spent a summer interning at the Cleveland Museum of Natural History studying paleobotany with Dr. Shya Chitaley, identifying 360 million year old spores present in Devonian-era shale. The next year, she conducted her senior thesis research with Dr. Philip Benfey on the identification of upstream regulators of the gene SHORT-ROOT in *Arabidopsis*. After graduating *summa cum laude* from Duke, Lindsay spent a year at the National Institutes of Health as a post-baccalaureate fellow in Dr. Alan Michelson's lab studying *Drosophila* heart cell migration.

Lindsay began her Ph.D. studies in the Department of Plant Breeding and Genetics at Cornell University in 2009, and after rotating with Neil Mattson and Philip Griffiths, joined the lab of Michael Mazourek. While at Cornell, she has been active in the Plant Breeding and Genetics graduate student group, Synapsis, serving as both president and events coordinator. She was a teaching assistant for an undergraduate Plant Genetics course for two semesters, developing new laboratory exercises for the course. She has also helped to organize three symposia, including the 2014 Student

Organic Seed Symposium, hosted at Cornell. In addition to plant breeding, Lindsay's interests include reading, cooking, canning and preserving, and playing the viola.

For my grandparents, Arnold and Carol Blair

ACKNOWLEDGMENTS

I am grateful for the funding that made my time at Cornell possible. Generous support from the Cornell University Presidential Life Sciences Fellowship, USDA AFRI NIFA National Needs Fellowship #2008-38420-04755, USDA AFRI NIFA Plant Breeding and Education Grant #2010-85117-20551, and USDA AFRI NIFA Graduate Fellowship #2013-67011-21122 funded my graduate studies. A grant from the Toward Sustainability Foundation and USDA AFRI NIFA Plant Breeding Competitive Grant #2013-67013-21232 supported much of my research activities. I am additionally grateful for specific project support provided by a variety of sources, detailed in the subsequent chapters.

I would like to thank my advisor and mentor, Michael Mazourek, for his continual support and guidance throughout my Ph.D. career. I truly appreciate his time spent teaching me both the artistic and the practical aspects of plant breeding, the professional development opportunities he has provided for me, and for allowing me to assume a leadership role in our breeding program. My graduate committee members Christine Smart and Li Li have provided me with encouragement and advice and have been valuable collaborators in our research activities. Maryann Fink and Michael Glos have been instrumental to my success as a graduate student, providing great assistance in field and greenhouse work and also as friends and mentors. Our fantastic summer workers and the Cornell farm crew have also been a huge help in my field research. In the lab, Nancy Eannetta taught me many molecular biology skills and was a great source of support as I began graduate school. Suzy Strickler has been a great collaborator for my transcriptome sequencing work and is a pleasure to work

with. My fellow graduate students have been fantastic colleagues and have provided both emotional and scientific support. I especially appreciate the friendship of Bill Holdsworth in our five years together in the Mazourek lab. We have had many wonderful undergraduate lab members who have participated in our lab activities throughout my time at Cornell. Samantha Klasfeld and Nicholas Biebel in particular provided valuable help with my research. I am also grateful for the support of the Plant Breeding & Genetics faculty, staff, and students, who made it a pleasure to come to Cornell every day. Finally, I would like to thank my family and friends for their love and encouragement.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Vegetables are an important part of a balanced diet, providing a low-calorie source of nutrients and dietary fiber. To support vegetable production and consumption in the United States through plant breeding and genetics, several areas of focus can be considered. First, advancing genetic and genomic knowledge about vegetable crops supports breeding activities through the identification of genes of interest, the development of molecular markers, and a deeper understanding of the genetic architecture of traits targeted for improvement. Many improvements can also be made through applied breeding activities; one important target is the development of disease-resistant cultivars that can benefit growers through increased yield and reduced pesticide use. Another area of focus for increasing vegetable consumption is the development of high-quality, unique, nutritious vegetable varieties that appeal to consumers through their taste and appearance. This dissertation describes genomics research and applied breeding targeting fruit quality and disease resistance in pepper and squash to support both vegetable production and consumption.

Pepper

Peppers (*Capsicum spp.*) are cultivated as a vegetable crop, as a spice, and for the industrial production of capsaicin (Crosby, 2008). They are distinctive because of their production of capsaicin, a molecule that causes the burning sensation of hot peppers. In addition to capsaicin, peppers also contain carotenoids, flavonoids, and ascorbic acid, making them a highly nutritious food (Crosby, 2008). They are native

to the New World tropics, with centers of diversity of the cultivated species ranging from Mexico to Peru, and were domesticated over many years in the Pre-Columbian era (Crosby, 2008). There are five major cultivated pepper species: *C. annuum, C. chinense, C. frutescens, C. baccatum,* and *C. pubescens*.

Pepper Genomic Resources

Capsicum species have a large genome with a continually growing amount of genetic resources. The pepper genome is diploid and has 12 chromosomes, with genome size estimates ranging from 2700-3480 Mbp per haploid genome (Arumuganathan and Earle, 1991; Kim et al., 2014). Many molecular marker sets and linkage maps have been created for pepper (Lefebvre et al., 2002; Prince et al., 1993). One notable map, the COSII map (Wu et al., 2009), used markers orthologous between the pepper and tomato genomes, allowing the exploration of the degree of synteny between pepper and tomato. More recent maps have incorporated simple sequence repeat (SSR) and single nucleotide polymorphism (SNP) markers generated from high-throughput sequencing data (Lu et al., 2012; Yarnes et al., 2013). In addition, a Pepper Affymetrix GeneChip microarray was created that can detect 33,401 SNPs across species and 6,426 SNPs within *C. annuum* (Hill et al., 2013).

Transcriptome sequencing in *C. annuum* has recently added a great deal of genomic knowledge. It has been used to identify approximately 25,000 unigenes expressed in pepper (Lu et al., 2011) and to design molecular markers that differentiate several popular commercial peppers from 'Criollo de Morelos 334', a landrace that is the source of several disease resistances (Ashrafi et al., 2012). The *C. frutescens* transcriptome was also sequenced in order to study capsaicinoid

biosynthesis (Liu et al., 2013b). Other pepper transcriptome studies have investigated stress response signaling (Lee and Choi, 2013), cytoplasmic male sterility (Liu et al., 2013a), and fruit development (Martinez-Lopez et al., 2014).

Recently, the pepper genome was sequenced by Kim et al. (Kim et al., 2014). They sequenced the 'Criollo de Morelos 334' genome at 186x coverage and assembled 3.06 Gb of the estimated 3.48 Gb genome. They then mapped the 37,989 scaffolds onto a genetic map and resequenced two additional *C. annuum* varieties and a *C. chinense* accession, PI159236. Approximately 11 million SNPs were identified between the *C. annuum* lines and 56 million SNPs between the two species. Over 75% of the pepper genome was found to consist of transposable elements and 34,903 protein-coding genes were predicted.

Pepper is a member of the Solanaceae family, which also includes tomato (*Solanum lycopersicum*), potato (*Solanum tuberosum*), eggplant (*Solanum melongena*), and tobacco (*Nicotiana benthamiana*). Of these species, tomato has the most genetic resources, including a sequenced genome (Sato et al., 2012) and sequenced transcriptomes (Gupta et al., 2013; Hamilton et al., 2012; Lee et al., 2012). Pepper and tomato have a high degree of synteny (Livingstone et al., 1999; Wu et al., 2009), which has made tomato a valuable resource for identifying pepper candidate genes.

Pungency

Pungency, caused by the biosynthesis of capsaicinoid compounds, is unique to the genus *Capsicum*. Peppers are valued as a spice due to their capsacinoid production and are also grown for the industrial production of capsaicinoids. The primary

capsaicinoids are capsaicin, dihydrocapsaicin, and nordihydrocapsaicin (Bennett and Kirby, 1968). Capsaicinoids cause their burning sensation by activating the TRPV1 (VR1) receptor (Caterina et al., 1997), which also responds to high heat, and therefore activating sensory neurons that send thermal pain signals to the brain. They are produced in blister-like structures in the placenta tissue of the fruit (Zamski et al., 1987) through the condensation of vanillylamine, derived from the phenylpropanoid pathway, and a branched chain fatty acid (Bennett and Kirby, 1968; Leete and Louden, 1968).

Several major genes that affect capsaicinoid production cause a wide variation in the capsaicinoid content found in pepper. The gene *Pun1* encodes a putative acyltransferase and is required for production of capsaicinoids (Stewart et al., 2005). Recessive mutant alleles of *Pun1* have been identified in *C. annuum, C. chinense*, and *C. frutescens* (Stellari et al., 2010; Stewart et al., 2007; Stewart et al., 2005). Plants possessing these mutant alleles do not produce capsaicinoids and consequently have non-pungent fruit. The *Pun1* allele in *C. annuum*, $pun1^{1}$, is the source of the non-pungent characteristic of most cultivated sweet peppers (Stewart et al., 2005). Another gene, *Pun2*, was identified in *C. chacoense*. When *Pun2* is mutant, there is a large reduction in capsaicinoid production (Stellari et al., 2010). Finally, mutations in the putative aminotransferase gene (*pAMT*) cause the production of non-pungent capsinoids instead of pungent capsaicinoids (Kobata et al., 1998; Kobata et al., 1999; Lang et al., 2009; Tanaka et al., 2010).

In addition, multiple loci have been identified that affect the pungency of peppers by quantitatively influencing the amount of capsaicinoids produced. The low-

pungency gene *cf* was identified in *C. frutescens* and reduces the level of capsaicinoids produced (Saritnum et al., 2008). Several studies have also mapped quantitative trait loci (QTLs) that affect the amount of capsaicinoid production. The first study identified a major QTL, *cap*, which explained 34-38% of the variation in capsaicinoid content in a cross between *C. annuum* and *C. frutescens* (Blum et al., 2003). A second study identified 6 QTLs that influence capsaicinoid content (Ben-Chaim et al., 2006). More recently, a study utilizing the Pepper GeneChip identified 12 QTL for capsaicinoid traits in a *C. annuum* x *C. frutescens* population (Yarnes et al., 2013). This quantitative variation in pungency is further influenced by environmental factors (Zewdie and Bosland, 2000). For example, environmental stresses to pepper plants, such as water stress, cause an increase in the pungency of their fruits (Estrada et al., 1999; Harvell and Bosland, 1997; Ruiz-Lau et al., 2011).

Phytophthora capsici

Phytophthora capsici Leonian is an important disease of peppers that was first discovered on chilies in New Mexico in 1918 (Leonian, 1922). *P. capsici* is an oomycete that can infect a wide range of vegetable crops, including squash (Crossan et al., 1953), tomato, eggplant, watermelon, and peppers (Polach and Webster, 1972). On peppers, *P. capsici* causes several distinct disease syndromes: foliar blight, root rot, stem blight, and fruit rot (Sy et al., 2005; Walker and Bosland, 1999). *P. capsici* can cause severe losses in many pepper-producing areas around the United States and the world (Bosland and Lindsey, 1991; Garcia-Rodriguez et al., 2010; Hwang and Kim, 1995; Oelke et al., 2003).

P. capsici is a heterothallic oomycete with A1 and A2 mating types (Smith et al., 1967). *P. capsici* reproduces sexually by forming oospores, which are capable of overwintering, and is a polycyclic pathogen that can complete multiple infection cycles in one growing season, reproducing asexually through the formation of sporangia (Ristaino, 1991). *P. capsici* sporangia, unlike those of some other *Phytophthora* species, are dispersed by water (Granke et al., 2009; Hausbeck and Lamour, 2004; Ristaino and Johnston, 1999; Schlub, 1983). Because zoospores are released with and move in water, soil moisture, rainfall, and irrigation have a significant impact on disease development (Bowers and Mitchell, 1990; Cafe and Duniway, 1995; Ristaino, 1991).

P. capsici is a diverse pathogen with multiple races of differential infectivity. Several studies have attempted to divide isolates into different races based on pathogenicity on a variety of different peppers (Glosier et al., 2008; Oelke et al., 2003; Sy et al., 2008). It is apparent from these studies that there is not yet one accepted designation of races for *P. capsici*. Additionally, it is likely that more races will be detected with further study, as isolates from only a fraction of *P. capsici*'s range have been categorized thus far.

Control strategies for *P. capsici* include modifying cultural practices, fungicide use, and the utilization of resistant pepper varieties. Cultural practices that can reduce disease consist of sanitation procedures to keep *P. capsici* from initially infesting a field (Hausbeck and Lamour, 2004), crop rotation (Ristaino and Johnston, 1999), and water management (Cafe and Duniway, 1995; Ristaino, 1991; Ristaino and Johnston, 1999). Mefenoxam is a widely used fungicide for *P. capsici* control (Ristaino and

Johnston, 1999), but due to the increasing use of a limited range of fungicides, mefenoxam-insensitive *P. capsici* isolates have been widely reported (Lamour and Hausbeck, 2001). Perhaps the most grower- and environmentally-friendly control strategy is the use of pepper varieties that are resistant to *P. capsici*, combined with cultural control. Unfortunately, there has been significant difficulty with breeding peppers resistant to *P. capsici* and there are currently no peppers available with the full complement of resistance to the four disease syndromes as well as the desired horticultural type (Oelke et al., 2003; Ristaino and Johnston, 1999; Thabuis et al., 2004b).

A variety of resistance sources for *P. capsici* have been identified in pepper. Two sources for root rot resistance, PI 201234 and PI 201232, were identified in 1960 (Kimble and Grogan, 1960) and still provide good resistance and continue to be studied (Ogundiwin et al., 2005; Ortega et al., 1992; Palloix et al., 1990). Recently, a large-scale screen of the USDA *Capsicum annuum* germplasm collection was conducted, measuring root rot resistance (Candole et al., 2010). Fourteen accessions were identified that displayed resistance in the initial greenhouse screen as well as in replicated testing and of these, two accessions, PI 201237 and PI 640532, had the best resistance and are potential resistance sources for future breeding (Candole et al., 2010). A Korean screen identified four additional accessions that can be used in breeding for resistance (Mo et al., 2014). The best resistance source thus far is a land race from Mexico, Criollo de Morelos 334 (CdM334) (Guerrero-Moreno and Laborde, 1980), which has consistently displayed resistance to root rot (Bosland and Lindsey, 1991), foliar blight (Alcantara and Bosland, 1994), stem blight (Sy et al., 2005), and

fruit rot (Egea-Gilabert et al., 2008) and is resistant to a wide range of *P. capsici* isolates (Oelke et al., 2003). CdM334 therefore has the most potential for use in breeding as a *P. capsici* resistance source.

There has been extensive research on the genetics of resistance to *P. capsici* in pepper. Early studies had varying results, likely confounded by the use of different *P. capsici* isolates, different resistance sources, and different screening methods. Resistance to root rot has been suggested to be controlled by one or two dominant genes (Smith et al., 1967), three dominant genes (Ortega et al., 1991), or two genes with epistasis (Reifschneider et al., 1992). Foliar blight resistance has been attributed to a single dominant or incompletely dominant gene, possibly with modifiers (Barksdale et al., 1984). Fruit rot resistance was predicted to correspond to a single dominant gene (Saini and Sharma, 1978). Epistasis was also predicted to be a contributing genetic factor for resistance (Bartual et al., 1991).

Recently, work has been done to differentiate the genetics of resistance to the different disease syndromes caused by *P. capsici*. Resistances to root rot and foliar blight were controlled by independent dominant genes in a cross between 'Early Jalapeño' and CdM334 (Walker and Bosland, 1999). Another independent dominant gene controlled stem blight resistance in the same cross (Sy et al., 2005). However, only one *P. capsici* isolate was used in these experiments. Thus, the results are specific to this isolate and do not identify all potential resistance genes. From this data, it is clear that each disease syndrome must be considered when breeding for resistance, as genes that provide resistance to one syndrome may not also provide resistance to the others. It also suggests that there are gene-for-gene interactions

providing resistance to *P. capsici* (Sy et al., 2005). Another study identified additional gene-for-gene resistance interactions; a set of RILs from a cross between 'Early Jalapeño' and CdM334 enabled the detection of five dominant, non-allelic R genes that provide resistance to one *P. capsici* race per R gene (Monroy-Barbosa and Bosland, 2008).

Although there are these major resistance genes that have specific interactions with different races of *P. capsici*, there are also quantitative resistance genes. Several studies have mapped QTL for *P. capsici* resistance. An early QTL study, using 'Perennial' as the resistance source, identified two major and 11 minor QTL providing root rot and stem blight resistance (Lefebvre and Palloix, 1996). Another study utilizing three resistance sources ('Vania', derived from PI 201234, 'Perennial', and CdM334) identified between five and nine QTL per population, most of which were population-specific (Thabuis et al., 2003). One QTL was common to 'Vania' and 'Perennial. Another QTL, *Phyto.5.2*, which provided the strongest effect, was common to all three resistance sources (Thabuis et al., 2003). Sixteen QTL were identified for root rot and foliar blight resistance in a population using PI 201234 as the resistance source (Ogundiwin et al., 2005). In the same study, five QTL were identified from CdM334 that provide root rot resistance. Most of the identified QTL only provided resistance to some of the *P. capsici* isolates used for resistance testing. One major QTL on chromosome five was present in both populations and corresponded to *Phyto.5.2* (Ogundiwin et al., 2005). A bulk segregant analysis comparing resistant and susceptible pepper lines identified one RAPD marker linked to resistance. This marker was also found to correspond to *Phyto.5.2* (Quirin et al.,

2005). Another study using PI 201234 as the resistance source found one major and two minor QTL (Sugita et al., 2006). The major QTL found in the study corresponded to *Phyto.5.2* (Sugita et al., 2006). Further confirmation for the importance of *Phyto.5.2* came from two additional studies, both using CdM334 as the resistance source, which each identified a QTL on chromosome five (Kim et al., 2008; Minamiyama et al., 2007). Another group studied this locus in more detail and found a cluster of three QTL on chromosome five, which contributed to resistance in all four of the resistant parents tested (Mallard et al., 2013). Liu et al. used bulked segregant analysis to investigate this QTL and found three candidate genes near the most significantly associated SNP marker, which will be useful for further study of this locus (Liu et al., 2014). Many markers have been developed for *Phyto.5.2*: a dominant SCAR marker (Quirin et al., 2005), AFLP and RAPD markers (Sugita et al., 2006), SSR markers (Minamiyama et al., 2007), a SNAP marker (Kim et al., 2008), SCAR and RAPD markers (Truong et al., 2012), and SNP markers (Liu et al., 2014).

Despite ever-increasing knowledge about the genetics of resistance to *P. capsici*, there have still been no fully successful breeding efforts to introduce the resistance from CdM334 into a commercial pepper line. The first issue encountered is linkage drag, in which resistance genes are linked to undesirable horticultural traits. In F_2 and F_3 populations from a cross between the susceptible 'Americano' and CdM334, hairs on the stem (a trait of CdM334) were strongly correlated with resistance to *P. capsici* (Egea-Gilabert et al., 2008). The authors concluded that the plants that were phenotypically most similar to CdM334 were also most likely to be resistant (Egea-Gilabert et al., 2008). Fruit shape was weakly correlated with fruit rot resistance in a

population derived from a cross between 'Early Jalapeno' and CdM334 (Naegele and Hausbeck, 2014). This correlation is further supported by the co-localization of QTLs for fruit rot resistance, fruit shape, and pericarp thickness (Naegele et al., 2014). The other problem experienced when breeding for resistance is the failure to transfer all needed resistance genes (Lefebvre and Palloix, 1996; Palloix et al., 1990). Backcross breeding (Saini and Sharma, 1978), recurrent selection (Palloix et al., 1990; Thabuis et al., 2004a), and marker-assisted QTL introgression (Thabuis et al., 2004b) have all been attempted, none with wholly successful results.

Squash

Squash (*Cucurbita* spp.) is an important vegetable crop that is popular worldwide. The three main cultivated species of squash are *Cucurbita pepo*, *Cucurbita moschata*, and *Cucurbita maxima*. Squash is native to North and South America and was domesticated up to 10,000 years ago (Ferriol and Picó, 2008). *C. pepo* fruits are consumed at both the immature and mature stages, while C. *moschata* and *C. maxima* fruits are primarily consumed at the mature stage (Ferriol and Picó, 2008). Mature squash, also known as winter squash, is sold and consumed as a fresh market crop and processed as frozen squash and canned pumpkin. Squash fruits are highly nutritious, containing extremely high levels of pro-vitamin A. In addition, they contain high levels of vitamin C and are sources of other nutrients such as calcium, potassium, and iron. Of further nutritional benefit are their high fiber and the low calorie content per serving.

Squash Genomic Resources

Cucurbita species are part of the *Cucurbitaceae* family, which also includes cucumber (*Cucumis sativus*), melon (*Cucumis melo*), and watermelon (*Citrullus lanatus*). The three cultivated *Cucurbita* species (2n = 40) have a genome of approximately 500 Mb (Arumuganathan and Earle, 1991) and have a small but growing number of genomic resources. In the *Cucurbitaceae* family, cucumber has the most extensive genomic resources, with a sequenced genome and many molecular markers (Huang et al., 2009; Zhang et al., 2012). Melon and watermelon also have sequenced genomes and high-quality maps (Clepet et al., 2011; Diaz et al., 2011; Garcia-Mas et al., 2012; Guo et al., 2013; Ren et al., 2012).

Squash has a growing number of genomic resources. Several genomic maps have been generated for *C. pepo* (Gong et al., 2008b; Zraidi et al., 2007), including a recent map by Esteras et al., which was the first to use SNP markers (Esteras et al., 2012). There is also a genetic map comprised of 205 SSR markers for *C. moschata* and a comparative map between *C. pepo* and *C. moschata* with 76 common SSR markers (Gong et al., 2008a). Another major source of genomic information is two sequenced *Cucurbita* transcriptomes. The first transcriptome was comprised of 49,610 unigenes derived from flower, leaf, and root tissue of two summer squash cultivars (Blanca et al., 2011). More than 60% of the unigenes were annotated, aiding in the potential identification of genes of interest in the squash genome. It was used to identify more than 10,000 potential simple sequence repeats (SSRs) and SNPs that can be used to generate molecular markers. In addition, a *C. moschata* transcriptome

consisting of 62,480 unigenes was sequenced from leaf, stem, and shoot tissue (Wu et al., 2014).

Fruit Development and Quality

Squash is primarily grown for the consumption of its fruit, which has a complex anatomy and development. Squash fruit consists of several layers: the exocarp is the outer layer (rind), the mesocarp is the middle layer (fruit flesh), and the endocarp is the thin inner layer around the seeds (Loy, 2004). From pollination to maturity, fruit development takes approximately 60 days. There are three major developmental stages (Irving et al., 1997). Early growth takes place during the first 30 days of fruit development and is when fruit expansion, starch accumulation, and dry matter accumulation occur. For the next 30 days, the fruit matures and sucrose accumulates. After 60 days, fruit ripening occurs and starch is metabolized while sucrose further accumulates.

Fruit quality is especially important for winter squash and is made up of the interaction of several factors including texture, flavor, sweetness, color, and nutrient content. Of the three primary squash species, *C. maxima* generally has the highest fruit quality, while *C. pepo* has the lowest. Although many components contribute to squash fruit quality, research has identified three main metabolites that are easily quantifiable and underlie variation in quality: carotenoid content determines fruit color (Gajewski et al., 2008; Itle and Kabelka, 2009), starch content and dry matter are highly correlated with fruit texture, with higher starch content and thus higher dry matter manifested as superior texture (Corrigan et al., 2001; Cumarasamy et al., 2002; Hurst et al., 2006), and sugar content is important for sweetness and consumer

acceptance and is generally associated with improved squash flavor (Cumarasamy et al., 2002). These three components are each determined by complex, yet well-understood biochemical networks.

In addition to contributing to fruit asthetics in the form of fruit mesocarp color, carotenoids additionally provide nutritional benefits to consumers. The primary carotenoids in squash are α-carotene, β-carotene, and lutein (Azevedo-Meleiro and Rodriguez-Amaya, 2007). Alpha-carotene and β-carotene are sources of pro-vitamin A and approximately 19 mg and 9.5 mg respectively are needed to meet the recommended daily allowance (RDA) for adults (NIH, 2013). Winter squash varies greatly in total pro-vitamin A content, ranging from 3% of the RDA (acorn squash) to over 93% of the RDA (butternut squash) for a one-cup serving, according to the USDA National Nutrient Database (USDA, 2011). Butternut squash is one of the best plant sources of pro-vitamin A, with only a few vegetables, including sweet potato and carrot, providing more pro-vitamin A per one-cup serving (USDA, 2011). Lutein does not have an official RDA, but the American Optometric Association recommends an intake of 10 mg per day to support eye health (AOA, 2014). Winter squash lutein content varies greatly, ranging from 0 mg to 8 mg per one-cup serving (Azevedo-Meleiro and Rodriguez-Amaya, 2007; Itle and Kabelka, 2009). Other high-lutein foods include kale, spinach, and collards, with 15-24 mg lutein per one-cup serving (AOA, 2014).

Squash Seeds and the Hull-less Gene

Although squash was originally domesticated for consumption of its nutritious seeds (Robinson and Decker-Walters, 1997), seeds are currently a neglected

component of the squash fruit. A primary reason for the under-utilization of squash seeds is their tough, unpalatable seed coat. There is, however, a mutation that causes the seeds to be "hull-less" (Stuart and Loy, 1983) that has great potential for increasing consumption of squash seeds. This mutation, known as the hull-less mutation, is thought to have appeared in Austria around 1880 (Teppner, 2000) and causes a dramatic reduction in the development of the seed coat. When this mutation is present, all of the layers of the seed coat are formed, but the outer layers of the seed coat do not undergo normal secondary cell wall development and thus have greatly reduced levels of lignin, cellulose, and hemicellulose (Stuart, 1981; Stuart and Loy, 1983). The hull-less gene is a major recessive gene with modifiers, and a complementation test confirmed that six representative *C. pepo* mutant lines all have mutations at the same locus (Wyatt et al., unpublished). The hull-less gene has been mapped (Gong et al., 2008b; Zraidi et al., 2007), but the published linked markers do not cosegregate with the mutation in all germplasm (Wyatt et al., unpublished).

The hull-less mutation is also of interest because of the high nutritional quality of squash seeds, which contain protein, fiber, unsaturated fatty acids, Vitamin E, phytosterols, and a range of essential minerals (Fruhwirth and Hermetter, 2007; Lazos, 1986; Stevenson et al., 2007). Both wild type and hull-less seeds have been found to vary in fatty acid profile, protein content, and vitamin E content (Idouraine et al., 1996; Murkovic et al., 1999). Hull-less seeds have an enriched fatty acid content as compared to wild type seeds (Fruhwirth and Hermetter, 2007), which provides an opportunity for selection for improved nutritional content.

In wild-type squash seed development, the seed coat develops from the outer integument of the ovule (Teppner, 2004). There are two main seed coat layers, the chlorenchyma, which is a thin, membranous-like, green inner layer containing proto-chlorophyll, and a thick, hard, outer layer made up of the epidermis, hypodermis, schlerenchyma, and aerenchyma (Loy, 2000; Stuart and Loy, 1983; Teppner, 2004). The hypodermis, sclerenchyma, and aerenchyma are lignified and form secondary cell walls, making the seed coat tough (Stuart and Loy, 1983). In plants with the hull-less mutation, however, this secondary cell wall development and lignification do not occur, even though all seed coat layers are present (Loy, 2000; Stuart and Loy, 1983). Specifically, lignin, cellulose, and hemicellulose are all reduced in hull-less seed coats (Loy, 2000; Stuart and Loy, 1983). Because of this lack of hardening and thickening, the outer layers collapse, leaving only the inner, green, chlorenchyma, which develops more cell layers than in wild-type seeds (Stuart and Loy, 1983) and causes the characteristic green appearance of hull-less seeds.

The hull-less gene has yet to be identified, but there are several hypotheses about its function. Our preferred hypothesis is that the hull-less mutation is in a transcription factor controlling secondary cell wall development in the seed coat. Such major regulatory transcription factors are known to control tissue-specific secondary cell wall development in many species such as Arabidopsis, poplar, rice, and Medicago (Mitsuda et al., 2005; Wang et al., 2011; Zhong et al., 2010; Zhong et al., 2011). One line of evidence supporting this hypothesis is that four genes involved in secondary cell wall biosynthesis were found to be down-regulated in seed coats of hull-less mutants (Bezold et al., 2005), as would be expected if the hull-less gene is a

transcription factor that regulates the expression of these and other secondary cell wall biosynthesis genes.

Scope of the Dissertation

In this dissertation, I explore multiple avenues of vegetable crop improvement. Chapter 2 details the development of molecular markers that differentiate the three *Pun1* alleles in pepper and enable breeders to select for or against pungency at the seedling stage. In Chapter 3, the disease resistance and yield of our *Phytophthora capsici*-resistant bell peppers are described in comparison to commercially-available varieties. Chapter 4 moves to the broader genomics level and discusses the sequencing and assembly of the first winter squash fruit and seed transcriptome. Building on this transcriptome, Chapter 5 looks at differential gene expression throughout fruit development in two winter squash cultivars bred for different purposes. Together, the research described contributes to both applied vegetable breeding and knowledge about the genomics underlying major traits of interest.

References

Alcantara, T. and P. Bosland, 1994. An Inexpensive Disease Screening Technique for Foliar Blight of Chile Pepper Seedlings. *Hortscience* 29:1182-1183.

American Optometric Association, 2014. Lutein & Zeaxanthin. American Optometric Association, St. Louis, MO. 21 April 2015. < http://www.aoa.org/patients-and-public/caring-for-your-vision/diet-and-nutrition/lutein?sso=y>

Arumuganathan, K. and E.D. Earle, 1991. Nuclear DNA Content of Some Important Plant Species. *Plant Molecular Biology Reporter* 9:208-218.

Ashrafi, H., T. Hill, K. Stoffel, A. Kozik, J.Q. Yao, S.R. Chin-Wo, and A. Van Deynze, 2012. *De novo* assembly of the pepper transcriptome (*Capsicum annuum*): a benchmark for in silico discovery of SNPs, SSRs and candidate genes. *BMC Genomics* 13.

Azevedo-Meleiro, C.H. and D.B. Rodriguez-Amaya, 2007. Qualitative and quantitative differences in carotenoid composition among *Cucurbita moschata*, *Cucurbita maxima*, and *Cucurbita pepo*. *Journal of Agricultural and Food Chemistry* 55:4027-4033.

Barksdale, T.H., G.C. Papavizas, and S.A. Johnston, 1984. Resistance to Foliar Blight and Crown Rot of Pepper Caused by *Phytophthora capsici*. *Plant Disease* 68:506-509.

Bartual, R., E. Carbonell, J. Marsal, J. Tello, and T. Campos, 1991. Gene action in the resistance of peppers *(Capsicum annuum)* to Phytophthora stem blight *(Phytophthora capsici* L.). *Euphytica* 54:195-200.

Ben-Chaim, A., Y. Borovsky, M. Falise, M. Mazourek, B.C. Kang, I. Paran, and M. Jahn, 2006. QTL analysis for capsaicinoid content in *Capsicum*. *Theoretical and Applied Genetics* 113:1481-1490.

Bennett, D.J. and G.W. Kirby, 1968. Constitution and Biosynthesis of Capsaicin. *Journal of the Chemical Society C-Organic* 442-446.

Bezold, T.N., D. Mathews, J.B. Loy, and S.C. Minocha, 2005. Molecular analysis of the hull-less seed trait in pumpkin: expression profiles of genes related to seed coat development. *Seed Science Research* 15:205-217.

Blanca, J., J. Canizares, C. Roig, P. Ziarsolo, F. Nuez, and B. Pico, 2011. Transcriptome characterization and high throughput SSRs and SNPs discovery in *Cucurbita pepo* (Cucurbitaceae). *BMC Genomics* 12.

Blum, E., M. Mazourek, M. O'Connell, J. Curry, T. Thorup, K.D. Liu, M. Jahn, and I. Paran, 2003. Molecular mapping of capsaicinoid biosynthesis genes and quantitative trait loci analysis for capsaicinoid content in *Capsicum*. *Theoretical and Applied Genetics* 108:79-86.

Bosland, P.W. and D.L. Lindsey, 1991. A Seedling Screen for Phytophthora Root Rot of Pepper, *Capsicum annuum*. *Plant Disease* 75:1048-1050.

Bowers, J.H. and D.J. Mitchell, 1990. Effect of Soil-Water Matric Potential and Periodic Flooding on Mortality of Pepper Caused by *Phytophthora capsici*. *Phytopathology* 80:1447-1450.

Cafe, A.C. and J.M. Duniway, 1995. Effects of Furrow Irrigation Schedules and Host Genotype on Phytophthora Root Rot of Pepper. *Plant Disease* 79:39-43.

Candole, B., P. Conner, and P. Ji, 2010. Screening Capsicum annuum Accessions for Resistance to Six Isolates of *Phytophthora capsici*. *Hortscience* 45:254-259.

Caterina, M.J., M.A. Schumacher, M. Tominaga, T.A. Rosen, J.D. Levine, and D. Julius, 1997. The capsaicin receptor: a heat-activated ion channel in the pain pathway. *Nature* 389:816-824.

Clepet, C., T. Joobeur, Y. Zheng, D. Jublot, M.Y. Huang, V. Truniger, A. Boualem, M.E. Hernandez-Gonzalez, R. Dolcet-Sanjuan, V. Portnoy, A. Mascarell-Creus, A.I. Cano-Delgado, N. Katzir, A. Bendahmane, J.J. Giovannoni, M.A. Aranda, J. Garcia-Mas, and Z.J. Fei, 2011. Analysis of expressed sequence tags generated from full-length enriched cDNA libraries of melon. *BMC Genomics* 12.

Corrigan, V.K., P.L. Hurst, and J.F. Potter, 2001. Winter squash (*Cucurbita maxima*) texture: sensory, chemical, and physical measures. *New Zealand Journal of Crop and Horticultural Science* 29:111-124.

Crosby, K.M., 2008. Pepper, p. 221-248, Vegetables II. Springer New York.

Crossan, D.F., F.A. Haasis, and D.E. Ellis, 1953. Phytophthora Blight of Summer Squash in North Carolina. *Phytopathology* 43:469-469.

Cumarasamy, R., V. Corrigan, P. Hurst, and M. Bendall, 2002. Cultivar differences in New Zealand "Kabocha" (buttercup squash, *Cucurbita maxima*). *New Zealand Journal of Crop and Horticultural Science* 30:197-208.

Diaz, A., M. Fergany, G. Formisano, P. Ziarsolo, J. Blanca, Z.J. Fei, J.E. Staub, J.E. Zalapa, H.E. Cuevas, G. Dace, M. Oliver, N. Boissot, C. Dogimont, M. Pitrat, R. Hofstede, P. van Koert, R. Harel-Beja, G. Tzuri, V. Portnoy, S. Cohen, A.

Schaffer, N. Katzir, Y. Xu, H.Y. Zhang, N. Fukino, S. Matsumoto, J. Garcia-Mas, and A.J. Monforte, 2011. A consensus linkage map for molecular markers and Quantitative Trait Loci associated with economically important traits in melon (*Cucumis melo* L.). *BMC Plant Biology* 11.

Egea-Gilabert, C., G. Bilotti, M. Requena, M. Ezziyyani, J. Vivo-Molina, and M. Candela, 2008. Pepper morphological traits related with resistance to *Phytophthora capsici*. *Biologia Plantarum* 52:105-109.

Esteras, C., P. Gomez, A.J. Monforte, J. Blanca, N. Vicente-Dolera, C. Roig, F. Nuez, and B. Pico, 2012. High-throughput SNP genotyping in *Cucurbita pepo* for map construction and quantitative trait loci mapping. *BMC Genomics* 13.

Estrada, B., F. Pomar, J. Diaz, F. Merino, and M.A. Bernal, 1999. Pungency level in fruits of the Padron pepper with different water supply. *Scientia Horticulturae* 81:385-396.

Ferriol, M. and B. Picó, 2008. Pumpkin and Winter Squash, p. 317-349. In: J. Prohens and F. Nuez (eds.), Vegetables I. Springer New York, New York.

Fruhwirth, G.O. and A. Hermetter, 2007. Seeds and oil of the Styrian oil pumpkin: Components and biological activities. *European Journal of Lipid Science and Technology* 109:1128-1140.

Gajewski, M., J. Radzanowska, H. Danilcenko, E. Jariene, and J. Cerniauskiene, 2008. Quality of Pumpkin Cultivars in Relation to Sensory Characteristics. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca* 36:73-79.

Garcia-Mas, J., A. Benjak, W. Sanseverino, M. Bourgeois, G. Mir, V.M. Gonzalez, E. Henaff, F. Camara, L. Cozzuto, E. Lowy, T. Alioto, S. Capella-Gutierrez, J. Blanca, J. Canizares, P. Ziarsolo, D. Gonzalez-Ibeas, L. Rodriguez-Moreno, M. Droege, L. Du, M. Alvarez-Tejado, B. Lorente-Galdos, M. Mele, L.M. Yang, Y.Q. Weng, A. Navarro, T. Marques-Bonet, M.A. Aranda, F. Nuez, B. Pico, T. Gabaldon, G. Roma, R. Guigo, J.M. Casacuberta, P. Arus, and P. Puigdomenech, 2012. The genome of melon (*Cucumis melo* L.). *Proceedings of the National Academy of Sciences of the United States of America* 109:11872-11877.

Garcia-Rodriguez, M.R., E. Chiquito-Almanza, P.D. Loeza-Lara, H. Godoy-Hernandez, E.V. Pineda, J.L. Pons-Hernandez, M.M. Gonzalez-Chavira, and J.L. Anaya-Lopez, 2010. Production of Ancho Chili Graft on Criollo de Morelos 334 for the Control of *Phytophthora capsici*. *Agrociencia* 44:701-709.

Glosier, B., E. Ogundiwin, G. Sidhu, D. Sischo, and J. Prince, 2008. A differential series of pepper (*Capsicum annuum*) lines delineates fourteen physiological races of *Phytophthora capsici*. *Euphytica* 162:23-30.

Gong, L., M. Pachner, K. Kalai, and T. Lelley, 2008a. SSR-based genetic linkage map of *Cucurbita moschata* and its synteny with *Cucurbita pepo*. *Genome* 51:878-887.

Gong, L., G. Stift, R. Kofler, M. Pachner, and T. Lelley, 2008b. Microsatellites for the genus *Cucurbita* and an SSR-based genetic linkage map of *Cucurbita pepo* L. *Theoretical and Applied Genetics* 117:37-48.

Granke, L.L., S.T. Windstam, H.C. Hoch, C.D. Smart, and M.K. Hausbeck, 2009. Dispersal and Movement Mechanisms of *Phytophthora capsici* Sporangia. *Phytopathology* 99:1258-1264.

Guerrero-Moreno, A. and J. Laborde, 1980. Current status of pepper breeding for resistance to *Phytophthora capsici* in Mexico. *Synopses of the IVth Eucarpia Meeting on Capsicum Wageningen (Netherlands)* 52-56.

Guo, S.G., J.G. Zhang, H.H. Sun, J. Salse, W.J. Lucas, H.Y. Zhang, Y. Zheng, L.Y. Mao, Y. Ren, Z.W. Wang, J.M. Min, X.S. Guo, F. Murat, B.K. Ham, Z.L. Zhang, S. Gao, M.Y. Huang, Y.M. Xu, S.L. Zhong, A. Bombarely, L.A. Mueller, H. Zhao, H.J. He, Y. Zhang, Z.H. Zhang, S.W. Huang, T. Tan, E.L. Pang, K. Lin, Q. Hu, H.H. Kuang, P.X. Ni, B. Wang, J.A. Liu, Q.H. Kou, W.J. Hou, X.H. Zou, J. Jiang, G.Y. Gong, K. Klee, H. Schoof, Y. Huang, X.S. Hu, S.S. Dong, D.Q. Liang, J. Wang, K. Wu, Y. Xia, X. Zhao, Z.Q. Zheng, M. Xing, X.M. Liang, B.Q. Huang, T. Lv, J.Y. Wang, Y. Yin, H.P. Yi, R.Q. Li, M.Z. Wu, A. Levi, X.P. Zhang, J.J. Giovannoni, Y.F. Li, Z.J. Fei, and Y. Xu, 2013. The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions. *Nature Genetics* 45:51-58.

Gupta, S., X.L. Shi, I.E. Lindquist, N. Devitt, J. Mudge, and A.M. Rashotte, 2013. Transcriptome profiling of cytokinin and auxin regulation in tomato root. *Journal of Experimental Botany* 64:695-704.

Hamilton, J.P., S.C. Sim, K. Stoffel, A. Van Deynze, C.R. Buell, and D.M. Francis, 2012. Single Nucleotide Polymorphism Discovery in Cultivated Tomato via Sequencing by Synthesis. *Plant Genome* 5:17-29.

Harvell, K.P. and P.W. Bosland, 1997. The environment produces a significant effect on pungency of chiles. *Hortscience* 32:1292-1292.

Hausbeck, M. and K. Lamour, 2004. *Phytophthora capsici* on vegetable crops: Research progress and management challenges. *Plant Disease* 88:1292-1303.

Hill, T.A., H. Ashrafi, S. Reyes-Chin-Wo, J.Q. Yao, K. Stoffel, M.J. Truco, A. Kozik, R.W. Michelmore, and A. Van Deynze, 2013. Characterization of *Capsicum annuum* Genetic Diversity and Population Structure Based on Parallel Polymorphism Discovery with a 30K Unigene Pepper GeneChip. *Plos One* 8.

Huang, S.W., R.Q. Li, Z.H. Zhang, L. Li, X.F. Gu, W. Fan, W.J. Lucas, X.W. Wang, B.Y. Xie, P.X. Ni, Y.Y. Ren, H.M. Zhu, J. Li, K. Lin, W.W. Jin, Z.J. Fei, G.C. Li, J. Staub, A. Kilian, E.A.G. van der Vossen, Y. Wu, J. Guo, J. He, Z.Q. Jia, Y. Ren, G. Tian, Y. Lu, J. Ruan, W.B. Qian, M.W. Wang, Q.F. Huang, B. Li, Z.L. Xuan, J.J. Cao, Asan, Z.G. Wu, J.B. Zhang, Q.L. Cai, Y.Q. Bai, B.W. Zhao, Y.H. Han, Y. Li, X.F. Li, S.H. Wang, Q.X. Shi, S.Q. Liu, W.K. Cho, J.Y. Kim, Y. Xu, K. Heller-Uszynska, H. Miao, Z.C. Cheng, S.P. Zhang, J. Wu, Y.H. Yang, H.X. Kang, M. Li, H.Q. Liang, X.L. Ren, Z.B. Shi, M. Wen, M. Jian, H.L. Yang, G.J. Zhang, Z.T. Yang, R. Chen, S.F. Liu, J.W. Li, L.J. Ma, H. Liu, Y. Zhou, J. Zhao, X.D. Fang, G.Q. Li, L. Fang, Y.R. Li, D.Y. Liu, H.K. Zheng, Y. Zhang, N. Qin, Z. Li, G.H. Yang, S. Yang, L. Bolund, K. Kristiansen, H.C. Zheng, S.C. Li, X.Q. Zhang, H.M. Yang, J. Wang, R.F. Sun, B.X. Zhang, S.Z. Jiang, Y.C. Du, and S.G. Li, 2009. The genome of the cucumber, *Cucumis sativus* L. *Nature Genetics* 41:1275-1281.

Hurst, P.L., V.K. Corrigan, and J. Koolaard, 2006. Genetic analysis of sweetness and textural attributes in winter squash (*Cucurbita maxima*). *New Zealand Journal of Crop and Horticultural Science* 34:359-367.

Hwang, B.K. and C.H. Kim, 1995. Phytophthora Blight of Pepper and Its Control in Korea. *Plant Disease* 79:221-227.

Idouraine, A., E.A. Kohlhepp, C.W. Weber, W.A. Warid, and J.J. Martinez-Tellez, 1996. Nutrient constituents from eight lines of naked seed squash (*Cucurbita pepo* L). *Journal of Agricultural and Food Chemistry* 44:721-724.

Irving, D.E., P.L. Hurst, and J.S. Ragg, 1997. Changes in carbohydrates and carbohydrate metabolizing enzymes during the development, maturation, and ripening of buttercup squash (*Cucurbita maxima* D 'Delica'). *Journal of the American Society for Horticultural Science* 122:310-314.

Itle, R.A. and E.A. Kabelka, 2009. Correlation Between L*a*b* Color Space Values and Carotenoid Content in Pumpkins and Squash (*Cucurbita* spp.). *Hortscience* 44:633-637.

Kim, H.J., S.H. Nahm, H.R. Lee, G.B. Yoon, K.T. Kim, B.C. Kang, D. Choi, O. Kweon, M.C. Cho, J.K. Kwon, J.H. Han, J.H. Kim, M. Park, J. Ahn, S. Choi, N. Her, J.H. Sung, and B.D. Kim, 2008. BAC-derived markers converted from RFLP linked to *Phytophthora capsici* resistance in pepper (*Capsicum annuum* L.). *Theoretical and Applied Genetics* 118:15-27.

Kim, S., M. Park, S.I. Yeom, Y.M. Kim, J.M. Lee, H.A. Lee, E. Seo, J. Choi, K. Cheong, K.T. Kim, K. Jung, G.W. Lee, S.K. Oh, C. Bae, S.B. Kim, H.Y. Lee, S.Y. Kim, M.S. Kim, B.C. Kang, Y.D. Jo, H.B. Yang, H.J. Jeong, W.H. Kang, J.K. Kwon, C. Shin, J.Y. Lim, J.H. Park, J.H. Huh, J.S. Kim, B.D. Kim, O. Cohen, I. Paran, M.C. Suh, S.B. Lee, Y.K. Kim, Y. Shin, S.J. Noh, J. Park, Y.S. Seo, S.Y. Kwon, H.A. Kim, J.M. Park, H.J. Kim, S.B. Choi, P.W. Bosland, G. Reeves, S.H. Jo, B.W. Lee, H.T. Cho, H.S. Choi, M.S. Lee, Y. Yu, Y. Do Choi, B.S. Park, A. van Deynze, H. Ashrafi, T. Hill, W.T. Kim, H.S. Pai, H.K. Ahn, I. Yeam, J.J. Giovannoni, J.K.C. Rose, I. Sorensen, S.J. Lee, R.W. Kim, I.Y. Choi, B.S. Choi, J.S. Lim, Y.H. Lee, and D. Choi, 2014. Genome sequence of the hot pepper provides insights into the evolution of pungency in *Capsicum* species. *Nature Genetics* 46:270-278.

Kimble, K. and R. Grogan, 1960. Resistance to Phytophthora Root Rot in Pepper. *Plant Disease Reporter* 44:872-873.

Kobata, K., K. Sutoh, T. Todo, S. Yazawa, K. Iwai, and T. Watanabe, 1999. Nordihydrocapsiate, a new capsinoid from the fruits of a nonpungent pepper, *Capsicum annuum*. *Journal of Natural Products* 62:335-336.

Kobata, K., T. Todo, S. Yazawa, K. Iwai, and T. Watanabe, 1998. Novel capsaicinoid-like substances, capsiate and dihydrocapsiate, from the fruits of a nonpungent cultivar, CH-19 sweet, of pepper (*Capsicum annuum* L.). *Journal of Agricultural and Food Chemistry* 46:1695-1697.

Lamour, K.H. and M.K. Hausbeck, 2001. The dynamics of mefenoxam insensitivity in a recombining population of *Phytophthora capsici* characterized with amplified fragment length polymorphism markers. *Phytopathology* 91:553-557.

Lang, Y.Q., H. Kisaka, R. Sugiyama, K. Nomura, A. Morita, T. Watanabe, Y. Tanaka, S. Yazawa, and T. Miwa, 2009. Functional loss of pAMT results in biosynthesis of capsinoids, capsaicinoid analogs, in *Capsicum annuum* cv. CH-19 Sweet. *Plant Journal* 59:953-961.

Lazos, E.S., 1986. Nutritional, Fatty Acid, and Oil Characteristics of Pumpkin and Melon Seeds. *Journal of Food Science* 51:1382-1383.

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

Lee, S. and D. Choi, 2013. Comparative transcriptome analysis of pepper (*Capsicum annuum*) revealed common regulons in multiple stress conditions and hormone treatments. *Plant Cell Reports* 32:1351-1359.

Leete, E. and M.C.L. Louden, 1968. Biosynthesis of Capsaicin and Dihydrocapsaicin in *Capsicum frutescens*. *Journal of the American Chemical Society* 90:6837-6841.

Lefebvre, V. and A. Palloix, 1996. Both epistatic and additive effects of QTLs are involved in polygenic induced resistance to disease: A case study, the interaction pepper - *Phytophthora capsici* Leonian. *Theoretical and Applied Genetics* 93:503-511.

Lefebvre, V., S. Pflieger, A. Thabuis, C. Caranta, A. Blattes, J.C. Chauvet, A.M. Daubeze, and A. Palloix, 2002. Towards the saturation of the pepper linkage map by alignment of three intraspecific maps including known-function genes. *Genome* 45:839-854.

Leonian, L., 1922. Stem and fruit blight of peppers caused by *Phytophthora capsici* sp nov. *Phytopathology* 12:401-408.

Liu, C., N. Ma, P.Y. Wang, N. Fu, and H.L. Shen, 2013a. Transcriptome Sequencing and De Novo Analysis of a Cytoplasmic Male Sterile Line and Its Near-Isogenic Restorer Line in Chili Pepper (*Capsicum annuum* L.). *Plos One* 8.

Liu, S.Q., W.S. Li, Y.M. Wu, C.M. Chen, and J.J. Lei, 2013b. De Novo Transcriptome Assembly in Chili Pepper (*Capsicum frutescens*) to Identify Genes Involved in the Biosynthesis of Capsaicinoids. *Plos One* 8.

Liu, W.Y., J.H. Kang, H.S. Jeong, H.J. Choi, H.B. Yang, K.T. Kim, D. Choi, G.J. Choi, M. Jahn, and B.C. Kang, 2014. Combined use of bulked segregant analysis and microarrays reveals SNP markers pinpointing a major QTL for resistance to *Phytophthora capsici* in pepper. *Theoretical and Applied Genetics* 127:2503-2513.

Livingstone, K.D., V.K. Lackney, J.R. Blauth, R. van Wijk, and M.K. Jahn, 1999. Genome mapping in *Capsicum* and the evolution of genome structure in the Solanaceae. *Genetics* 152:1183-1202.

Loy, J.B., 2000. Seed Development in *Cucurbita pepo*: An Overview with Emphasis on Hull-less Seeded Genotypes of Pumpkin. *Cucurbit Genetics Cooperative Report* 23:89-95.

Loy, J.B., 2004. Morpho-physiological aspects of productivity and quality in squash and pumpkins (*Cucurbita* spp.). *Critical Reviews in Plant Sciences* 23:337-363.

Lu, F.H., S.W. Kwon, M.Y. Yoon, K.T. Kim, M.C. Cho, M.K. Yoon, and Y.J. Park, 2012. SNP Marker Integration and QTL Analysis of 12 Agronomic and Morphological Traits in F-8 RILs of Pepper (*Capsicum annuum* L.). *Molecules and Cells* 34:25-34.

Lu, F.H., M.Y. Yoon, Y.I. Cho, J.W. Chung, K.T. Kim, M.C. Cho, S.R. Cheong, and Y.J. Park, 2011. Transcriptome analysis and SNP/SSR marker information of red pepper variety YCM334 and Taean. *Scientia Horticulturae* 129:38-45.

Mallard, S., M. Cantet, A. Massire, A. Bachellez, S. Ewert, and V. Lefebvre, 2013. A key QTL cluster is conserved among accessions and exhibits broad-spectrum resistance to *Phytophthora capsici*: a valuable locus for pepper breeding. *Molecular Breeding* 32:349-364.

Martinez-Lopez, L.A., N. Ochoa-Alejo, and O. Martinez, 2014. Dynamics of the chili pepper transcriptome during fruit development. *BMC Genomics* 15.

Minamiyama, Y., M. Tsuro, T. Kubo, and M. Hirai, 2007. QTL analysis for resistance to *Phytophthora capsici* in pepper using a high density SSR-based map. *Breeding Science* 57:129-134.

Mitsuda, N., M. Seki, K. Shinozaki, and M. Ohme-Takagi, 2005. The NAC transcription factors NST1 and NST2 of Arabidopsis regulate secondary wall thickenings and are required for anther dehiscence. *Plant Cell* 17:2993-3006.

Mo, H., S. Kim, K.P.P. Wai, M.I. Siddique, H. Yoo, and B.S. Kim, 2014. New sources of resistance to *Phytophthora capsici* in *Capsicum* spp. *Horticulture Environment and Biotechnology* 55:50-55.

Monroy-Barbosa, A. and P.W. Bosland, 2008. Genetic Analysis of Phytophthora Root Rot Race-specific Resistance in Chile Pepper. *Journal of the American Society for Horticultural Science* 133:825-829.

Murkovic, M., A. Hillebrand, S. Draxl, J. Winkler, and W. Pfannhauser. 1999. Distribution of fatty acids and Vitamin E content in pumpkin seeds (*Cucurbita pepo* L.) in breeding lines. *Proceedings of the 1st International Symposium on Cucurbits* 47-55.

Naegele, R.P., H. Ashrafi, T.A. Hill, S.R. Chin-Wo, A.E. Van Deynze, and M.K. Hausbeck, 2014. QTL Mapping of Fruit Rot Resistance to the Plant Pathogen

Phytophthora capsici in a Recombinant Inbred Line Capsicum annuum Population. *Phytopathology* 104:479-483.

Naegele, R.P. and M.K. Hausbeck, 2014. Evaluation of Pepper Fruit for Resistance to *Phytophthora capsici* in a Recombinant Inbred Line Population, and the Correlation with Fruit Shape. *Plant Disease* 98:885-890.

National Institutes of Health, 2013. Vitamin A - Fact Sheet for Health Professionals. National Institutes of Health Office of Dietary Supplements, Bethesda, MD. 21 April 2015. < http://ods.od.nih.gov/factsheets/VitaminA-HealthProfessional/#h2>

Oelke, L., P. Bosland, and R. Steiner, 2003. Differentiation of race specific resistance to phytophthora root rot and foliar blight in *Capsicum annuum*. *Journal of the American Society for Horticultural Science* 128:213-218.

Ogundiwin, E., T. Berke, M. Massoudi, L. Black, G. Huestis, D. Choi, S. Lee, and J. Prince, 2005. Construction of 2 intraspecific linkage maps and identification of resistance QTLs for *Phytophthora capsici* root-rot and foliar-blight diseases of pepper (*Capsicum annuum* L.). *Genome* 48:698-711.

Ortega, R.G., C.P. Espanol, and J.C. Zueco, 1991. Genetics of Resistance to *Phytophthora capsici* in the Pepper Line 'SCM-334'. *Plant Breeding* 107:50-55.

Ortega, R.G., C.P. Espanol, and J.C. Zueco, 1992. Genetic Relationships Among Four Pepper Genotypes Resistant to *Phytophthora capsici*. *Plant Breeding* 108:118-125.

Palloix, A., A. Daubeze, T. Phaly, and E. Pochard, 1990. Breeding transgressive lines of pepper for resistance to *Phytophthora capsici* in a recurrent selection system. *Euphytica* 51:141-150.

Polach, F. and R. Webster, 1972. Identification of Strains and Inheritance of Pathogenicity in *Phytophthora capsici*. *Phytopathology* 62:20-26.

Prince, J.P., E. Pochard, and S.D. Tanksley, 1993. Construction of a molecular linkage map of pepper and a comparison of synteny with tomato. *Genome* 36:404-417.

Quirin, E., E. Ogundiwin, J. Prince, M. Mazourek, M. Briggs, T. Chlanda, K. Kim, M. Falise, B. Kang, and M. Jahn, 2005. Development of sequence characterized amplified region (SCAR) primers for the detection of Phyto.5.2, a major QTL for resistance to *Phytophthora capsici* Leon. in pepper. *Theoretical and Applied Genetics* 110:605-612.

Reifschneider, F.J.B., L.S. Boiteux, P.T. Dellavecchia, J.M. Poulos, and N. Kuroda, 1992. Inheritance of adult-plant resistance to *Phytophthora capsic*i in pepper *Euphytica* 62:45-49.

Ren, Y., H. Zhao, Q. Kou, J. Jiang, S. Guo, H. Zhang, W. Hou, X. Zou, H. Sun, G. Gong, A. Levi, and Y. Xu, 2012. A high resolution genetic map anchoring scaffolds of the sequenced watermelon genome. *Plos One* 7.

Ristaino, J.B., 1991. Influence of Rainfall, Drip Irrigation, and Inoculum Density on the Development of Phytophthora Root and Crown Rot Epidemics and Yield in Bell Pepper. *Phytopathology* 81:922-929.

Ristaino, J.B. and S.A. Johnston, 1999. Ecologically based approaches to management of Phytophthora blight on Bell pepper. *Plant Disease* 83:1080-1089.

Robinson, R.W. and D.S. Decker-Walters, 1997. Cucurbits. CAB International, New York.

Ruiz-Lau, N., F. Medina-Lara, Y. Minero-Garcia, E. Zamudio-Moreno, A. Guzman-Antonio, I. Echevarria-Machado, and M. Martinez-Estevez, 2011. Water Deficit Affects the Accumulation of Capsaicinoids in Fruits of *Capsicum chinense* Jacq. *Hortscience* 46:487-492.

Saini, S.S. and P.P. Sharma, 1978. Inheritance of resistance to fruit rot Leon.) and induction of resistance in bell pepper (*Capsicum annuum* L.). *Euphytica* 27:721-723.

Saritnum, O., M. Minami, K.I. Matsushima, Y. Minamiyama, M. Hirai, T. Baba, H. Bansho, and K. Nemoto, 2008. Inheritance of few-pungent trait in chili pepper 'S3212' (*Capsicum frutescens*). *Journal of the Japanese Society for Horticultural Science* 77:265-269.

Sato, S., S. Tabata, H. Hirakawa, E. Asamizu, K. Shirasawa, S. Isobe, T. Kaneko, Y. Nakamura, D. Shibata, K. Aoki, M. Egholm, J. Knight, R. Bogden, C.B. Li, Y. Shuang, X. Xu, S.K. Pan, S.F. Cheng, X. Liu, Y.Y. Ren, J. Wang, A. Albiero, F. Dal Pero, S. Todesco, J. Van Eck, R.M. Buels, A. Bombarely, J.R. Gosselin, M.Y. Huang, J.A. Leto, N. Menda, S. Strickler, L.Y. Mao, S. Gao, I.Y. Tecle, T. York, Y. Zheng, J.T. Vrebalov, J. Lee, S.L. Zhong, L.A. Mueller, W.J. Stiekema, P. Ribeca, T. Alioto, W.C. Yang, S.W. Huang, Y.C. Du, Z.H. Zhang, J.C. Gao, Y.M. Guo, X.X. Wang, Y. Li, J. He, C.Y. Li, Z.K. Cheng, J.R. Zuo, J.F. Ren, J.H. Zhao, L.H. Yan, H.L. Jiang, B. Wang, H.S. Li, Z.J. Li, F.Y. Fu, B.T. Chen, B. Han, Q. Feng, D.L. Fan, Y. Wang, H.Q. Ling, Y.B.A. Xue, D. Ware, W.R. McCombie, Z.B. Lippman, J.M. Chia, K. Jiang, S. Pasternak, L. Gelley, M. Kramer, L.K. Anderson, S.B. Chang, S.M. Royer, L.A. Shearer, S.M. Stack, J.K.C. Rose, Y.M. Xu, N. Eannetta, A.J. Matas, R.

McQuinn, S.D. Tanksley, F. Camara, R. Guigo, S. Rombauts, J. Fawcett, Y. Van de Peer, D. Zamir, C.B. Liang, M. Spannagl, H. Gundlach, R. Bruggmann, et al., 2012. The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* 485:635-641.

Schlub, R.L., 1983. Epidemiology of *Phytophthora capsici* on bell pepper. *Journal of Agricultural Science* 100:7-11.

Smith, P., K. Kimble, R. Grogan, and A. Millet, 1967. Inheritance of Resistance in Peppers to Phytophthora Root Rot. *Phytopathology* 57:377-379.

Stellari, G.M., M. Mazourek, and M.M. Jahn, 2010. Contrasting modes for loss of pungency between cultivated and wild species of *Capsicum*. *Heredity* 104:460-471.

Stevenson, D.G., F.J. Eller, L.P. Wang, J.L. Jane, T. Wang, and G.E. Inglett, 2007. Oil and tocopherol content and composition of pumpkin seed oil in 12 cultivars. *Journal of Agricultural and Food Chemistry* 55:4005-4013.

Stewart, C., B.C. Kang, K. Liu, M. Mazourek, S.L. Moore, E.Y. Yoo, B.D. Kim, I. Paran, and M.M. Jahn, 2005. The *Pun1* gene for pungency in pepper encodes a putative acyltransferase. *Plant Journal* 42:675-688.

Stewart, C., M. Mazourek, G.M. Stellari, M. O'Connell, and M. Jahn, 2007. Genetic control of pungency in *C. chinense* via the *Pun1* locus. *Journal of Experimental Botany* 58:979-991.

Stuart, S.G., 1981. Comparative studies of testa development in normal and hull-less seed strains of *Cucurbita pepo* L. University of New Hampshire, MS thesis.

Stuart, S.G. and J.B. Loy, 1983. Comparison of testa development in normal and hull-less seeded strains of *Cucurbita pepo* L. *Botanical Gazette* 144:491-500.

Sugita, T., K. Yamaguchi, T. Kinoshita, K. Yuji, Y. Sugimura, R. Nagata, S. Kawasaki, and A. Todoroki, 2006. QTL analysis for resistance to phytophthora blight (*Phytophthora capsici* Leon.) using an intraspecific doubled-haploid population of *Capsicum annuum*. *Breeding Science* 56:137-145.

Sy, O., P. Bosland, and R. Steiner, 2005. Inheritance of phytophthora stem blight resistance as compared to phytophthora root rot and phytophthora foliar blight resistance in *Capsicum annuum* L. *Journal of the American Society for Horticultural Science*:75-78.

Sy, O., R. Steiner, and P. Bosland, 2008. Recombinant inbred line differential identifies race-specific resistance to Phytophthora root rot in *Capsicum annuum*. *Phytopathology* 98:867-870.

Tanaka, Y., M. Hosokawa, T. Miwa, T. Watanabe, and S. Yazawa, 2010. Novel Loss-of-Function putative aminotransferase Alleles Cause Biosynthesis of Capsinoids, Nonpungent Capsaicinoid Analogues, in Mildly Pungent Chili Peppers (*Capsicum chinense*). *Journal of Agricultural and Food Chemistry* 58:11762-11767.

Teppner, H., 2000. *Cucurbita pepo* (Cucurbitaceae) - History, seed coat types, thin coated seeds and their genetics. *Phyton-Annales Rei Botanicae* 40:1-42.

Teppner, H., 2004. Notes on *Lagenaria* and *Cucurbita* (Cucurbitaceae) - Review and new contributions. *Phyton-Annales Rei Botanicae* 44:245-308.

Thabuis, A., V. Lefebvre, G. Bernard, A. Daubeze, T. Phaly, E. Pochard, and A. Palloix, 2004a. Phenotypic and molecular evaluation of a recurrent selection program for a polygenic resistance to *Phytophthora capsici* in pepper. *Theoretical and Applied Genetics*:342-351.

Thabuis, A., A. Palloix, S. Pflieger, A. Daubeze, C. Caranta, and V. Lefebvre, 2003. Comparative mapping of *Phytophthora* resistance loci in pepper germplasm: evidence for conserved resistance loci across Solanaceae and for a large genetic diversity. *Theoretical and Applied Genetics* 109:1473-1485.

Thabuis, A., A. Palloix, B. Servin, A. Daubeze, P. Signoret, F. Hospital, and V. Lefebvre, 2004b. Marker-assisted introgression of 4 *Phytophthora capsici* resistance QTL alleles into a bell pepper line: validation of additive and epistatic effects. *Molecular Breeding* 14:9-20.

Truong, H.T.H., K.T. Kim, D.W. Kim, S. Kim, Y. Chae, J.H. Park, D.G. Oh, and M.C. Cho, 2012. Identification of isolate-specific resistance QTLs to phytophthora root rot using an intraspecific recombinant inbred line population of pepper (*Capsicum annuum*). *Plant Pathology* 61:48-56.

U. S. Department of Agriculture, 2011. USDA National Nutrient Database for Standard Reference. U.S. Department of Agriculture, Washington, DC. 21 April 2015. < http://ndb.nal.usda.gov//>

Walker, S. and P. Bosland, 1999. Inheritance of Phytophthora root rot and foliar blight resistance in pepper. *Journal of the American Society for Horticultural Science* 124:14-18.

Wang, H.Z., Q. Zhao, F. Chen, M.Y. Wang, and R.A. Dixon, 2011. NAC domain function and transcriptional control of a secondary cell wall master switch. *Plant Journal* 68:1104-1114.

Wu, F.N., N.T. Eannetta, Y.M. Xu, R. Durrett, M. Mazourek, M.M. Jahn, and S.D. Tanksley, 2009. A COSII genetic map of the pepper genome provides a detailed picture of synteny with tomato and new insights into recent chromosome evolution in the genus *Capsicum*. *Theoretical and Applied Genetics* 118:1279-1293.

Wu, T., S. Luo, R. Wang, Y. Zhong, X. Xu, Y.e. Lin, X. He, B. Sun, and H. Huang, 2014. The first Illumina-based *de novo* transcriptome sequencing and analysis of pumpkin (*Cucurbita moschata* Duch.) and SSR marker development. *Molecular Breeding* 34:1437-1447.

Yarnes, S.C., H. Ashrafi, S. Reyes-Chin-Wo, T.A. Hill, K.M. Stoffel, and A. Van Deynze, 2013. Identification of QTLs for capsaicinoids, fruit quality, and plant architecture-related traits in an interspecific *Capsicum* RIL population. *Genome* 56:61-74.

Zamski, E., O. Shoham, D. Palevitch, and A. Levy, 1987. Ultrastructure of capsacinoid-secreting cells in pungent and nonpungent red pepper (*Capsicum annuum* L.) cultivars. *Botanical Gazette* 148:1-6.

Zewdie, Y. and P.W. Bosland, 2000. Evaluation of genotype, environment, and genotype-by-environment interaction for capsaicinoids in *Capsicum annuum* L. *Euphytica* 111:185-190.

Zhang, W.W., J.S. Pan, H.L. He, C. Zhang, Z. Li, J.L. Zhao, X.J. Yuan, L.H. Zhu, S.W. Huang, and R. Cai, 2012. Construction of a high density integrated genetic map for cucumber (*Cucumis sativus* L.). *Theoretical and Applied Genetics* 124:249-259.

Zhong, R.Q., C. Lee, R.L. McCarthy, C.K. Reeves, E.G. Jones, and Z.H. Ye, 2011. Transcriptional Activation of Secondary Wall Biosynthesis by Rice and Maize NAC and MYB Transcription Factors. *Plant and Cell Physiology* 52:1856-1871.

Zhong, R.Q., C.H. Lee, and Z.H. Ye, 2010. Functional Characterization of Poplar Wood-Associated NAC Domain Transcription Factors. *Plant Physiology* 152:1044-1055.

Zraidi, A., G. Stift, M. Pachner, A. Shojaeiyan, L. Gong, and T. Lelley, 2007. A consensus map for *Cucurbita pepo*. *Molecular Breeding* 20:375-388.

CHAPTER 2

DEVELOPMENT AND APPLICATION OF A SUITE OF NON-PUNGENCY MARKERS FOR THE *PUN1* GENE IN PEPPER (*CAPSICUM* SPP.)[1]

Abstract

Pungency in peppers is due to the presence of capsaicinoid molecules, which are only produced in *Capsicum* species. The major gene *Pun1* is required for the production of capsaicinoids. Three distinct mutant alleles of *Pun1* have been found in three cultivated *Capsicum* species, one of which has been widely utilized by breeders. Although these mutations have been previously identified, a robust collection of molecular markers for the set of alleles is not available. This has been hindered by the existence of at least one paralogous locus that tends to amplify with *Pun1*. We present a suite of markers that can differentiate the four *Pun1* alleles and test them on a diverse panel of pepper lines and in an F2 population segregating for pungency. These markers will be useful for pepper breeding, germplasm characterization, and seed purity testing.

Introduction

Capsaicinoids, the molecules that cause a pungent, burning sensation when hot peppers are consumed, are produced exclusively in the genus *Capsicum*. This organoleptic quality is due to the activation of the TRPV1 (VR1) receptor (Caterina et al., 1997). The primary capsaicinoids are capsaicin, dihydrocapsaicin, and

[1] This chapter was originally published in *Molecular Breeding* and is re-printed with kind permission from Springer Science+Business Media: Wyatt, L., Eanetta, N., Stellari, G. and M. Mazourek. 2012. Development and application of a suite of non-pungency markers for the *Pun1* gene in pepper (*Capsicum* spp.). *Molecular Breeding* 30: 1525-1529.

nordihydrocapsaicin (Bennett and Kirby, 1968). The presence of capsaicinoids makes pungent peppers valuable as a spice. In contrast, the absence of capsaicinoids is important when non-pungent peppers are grown as a vegetable crop.

Although several genes are known to affect capsaicinoid levels, mutations in the *Pun1* locus are the most well-known and widely deployed in pepper germplasm. Three mutant alleles of *Pun1* have been identified. The first, $pun1^1$, has been found exclusively in *C. annuum* and is the allele that has been widely used to breed non-pungent bell peppers. The $pun1^1$ allele has a 2.5 kb deletion that eliminates the putative promoter region and most of the first exon, causing the gene to be neither transcribed nor translated (Stewart et al., 2005). The $pun1^2$ allele is characterized by a 4 bp deletion in the first exon that creates a premature stop codon. This allele is transcribed, but no protein product is produced (Stewart et al., 2007). The third known allele is $pun1^3$, which has a large deletion of the 3' end of the gene resulting in a loss of the final 70 amino acids of the *Pun1* protein (Stellari et al., 2010). The $pun1^3$ allele is not transcribed or translated (Stellari et al., 2010).

Either the presence or absence of capsaicin, caused by the presence of the wild type or mutant alleles of *Pun1* respectively, can be desirable in pepper breeding programs. Because pungent and non-pungent germplasm are often intermated in breeding programs, it is important to be able to determine the allelic state of *Pun1* in breeding populations. Multiple markers have been developed for $pun1^1$ (Blum et al., 2002; Minamiyama et al., 2005; Lee et al., 2005; Truong et al., 2009). In addition, PCR markers have been developed for the $pun1^2$ allele (Stewart et al., 2007) and the $pun1^3$ allele (Stellari et al., 2010). Still, there remains a need for a robust suite of

molecular markers that can detect the functional nucleotide polymorphisms of the three *Pun1* alleles, thus eliminating the possibility of recombination between marker and gene, and are adaptable to a variety of genotyping platforms. Additional requirements in designing markers for modern high-throughput assays are shorter amplicon lengths and eliminating the need for restriction digests or other manual steps. The development of such a suite of markers has been hindered by the presence of the pseudogene *catf2*. *catf2* is paralogous to *Pun1*, with a 82% sequence identity (Lang et al., 2006). PCR primers designed from the *Pun1* sequence tend to also amplify from *catf2* and confound marker results. In this study, we created an effective and easy to use marker set that reliably detects the $pun1^1$ and $pun1^2$ allele functional polymorphisms and the means to transfer these markers to other genotyping platforms.

Methods, Results, and Discussion

To reliably differentiate the $pun1^1$ allele (GenBank # AY81903) from the other *Pun1* alleles (ex. GenBank # AY819027), we designed a 3-primer PCR marker (Figure 2.1a). By positioning the first forward primer within the deleted region, amplification only occurs in the absence of the $pun1^1$ mutation. A second forward primer spans the deleted region, so it can only bind and amplify when the $pun1^1$ deletion is present. The resulting PCR fragments can easily be distinguished on a 1% agarose gel.

A second molecular marker was designed to differentiate the $pun1^2$ allele (GenBank # EF104910) from the other *Pun1* alleles. This marker detects the 4 bp deletion that defines the $pun1^2$ allele. A primer pair was designed in exon 1 that flanks

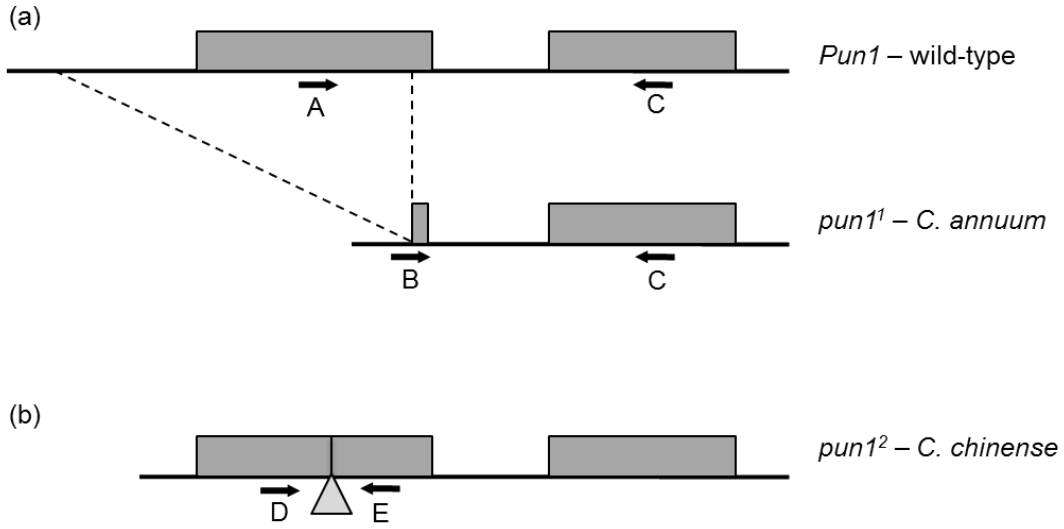


Figure 2.1 Allele-specific marker design. Exons are represented by boxes. **(a)** The *pun1*^{1} marker detects the large deletion in *pun1*^{1}, depicted by the dashed lines. Primers A and C amplify a larger fragment in the wild-type *Pun1* allele. A and C cannot amplify in the *pun1*^{1} allele because A is located in the deleted region. Primer B spans the deleted region and so can only bind to the *pun1*^{1} allele. It combines with Primer C to amplify a smaller fragment. **(b)** The *pun1*^{2} marker detects the 4 bp deletion in the *pun1*^{2} allele, depicted by the triangle

the deleted region (Figure 2.1b). The amplified fragments differ in size by 4 bp, with the *pun1*^{2} allele producing a 392 bp fragment and the *Pun1* and *pun1*^{3} alleles producing a 396 bp fragment. No fragment is amplified from the *pun1*^{1} allele because its large deletion includes the amplified region. While this 4bp polymorphism can be detected by PAGE with silver staining, the use of a fluorescently labeled forward primer allows this process to be automated using fragment analysis systems. The key innovation of these two new markers is the placement of a specific reverse PCR primer in a region that is unique to the *Pun1* locus and not present in the paralogous *catf2* pseudogene (GenBank # AB206920). In the *pun1*^{1} assay there are six mismatches between *Pun1* and *catf2* at the location of the reverse primer (Figure 2.2a). In the *pun1*^{2} assay, the reverse primer location has a 3 bp insertion in *catf2* as

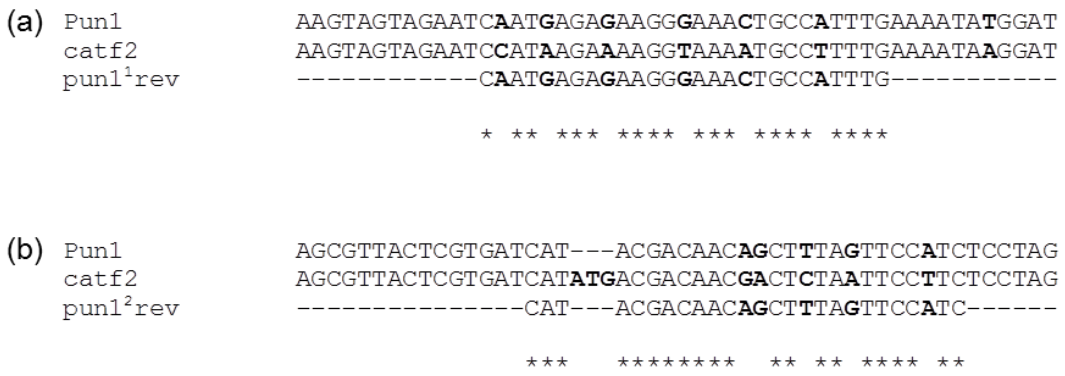


Figure 2.2 Paralog-specific primer design. Regions of the *Pun1* gene used to design PCR primers are shown aligned with the corresponding region of the *catf2* pseudogene. Primers were localized to regions that contained multiple mismatches with the *catf2* pseudogene. **(a)** The reverse primer for the *pun1*^{1}-specific marker aligned with corresponding genomic regions. **(b)** The reverse primer for the *pun1*^{2}-specific marker aligned with corresponding genomic regions. GenBank accession numbers are as follows: *Pun1* - AY819027; *pun1*^{1} - AY81903; *pun1*^{2} - EF104910; *catf2* - AB206920

well as 5 mismatches (Figure 2.2b). These primers prevent amplification from *catf2*, which would confound the results.

These new markers, along with the previously developed *pun1*^{3} marker (Stellari et al., 2010) were tested on a panel of 13 pungent and non-pungent pepper lines and representative hybrids. In addition to pepper lines that we have previously used as standard controls for the three *Pun1* alleles (Stewart et al., 2005; Stewart et al., 2007; Stellari et al., 2010), 4 *Capsicum* accessions described as non-pungent from the USDA ARS Germplasm Resources Information Network's National Plant Germplasm System (GRIN) (Griffin, GA) were included (PI 355819, Grif 9182, PI 224428, PI 653681). Three plants were grown from selfed seed in the Guterman greenhouse facility at Cornell University, Ithaca, NY, USA using standard horticultural practices. To validate the phenotype of the new non-pungent accessions, ripe fruit from each plant were tested organoleptically for pungency. Capsaicinoid content was also

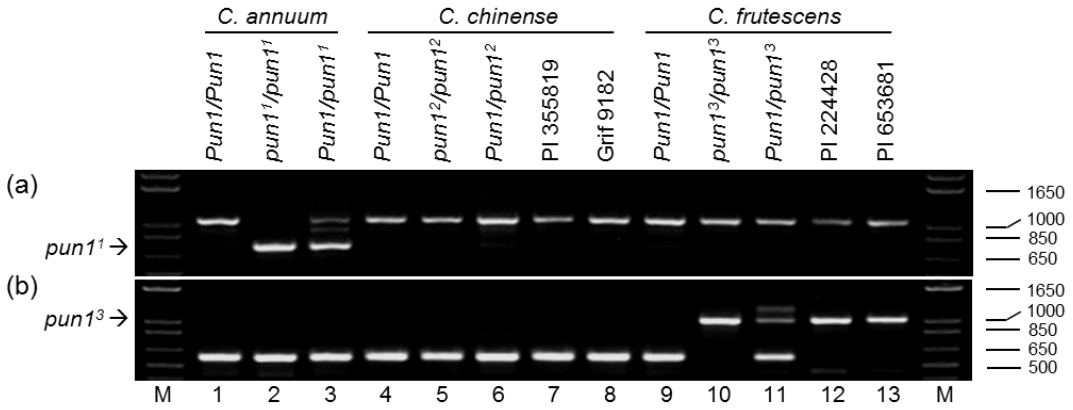
quantitatively measured in a sample of 10 mature, dried fruit from each of the three plants per accession using a capsaicinoid-detecting enzyme-linked immunosorbent assay (ELISA) kit from Beacon Analytical (Beacon Analytical, Portland, ME, USA) according to the manufacturer's protocol with modifications from Stewart et al. (2007). All four accessions were non-pungent as judged by tasting and a capsaicinoid content measured at less than 1 part per million (data not shown), which corresponds to the limits of detection of the kit and organoleptic assays (Stewart et al., 2007).

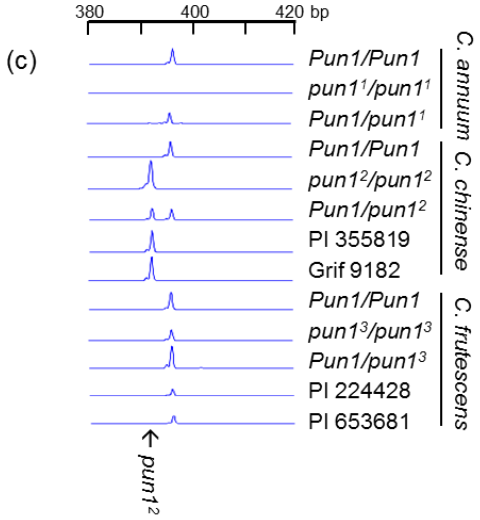
To perform the marker assays, genomic DNA from S1 plant leaf tissue was extracted using the method of Doyle and Doyle (1990). All primers sequences are listed in Table 2.1. ***pun1*^{1}-specific agarose gel marker:** PCR conditions were as follows: 2.5 µl of 10x PCR buffer, 1 µl of 2.5 mM dNTPs, 0.25 µl of 10 µM *pun1*^{1} fwd 1 primer, 0.25 µl of 10 µM *pun1*^{1} fwd 2 primer, 0.25 µl of 10 µM *pun1*^{1} rev primer, 0.25 µl of Taq polymerase, 10 µl of 20 ng/µl genomic DNA solution and H_2O to a final volume of 25 µl. PCR cycles were as described above, with an annealing temperature of 60 °C. PCR products were visualized on a 1% agarose gel stained with

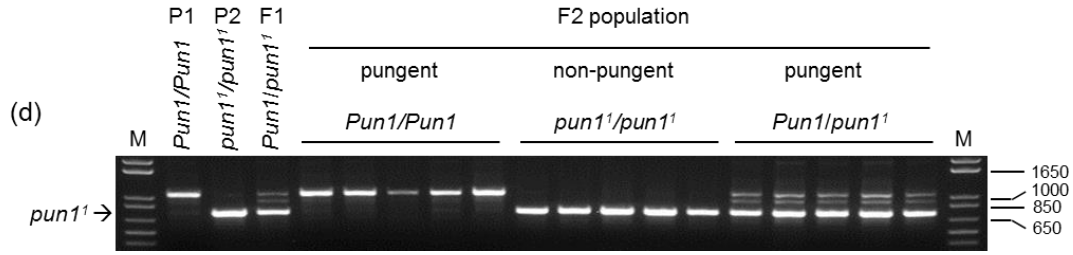
Table 2.1 PCR primers used for *Pun1* markers.

Primers	**Primer sequence (5' to 3')**
pun1^{1} fwd 1	TCCTCATGCATCTCTTGCAG
pun1^{1} fwd 2	GCTCCACGGAAAAGACTCAT
pun1^{1} rev	CAAATGGCAGTTTCCCTTCTCTCATT
pun1^{2} fwd	6-FAM TTCCCATATAGCCCACTTGC
pun1^{2} rev	GATGGAACTAAAGCTGTTGTCGTATG
pun1^{3} fwd	GTAGTTTTTCGGAAATGAAAAGTACT
pun1^{3} rev 1	CACGCCTTGCCCAGCTTTGTAATCTT
pun1^{3} rev 2	TCATGTCCATTCGGCCAAACAGTG

Figure 2.3 *Pun1* allele-specific PCR markers. Codominant PCR-based markers were developed to distinguish each of the four known *Pun1* alleles. Lanes and corresponding electropherogram: 1, 'Czech Black' (p); 2, 'Cal Wonder Orange' (np); 3, 'Czech Black' x 'Cal Wonder Orange' F_1 (p); 4, 'Habanero' (p); 5, NMCA 30036 (np); 6, 'Habanero' x 'NMCA 30036' F_1 (p); 7, PI 355819 (np); 8, Grif 9182 (np); 9, BG2814-6 (p); 10, PI 594141-np (np); 11, BG2814-6 x PI 594141-np F_1 (p); 12, PI 224428 (np); 13, PI 653681 (np); M, 1 kb+ ladder. (**a)** The *pun1*^{1}-specific marker visualized on a 1% agarose gel stained with ethidium bromide. The *pun1*^{1} allele yields a smaller band of 746 bp. A larger band of 1063 bp indicates the presence of the *Pun1*, *pun1*^{2}, or *pun1*^{3} allele. **(b)** The *pun1*^{3}-specific marker visualized on a 1% agarose gel stained with ethidium bromide. The *pun1*^{3} allele yields a larger band of 1033. A smaller band of 586 bp indicates the presence of the *Pun1*, *pun1*^{1}, or *pun1*^{2} allele. **(c)** The *pun1*^{2}-specific fragment analysis marker visualized as electropherograms. The *pun1*^{2} allele yields a smaller fragment of 392 bp. A larger fragment of 396 bp indicates the presence of the *Pun1* or *pun1*^{3} allele. No fragment is amplified from the *pun1*^{1} allele due to its large deletion. **(d)** The *pun1*^{1}-specific marker can clearly distinguish *Pun1* genotypes as shown by representative individuals of a *C. annuum* F2 population segregating for *Pun1*, demonstrating its utility in pepper breeding programs. P1, 'Czech Black;' P2, 'Cal Wonder Orange;' M, 1 kb+ ladder; p, pungent; np, non-pungent


C. annuum
C. chinense
C. frutescens
Pun1/Pun1
pun1¹/pun1¹
Pun1/pun1¹
Pun1/Pun1
pun1²/pun1²
Pun1/pun1²
PI 355819
Grif 9182
Pun1/Pun1
pun1³/pun1³
Pun1/pun1³
PI 224428
PI 653681
(a)
pun1¹
(b)
pun1³
1650
1000
850
650
500
M 1 2 3 4 5 6 7 8 9 10 11 12 13 M


380
400
420 bp
(c)
Pun1/Pun1
pun1¹/pun1¹
Pun1/pun1¹
Pun1/Pun1
pun1²/pun1²
Pun1/pun1²
PI 355819
Grif 9182
Pun1/Pun1
pun1³/pun1³
Pun1/pun1³
PI 224428
PI 653681
C. annuum
C. chinense
C. frutescens
pun1²


P1
P2
F1
F2 population
pungent
non-pungent
pungent
Pun1/Pun1
pun1¹/pun1¹
Pun1/pun1¹
(d)
M
pun1¹
1650
1000
850
650
M

ethidium bromide (Figure 2.3a). ***pun1^2*-specific fragment analysis marker:** PCR conditions were as follows: 2 µl of 10x PCR buffer, 1.6 µl of 2.5 mM dNTPs, 0.8 µl of 10 µM 6-FAM labeled *pun1^2* fwd primer, 0.8 µl of 10 µM *pun1^2* rev primer, 0.4 µl of Taq polymerase, 10 µl of 20 ng/µl genomic DNA solution and H_2O to a final volume of 20 µl. PCR cycles were as follows: 94 °C for 2 min, 41 cycles of 94 °C for 1 min, 55 °C for 1 min, 72 °C for 1 min, with a final extension at 72 °C for 5 min. PCR products were adjusted to the appropriate concentration with water and mixed with a Liz 500 size standard (Gene Scan) in formamide and denatured at 95 °C for 5 min prior to fragment analysis on a 3730xl DNA Analyzer (Applied Biosystems) (BRC, Cornell University, Ithaca, NY). Electropherograms were processed using Peak Scanner 1.0 (Applied Biosystems, 2006) (Figure 2.3c). ***pun1^3*-specific agarose gel marker:** As described in Stellari et al. (2010), PCR conditions were as follows: 2.5 µl of 10x PCR buffer, 2 µl of 2.5 mM dNTPs, 1 µl of 10 µM *pun1^3* fwd primer, 1 µl of 10 µM *pun1^3* rev 1 primer, 1 µl of 10 µM *pun1^3* rev 2 primer, 0.5 µl of Taq polymerase, 10 µl of 20 ng/µl genomic DNA solution and H_2O to a final volume of 25 µl. PCR cycles were as follows: 94 °C for 3 min, 34 cycles of 94 °C for 30 s, 55 °C for 30 s, 72 °C for 1 min 30 s, with a final extension at 72 °C for 15 min. PCR products were visualized on a 1% agarose gel stained with ethidium bromide (Figure 2.3b).

When the two newly developed markers are combined with the existing *pun1^3*-specific marker, a range of *Pun1* genotypes can be easily and accurately distinguished from each other, as demonstrated with our panel of peppers (Figure 2.3a-c). These markers have been found to be useful for phenotype prediction in segregating

populations. The *pun1*^{2} marker has been used successfully in our breeding program for several years (data not shown). The *pun1*^{1} marker was tested in a F2 population of 64 individuals derived from the highly pungent heirloom cultivar 'Czech Black' and the non-pungent 'Cal Wonder Orange' and has also been successfully used in a number of breeding projects. The marker co-segregated with pungency in the F2 population and no confounding *catf2* artifacts were observed (Figure 2.3d). The consistent identification of the *Pun1* alleles in varied pepper varieties and accessions is expected for these markers, as they are diagnostic of the functional polymorphisms that define the alleles.

The set of markers presented in this paper are unique in their ability to detect the functional nucleotide polymorphisms of the three *Pun1* alleles. This set of *Pun1* markers will aid diversity studies through the easy identification of the three known *Pun1* mutants in a wide range of germplasm. Additionally, the markers are useful for seed lot testing in seed purity programs. With a trait such as pungency in fruit, which can cause a painful sensation upon contact, it is critical to maintain the purity of non-pungent seed stocks. Finally, these markers will be highly useful in breeding programs because they provide an easy method to genotype populations and quickly identify plants with the desired pungency state.

Acknowledgments

We thank Nick Vail for providing plant care and Samantha Klasfeld and William Blessing for assisting with assays. This research was supported through start-up funds from Cornell University, the Vegetable Breeding Institute, and an

Agriculture and Food Research Initiative Competitive Grant from the USDA National Institute of Food and Agriculture (2010-8517-20551). Support for Lindsay Wyatt was provided by a Cornell University Presidential Life Sciences Fellowship and USDA National Needs Graduate Fellowship Competitive Grant No. 2008-38420-04755 from the National Institute of Food and Agriculture.

References

Bennett, D.J. and G.W. Kirby, 1968. Constitution and Biosynthesis of Capsaicin. *Journal of the Chemical Society C-Organic* 442-446.

Blum, E., K. Liu, M. Mazourek, E.Y. Yoo, M. Jahn, and I. Paran, 2002. Molecular mapping of the C locus for presence of pungency in Capsicum. *Genome* 45:702-705.

Caterina, M.J., M.A. Schumacher, M. Tominaga, T.A. Rosen, J.D. Levine, and D. Julius, 1997. The capsaicin receptor: a heat-activated ion channel in the pain pathway. *Nature* 389:816-824.

Doyle, J.J. and J.L. Doyle, 1990. Isolation of plant DNA from fresh tissue. *Focus* 12:13-15.

Lang, Y.Q., S. Yanagawa, T. Sasanuma, and T. Sasakuma, 2006. A gene encoding a putative acyl-transferase involved in pungency of Capsicum. *Breeding Science* 56:55-62.

Lee, C.J., E. Yoo, J.H. Shin, J. Lee, H.S. Hwang, and B.D. Kim, 2005. Non-pungent Capsicum contains a deletion in the capsaicinoid synthetase gene, which allows early detection of pungency with SCAR markers. *Molecules and Cells* 19:262-267.

Minamiyama, Y., S. Kinoshita, K. Inaba, and M. Inoue, 2005. Development of a cleaved amplified polymorphic sequence (CAPS) marker linked to pungency in pepper. *Plant Breeding* 124:288-291.

Stellari, G.M., M. Mazourek, and M.M. Jahn, 2010. Contrasting modes for loss of pungency between cultivated and wild species of Capsicum. *Heredity* 104:460-471.

Stewart, C., B.C. Kang, K. Liu, M. Mazourek, S.L. Moore, E.Y. Yoo, B.D. Kim, I. Paran, and M.M. Jahn, 2005. The Pun1 gene for pungency in pepper encodes a putative acyltransferase. *Plant Journal* 42:675-688.

Stewart, C., M. Mazourek, G.M. Stellari, M. O'Connell, and M. Jahn, 2007. Genetic control of pungency in C-chinense via the Pun1 locus. *Journal of Experimental Botany* 58:979-991.

Truong, H.T.H., K.-T. Kim, S. Kim, H.-R. Kim, M.-C. Cho, and J.-G. Woo, 2009. Development of Gene-Based Markers for The *Pun1* Pungency Gene in Pepper (*Capsicum* spp.) for Marker-Assisted Selection. *Horticulture, Environment, and Biotechnology* 50:358-365.

CHAPTER 3

RED HARVEST YIELD AND FRUIT CHARACTERISTICS OF *PHYTOPHTHORA CAPSICI*-RESISTANT BELL PEPPER INBRED LINES IN NEW YORK[2]

Abstract

Phytophthora capsici is an oomycete pathogen that causes disease on bell pepper (*Capsicum annuum*) and many other vegetable crops globally. Newly developed bell pepper inbred lines have been shown to be resistant to *P. capsici* and have been previously evaluated for green harvest yield. Nine *P. capsici*-resistant inbred lines and three commercial cultivars were evaluated for red harvest yield and fruit characteristics at three sites and disease resistance was evaluated through field inoculation studies. Three of the *P. capsici*-resistant lines were further evaluated as hybrid parents by measuring hybrid yield and disease resistance. *P. capsici*-resistant lines had excellent disease resistance and provided high levels of resistance to F_1 hybrids. Inbred lines had comparable yields to the commercial cultivars, but fruit were smaller in size and weight. These lines are suitable for use as inbred lines for markets where small fruit size is acceptable and have potential for use as hybrid parents.

Introduction

Bell peppers are an important vegetable crop in the United States. More than 1.8 billion pounds of bell peppers were grown in 2012 at a value of $628 million (U.S.

[2] This chapter was originally published in *HortTechnology* and is re-printed with kind permission from the American Society for Horticultural Science: Wyatt, L., Dunn, A., Falise, M., Reiners, S., Jahn, M., Smart, C., and M. Mazourek. 2013. Red Harvest Yield and Fruit Characteristics of *Phytophthora capsici*-Resistant Bell Pepper Inbred Lines in New York. *HortTechnology* 23: 356-363.

Department of Agriculture, 2013). Although most bell peppers are harvested and sold at the immature green stage, bell peppers at the mature ripe stage are also produced. Red bell peppers have excellent nutritional properties, with carotenoids, flavonoids, and vitamin C (Greenleaf, 1986).

Phytophthora blight is a disease of pepper (*Capsicum sp.*) that was first discovered in New Mexico in 1918 (Leonian, 1922). It causes severe losses in many bell pepper-producing areas around the United States and the world (Bosland and Lindsey, 1991; García-Rodríguez et al., 2010; Hwang and Kim, 1995; Oelke et al., 2003). Phytophthora blight is caused by the oomycete pathogen *Phytophthora capsici*, which can infect a wide range of vegetable crops (Crossan et al., 1954; Polach and Webster, 1972).

Management of phytophthora blight is difficult because it can spread quickly throughout a field from an initial inoculation source. *P. capsici* reproduces asexually through the formation of sporangia and zoospores (Ristaino, 1991) which are dispersed by water (Granke et al., 2009). Because the sporangia and zoospores move in water, rainfall and irrigation events have a significant impact on disease development and can spread inoculum throughout a field (Bowers and Mitchell, 1990; Café-Filho and Duniway, 1995; Ristaino, 1991). *P. capsici* reproduces sexually by forming oospores, which are capable of overwintering in the field and acting as a persistent inoculum source (Lamour and Hausbeck, 2003). In New York, phytophthora blight is a growing issue for bell pepper production (Dunn et al., 2010), with zoospore-containing irrigation water infesting previously unaffected farms and recent flooding spreading inoculum between farms. Fungicide-insensitive populations

of *P. capsici* are increasingly common and there are no fully resistant bell pepper cultivars that are commercially available; field infestations can lead to total yield loss for multiple years.

Current control strategies for phytophthora blight include the application of fungicides, cultural practices aimed at reducing the spread of inoculum, and the use of resistant bell pepper cultivars (Hausbeck and Lamour, 2004; Ristaino and Johnston, 1999). Unfortunately, control strategies are not always fully effective. Due to the prolonged use of a limited number of chemistries, fungicide-insensitive *P. capsici* isolates have been reported. Insensitivity to the fungicide mefenoxam is common and resistance to other fungicides has also been reported (Lamour and Hausbeck, 2001; Lu et al., 2010). In addition, once a field is infected, inoculum can remain for years (Lamour and Hausbeck, 2003). The most grower- and environment-friendly control strategy is the use of bell pepper cultivars that are resistant to *P. capsici*, combined with cultural practices. Unfortunately, there has been significant difficulty with breeding bell peppers resistant to *P. capsici*, with no commercially available bell peppers with full resistance as well as the desired horticultural type (Oelke et al., 2003; Ristaino and Johnston, 1999; Thabuis et al. 2004). This difficulty is due to the multigenic nature of the resistance, as well as linkage drag of negative horticultural traits linked to resistance genes (Thabuis et al., 2004). Partial resistance of some commercially available bell peppers varies in effectiveness depending on the *P. capsici* isolate used for testing (Foster and Hausbeck, 2010). These resistant bell pepper cultivars also tend to have a problem with silvering, where the bell pepper cuticle separates from the epidermis causing a silver coloration (Kline et al., 2011).

A set of nine *P. capsici*-resistant bell pepper inbred lines (hereafter referred to as "PR lines") has been developed that provides a new option to growers with phytophthora blight in the northeast United States. These PR lines displayed excellent resistance when challenged in a greenhouse assay using four Michigan *P. capsici* isolates (Foster and Hausbeck, 2010). They have also been evaluated in inoculated field studies and were found to be resistant to the crown rot phase of phytophthora blight caused by *P. capsici* (Dunn et al., 2013). In that study, the number and weight of fruit per plot of green (immature) harvested fruit of the PR lines were comparable to commercial cultivars, especially under disease pressure, and they had a low incidence of silvering (Dunn et al., 2013). In this study, we evaluated the ripe red harvest yield and fruit characteristics of the PR lines and several commercial cultivars, compared their performances in organic, conventional, and conventional drought environments, and tested their potential to provide disease resistance as hybrid parents.

Materials and Methods

All of the cultivars and lines evaluated in this study are described in Table 3.1. Nine PR lines bred at Cornell University were evaluated in this experiment. Three commercial cultivars, two of which were described by the supplier as having "intermediate resistance" to phytophthora blight, were grown as controls. 'Aristotle' could not be obtained as untreated seed, so it was omitted from the organic trial site. In 2012 only, six F_1 hybrids created by manual cross-pollinations between a PR line and a commercial cultivar were also tested. The two commercial cultivars used as

Table 3.1 Bell pepper cultivars, lines, and F_1 hybrids used in yield and disease resistance trial.

Cultivar or line	Source	
'Aristotle'	Seminis	St. Louis, MO
'Paladin'	Syngenta	Greensboro, NC
'Revolution'	Harris Moran	Modesto, CA
Pcap-NY8001-1	Cornell University	Ithaca, NY
Pcap-NY8002-3	Cornell University	Ithaca, NY
Pcap-NY8003-2	Cornell University	Ithaca, NY
Pcap-NY8006-1	Cornell University	Ithaca, NY
Pcap-NY8006-4	Cornell University	Ithaca, NY
Pcap-NY8007-1	Cornell University	Ithaca, NY
Pcap-NY8007-2	Cornell University	Ithaca, NY
Pcap-NY8007-3	Cornell University	Ithaca, NY
Pcap-NY8007-4	Cornell University	Ithaca, NY
F_1 #1 - 'Keystone Giant' x Pcap-NY8002-3	F_1 hybrid (this study)	
F_1 #2 - 'Keystone Giant' x Pcap-NY8003-2	F_1 hybrid (this study)	
F_1 #3 - 'Keystone Giant' x Pcap-NY8007-3	F_1 hybrid (this study)	
F_1 #4 - 'King of the North' x Pcap-NY8002-3	F_1 hybrid (this study)	
F_1 #5 - 'King of the North' x Pcap-NY8003-2	F_1 hybrid (this study)	
F_1 #6 - 'King of the North' x Pcap-NY8007-3	F_1 hybrid (this study)	
'Keystone Giant'	Gourmet Seed	Tatum, NM
'King of the North'	High Mowing Organic Seeds	Wolcott, VT

hybrid parents, Keystone Giant and King of the North, were also grown in 2012 as controls.

These experiments were conducted at three field sites: yield trials were conducted at the East Ithaca Research Farm of Cornell University in Ithaca, NY in 2011 and 2012 (hereafter referred to as “East Ithaca”) and the Freeville Organic Research Farm of Cornell University in Freeville, NY (hereafter referred to as

"Freeville Organic") in 2012, while disease resistance trials were conducted at the Phytophthora Blight Farm at the New York State Agricultural Experiment Station in Geneva, NY in 2011 and 2012. The Phytophthora Blight Farm is used annually for field-based inoculation studies because it is quarantined from nearby farms, but does not have an overwintering population of *P. capsici*. The East Ithaca Research Farm has an Arkport sandy loam soil type, the Freeville Organic Research Farm has a Rhinebeck clay soil type, and the Phytophthora Blight Farm has Odessa silt loam soil. At all sites, raised beds were constructed which were 4 inches high, 30 inches wide, and 7 ft between centers. Beds were covered with 1.25-mil black embossed plastic mulch (Belle Terre Irrigation, Sodus, NY) and drip tape (emitter spacing of 12 inches and flow rate of 0.45 gal/min per 100 ft; Aqua-Traxx, Toro Co., Bloomington, MN) was used.

At the East Ithaca site, 10N-8.7P-16.6K fertilizer was applied before planting at a rate of 300 lb/acre (Arrow; Royster-Clark, Princeton, NC). At the Freeville Organic site, compost (2.4N-1.25P-0.9K) was applied to the field at a rate of 10 tons/acre before planting. At the Phytophthora Blight Farm, 300 lb/acre 10N-4.4P-8.3K fertilizer (Phelps Supply, Phelps, NY) was applied under the mulch at the time beds were built.

In 2011, total monthly rainfall at the East Ithaca site was 2.59, 1.99, and 4.63 inches for June, July, and August, respectively, and at the Phytophthora Blight Farm monthly rainfall was 2.34, 0.72, and 2.62 inches for June, July, and August, respectively. In 2012 at the East Ithaca and Freeville Organic sites, total monthly rainfall was 1.84, 1.59, and 3.58 inches for June, July, and August, respectively and

monthly rainfall was 2.59, 2.80, and 2.26 inches for June, July, and August, respectively at the Phytophthora Blight Farm. At East Ithaca in 2012, a low rate of supplementary irrigation (0.44 acre-inch) was provided using drip tape every other week to create significant drought stress for the plants, simulating conditions for growers with no supplemental irrigation. Drought stress was confirmed through visual observation of the plants. At the Freeville Organic site, the Phytophthora Blight Farm, and East Ithaca in 2011, irrigation was provided to prevent drought stress.

Bell pepper seedlings were grown in a greenhouse using natural and supplemental light. Six weeks after seeding, transplant seedlings were transferred to a cold frame for hardening off and 8 weeks after planting, seedlings were transplanted to the field. Transplanting dates were 3 June 2011 and 31 May 2012 at the East Ithaca site, 10 June 2011 and 15 June 2012 at the Phytophthora Blight Farm, and 30 May 2012 at the Freeville Organic site.

At the East Ithaca site, soluble fertilizer (Peters 10N-13.1P-16.6K; JR Peters, Inc., Allentown, PA) was applied at transplant, at an approximate rate of 0.083 gal/ft of row (7.8 g/gal of water). At the Freeville Organic Site, Neptune's Harvest fish emulsion (Hydrolyzed Fish 2N-4P-1K; Neptune's Harvest, Gloucester, MA) was applied immediately after transplanting at an approximate rate of 0.055 fl oz/ft of row. At the Phytophthora Blight Farm, soluble fertilizer (Peters Excel 21N-2.2P-16.6K; JR Peters, Inc., Allentown, PA) was applied at transplant at an approximate rate of 0.07 gal/ft of row (0.04 lb fertilizer/gal of water). At the Freeville Organic site, plants were fertilized once during the season by applying diluted Neptune's Harvest fish emulsion at the base of each plant at the approximate rate of 0.055 fl oz/ft of row.

Yield was measured in three environments: the conventional, irrigated field at the East Ithaca site in 2011 (hereafter referred to as "2011 Conventional"), the conventional, drought-stressed field at the East Ithaca site in 2012 (hereafter referred to as "2012 Drought-stressed"), and the organic, irrigated field in 2012 at the Freeville Organic site (hereafter referred to as "2012 Organic").

For the yield trial, plants were arranged in a three-replicate randomized complete block design. Each plot consisted of 12 plants with the middle 10 plants being harvested. Plants were spaced 18 inches apart in an offset double row. Beginning the week of 15 Aug. 2011 and the week of 6 Aug. 2012, plots were harvested weekly for 8 weeks. Mature fruit (at least 80% red or yellow) were harvested and graded into marketable and unmarketable classes. Marketable fruit had a diameter of greater than 2.5 inches and were blemish-free. Yield of both classes was measured both by number of fruit and by total weight of fruit. Yield data were recorded on a 10-plant plot basis and were added across the eight harvest dates to calculate cumulative yield, which was used for subsequent statistical analyses. In 2012, the numbers of fruit in the unmarketable class with silvering and blossom-end rot were also counted in order to confirm between-line differences initially observed in 2011.

Ten marketable fruit (if available) were arbitrarily selected from each plot and the following measurements were made on each fruit: fruit length (centimeters), fruit width at the widest point (centimeter), fruit wall thickness (millimeters), fruit weight [grams (in 2012 only)], number of lobes, and percent soluble solids, measured using a refractometer. In 2011, fruit weight was calculated as the mean weight of the 10 fruit

selected for analysis. The number of lobes per fruit was used to calculate the percent of marketable fruit with four lobes. Lines with less than seven marketable fruit for individual fruit measurements were excluded from analyses due to inadequate sample size.

At the Phytophthora Blight Farm, all trial entries were evaluated for susceptibility to *P. capsici*. The PR lines and the commercial cultivars were evaluated in 2011 and the F_1 hybrids and hybrid parents were evaluated in 2012. Plants were arranged in a randomized complete block design with four replicates in 2011 and three replicates in 2012. Each plot consisted of 10 plants and plants were spaced 18 inches apart in a single row. Inoculation and data collection procedures were as described in Dunn et al. (2013). Briefly, a New York isolate of *P. capsici* (NY 066-4, collected from a bell pepper on a commercial vegetable farm in central New York in 2006; Dunn et al., 2010) was used to inoculate bell peppers 2, 4, and 6 weeks after transplanting in 2011. Approximately 5 mL of a 4×10^4-zoospores/mL suspension were applied at the crown of each plant using a 1.5-gal hand-pump sprayer for the first inoculation and 5 mL of a 1×10^5- zoospores/mL suspension were applied higher up the stem of the plants for the later inoculations. In 2012, bell peppers were inoculated 5 days and 3 weeks after transplanting. Approximately 5 mL of a 1×10^5-zoospores/mL suspension were applied at the crown of each plant using a 1.5-gal hand-pump sprayer for the first inoculation and 5 mL of a 1×10^5- zoospores/mL suspension were applied higher up the stem of the plants for the later inoculation. Inoculation procedure was adapted each year based on weather conditions to achieve high levels of disease pressure. In both years, starting approximately 5 d after the first

inoculation, the proportion of plants showing symptoms of phytophthora blight (either wilting of at least one branch or plant death) was recorded on a per-plot basis. These ratings continued approximately twice weekly for 3 months. Disease resistance was quantified by calculating the area under the disease progress curve (AUDPC). Isolations were taken from a representative sample of diseased plants to confirm that the causal organism was *P. capsici*.

JMP statistical software (version 9.0.0; SAS Institute, Cary, NC) was used for all analyses. Separate analyses of variance (ANOVA) were conducted for all yield, fruit characteristics, and disease resistance data for each year and site in order to determine the effects of line on the assorted dependent variables. Following all significant ANOVAs ($P < 0.05$), Tukey's honestly significant difference (HSD) test (at alpha = 0.05) was used to separate means, except when numbers of entries were unequal, in which case Tukey-Kramer's HSD test (at alpha = 0.05) was used.

Results

The Cornell University PR lines had excellent resistance to phytophthora blight. Seven of the breeding lines displayed no disease symptoms after being inoculated three times with a New York isolate of *P. capsici* on the Phytophthora Blight Farm in 2011 and the remaining two had minimal disease symptoms (Figure 3.1). All of the PR lines were significantly more resistant to the crown rot phase of phytophthora blight than two commonly available commercial cultivars, Aristotle, which although not advertised as resistant to phytophthora blight, is generally regarded as moderately resistant (Keinath, 2012; Li, 2012; Louws et al., 2008), and

Revolution, which is described as having intermediate resistance to phytophthora blight (Figure 3.1). While the PR lines were not significantly more resistant than 'Paladin,' a commonly available cultivar that is highly resistant to phytophthora blight, they all had a lower mean AUDPC than 'Paladin,' showing that they had relatively less disease symptoms.

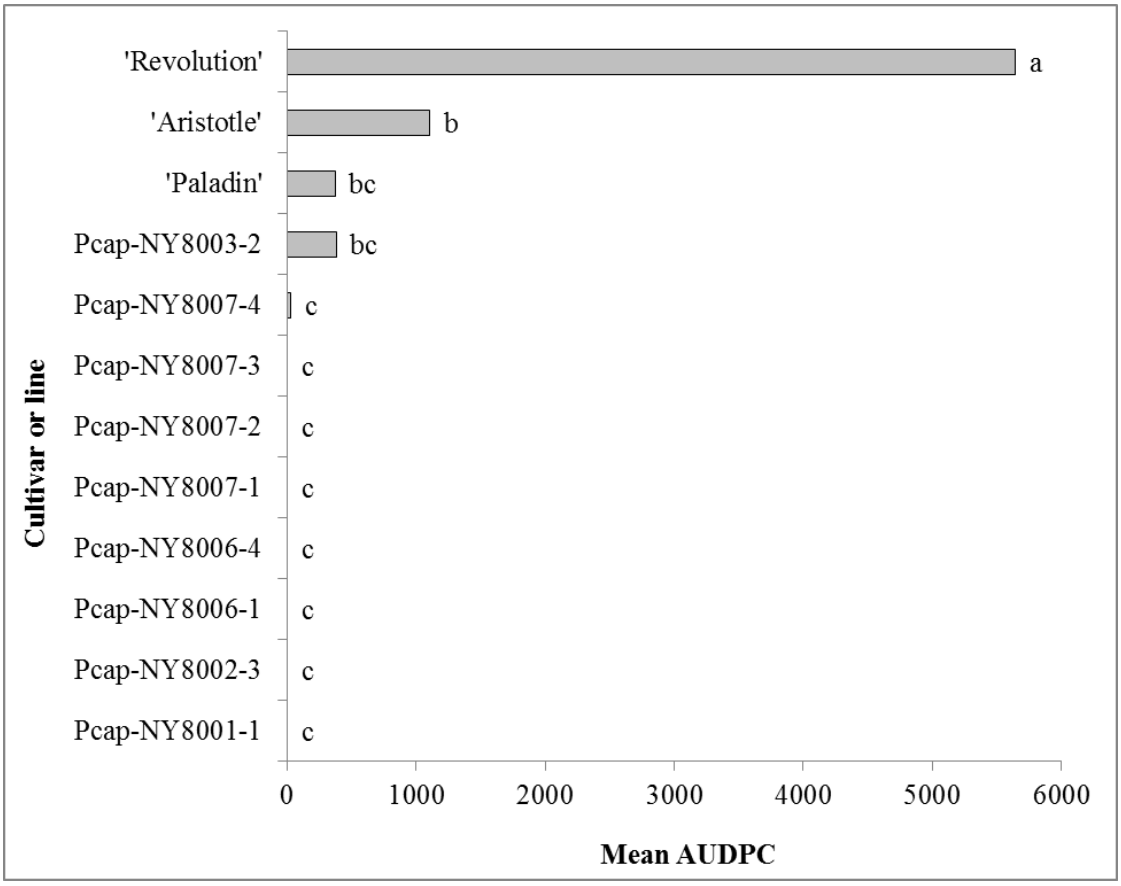


Figure 3.1 Resistance of bell pepper lines to phytophthora blight. Resistance is measured as area under the disease progress curve (AUDPC), calculated by biweekly disease incidence ratings of inoculated plants on the Phytophthora Blight Farm in 2011. The Phytophthora Blight Farm is a research farm in Geneva, NY used annually for field-based inoculation studies because it is quarantined from nearby farms, but does not have an overwintering population of *P. capsici*. Means followed by the same letter are not significantly different at $P < 0.05$ using Tukey's honestly significant difference test.

Overall marketable yield, measured as number of red or yellow fruit per 10-plant plot, varied between sites (Figure 3.2). Trial entries had significantly different fruit number per plot within each of the three sites. At the 2011 Conventional site, Pcap-NY8003-2, Pcap-NY8002-3 and Pcap-NY8006-1 had the highest yields, yielding significantly more than 'Revolution' (Figure 3.2). Pcap-NY8006-1 also had the highest yield at the 2012 Drought-stressed site and yielded significantly more than 'Revolution' and 'Paladin' (Figure 3.2). At the 2012 Organic site, 'Revolution' and 'Paladin' had the highest yields of all of the trial entries, while Pcap-NY8003-2 had the highest yield of the nine PR lines (Figure 3.2). Overall, three of the PR lines, Pcap-NY8001-1, Pcap-NY8003-2, and Pcap-NY8006-1, had numbers of marketable fruit comparable to the commercial cultivars across all sites.

The PR lines tended to have a high percentage of marketable fruit at the 2011 Conventional site, with Pcap-NY8003-2 and Pcap-NY8002-3having a significantly higher percentage of marketable fruit than the three commercial cultivars (Table 3.2). A similar trend was observed at the 2012 Drought-stressed site, where Pcap-NY8006-1 and Pcap-NY8006-4 had a significantly higher percentage of marketable fruit than 'Paladin' or 'Revolution.' At the 2012 Organic site, 'Revolution' had the highest percentage of marketable fruit, but six of the PR lines were not significantly lower. Based on observations of differences between lines in 2011, in 2012 the percent of fruit categorized as unmarketable because of either sunscald or blossom-end rot was also recorded at both sites (Table 3.2). At the 2012 Drought-stressed site, the three commercial cultivars had the highest percentage of fruit with sunscald, with Aristotle having a significantly higher percentage of fruit with sunscald than any of the PR

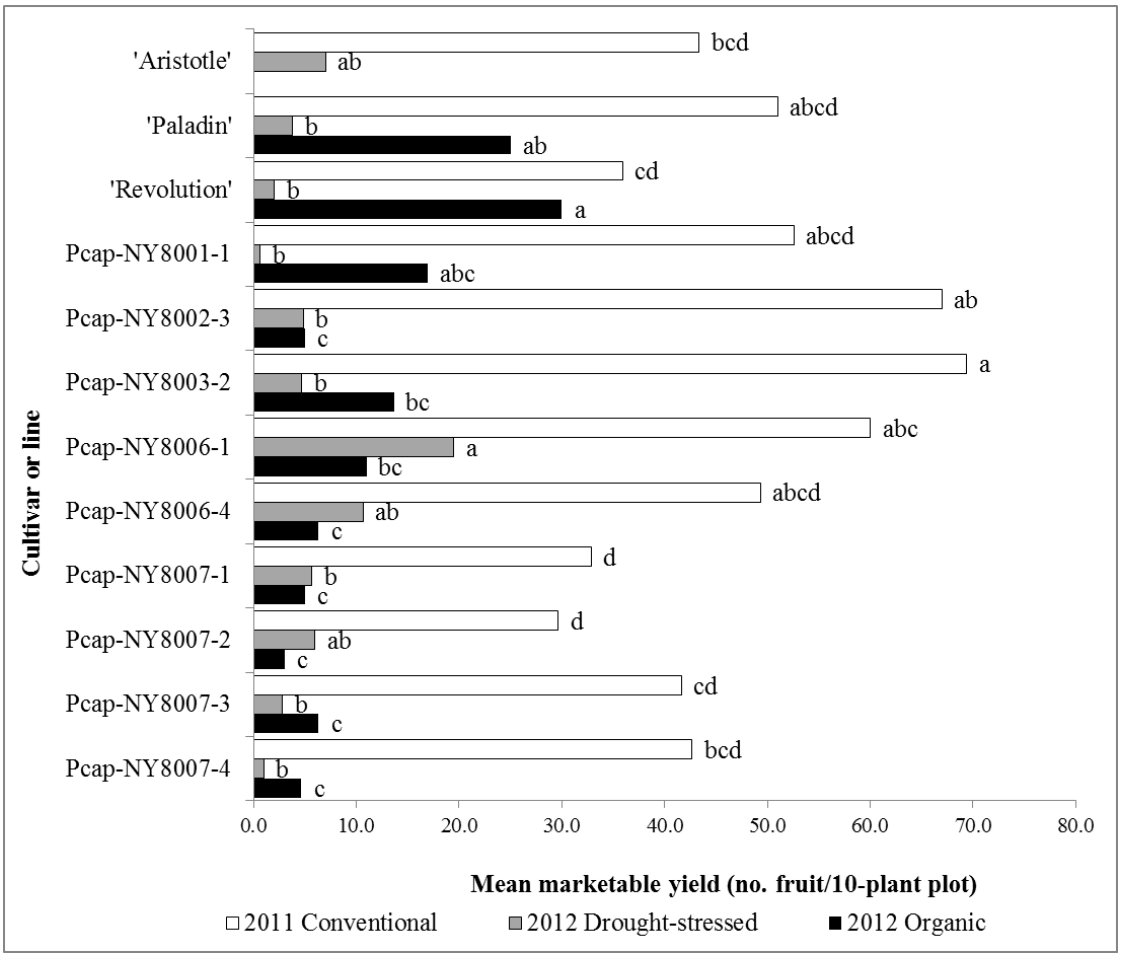


Figure 3.2 Cumulative marketable yield of red or yellow bell peppers. Yield was measured as the total number of red or yellow marketable fruit per 10-plant plot (10-plant plot = 20.625 ft^2 [1 ft^2 = 0.0929 m^2]; 1 fruit/10-plant plot = 0.0485 fruit/ft^2 = 0.5219 fruit/m^2) summed across eight weekly harvests. The three treatments were "2011 Conventional," a conventional, irrigated field at the East Ithaca, NY site in 2011, "2012 Drought-stressed," a conventional, drought-stressed field at the East Ithaca, NY site in 2012, and "2012 Organic," an organic, irrigated field in 2012 at the Freeville, NY Organic site. 'Aristotle' was omitted from 2012 Organic because untreated seed could not be obtained. Means for a given site followed by the same letter are not significantly different at $P < 0.05$ using Tukey's honestly significant difference test.

lines. At the 2012 Organic site, a similar trend was observed, but differences between lines were not significant. Pcap-NY8006-1 had the lowest occurrence of blossom-end rot at the 2012 Drought-stressed site, with a significantly lower percentage of fruit with blossom-end rot than 'Aristotle' or 'Revolution.'

Fruit size varied significantly between lines at both the 2011 Conventional and the 2012 Organic sites (Table 3.3). The commercial cultivars had larger fruit than the PR lines. 'Revolution' had the largest fruit measured both as fruit weight and fruit length x width and had a significantly larger length x width than all of the PR lines in both environments. Differences in fruit size were not significant at the 2012 Drought-stressed site because of low sample number, but the trends were similar across all sites (Table 3.3). However, all lines displayed an increased size at the 2012 Organic site compared to the other two sites. The differences between lines in fruit wall thickness were less pronounced (Table 3.3). The three commercial cultivars had relatively thick fruit walls, as did Pcap-NY8001-1, Pcap-NY-8002-3, and Pcap-NY8003-2.

The percent of fruit with four lobes did not differ significantly between lines in any of the environments (data not shown). In contrast, percent soluble solids was significantly different between lines in all three environments (Table 3.4). At the 2011 Conventional site, Pcap-NY8007-1, Pcap-NY8007-2, and Pcap-NY8007-4 had the highest percent soluble solids and were significantly higher than the three commercial cultivars. Similarly, Pcap-NY8007-1 had the highest percent soluble solids at the 2012 Organic site and had a significantly higher percent soluble solids than 'Paladin' and 'Revolution.' In contrast, the 2012 Drought-stressed site displayed a different trend, with 'Aristotle' having the highest percent soluble solids, although it was only significantly higher than one of the PR lines.

F_1 hybrids were made using three PR lines (Pcap-NY8002-3, Pcap-NY8003-2, and Pcap-NY8007-3) as paternal parents and two commercial open-pollinated cultivars (Keystone Giant and King of the North) as maternal parents. In 2012, their

Table 3.2. **Percent marketable bell pepper fruit, percent fruit with sunscald, and percent fruit with blossom-end rot**. Values are expressed as mean percent of total fruit number. Marketable fruit are those with no sunscald, blossom-end rot, or other blemishes. Bell peppers were evaluated in a conventional, irrigated environment, a conventional, drought-stressed environment (both in Ithaca, NY), and an organic, irrigated environment (in Freeville, NY).

	Mean sunscald (%)		Mean blossom-end rot (%)		Mean marketable fruit (%)		
Cultivar or line	2012 Drought-stressed[z]	2012 Organic[z]	2012 Drought-stressed	2012 Organic	2011 Conventional[z]	2012 Drought-stressed	2012 Organic
'Aristotle'[y]	10.6 a[x]	-	43.0 ab	-	46.4 cd	10.3 bc	-
'Paladin'	7.6 abc	5.3	28.9 abc	0.0	44.4 cd	4.8 c	44.5 abcd
'Revolution'	8.0 ab	2.4	45.4 ab	1.8	41.4 cd	3.2 c	69.0 a
Pcap-NY8001-1	0.0 d	0.0	48.0 a	1.3	59.3 abc	7.1 c	54.2 abc
Pcap-NY8002-3	2.8 bcd	0.0	29.4 abc	0.0	70.9 ab	18.1 abc	61.7 ab
Pcap-NY8003-2	0.0 d	0.0	48.8 a	0.0	74.4 a	23.8 abc	67.9 a
Pcap-NY8006-1	1.3 bcd	0.0	3.6 c	0.0	56.0 abcd	38.7 a	49.0 abcd
Pcap-NY8006-4	0.8 cd	0.0	12.8 bc	0.0	56.3 abcd	36.7 ab	25.2 cd
Pcap-NY8007-1	2.0 bcd	5.6	17.6 abc	2.8	40.8 cd	22.1 abc	36.8 abcd
Pcap-NY8007-2	1.0 cd	1.9	24.6 abc	0.0	35.8 d	20.5 abc	16.1 d
Pcap-NY8007-3	2.0 bcd	0.0	27.3 abc	0.0	47.5 cd	12.3 abc	26.1 bcd
Pcap-NY8007-4	1.4 bcd	0.0	14.4 abc	0.0	50.6 bcd	6.1 c	36.2 abcd
P-value[w]	<0.0001	NS	0.0007	NS	<0.0001	0.0005	0.0002

[z] The three treatments were "2011 Conventional," a conventional, irrigated field at the East Ithaca, NY site in 2011, "2012 Drought-stressed," a conventional, drought-stressed field at the East Ithaca, NY site in 2012, and "2012 Organic," an organic, irrigated field in 2012 at the Freeville, NY Organic site.

[y] 'Aristotle' was omitted from the 2012 Organic site because untreated seed could not be obtained.

[x] Within a column, means followed by the same letter are not significantly different at $P < 0.05$ using Tukey's honestly significant difference test.

[w] *P*-value for analysis of variance testing for significant differences among lines. NS indicates lines were not significantly different at $P < 0.05$.

resistance to phytophthora blight was evaluated at the Phytophthora Blight Farm. The F_1 hybrids were highly tolerant to phytophthora blight (Figure 3.3), with five out of the six F_1 hybrids not significantly different from the Pcap-resistant parents. All of the hybrids were significantly more resistant than the commercial cultivar parents, which were extremely susceptible as displayed by their high AUDPCs.

Hybrid yields varied by hybrid combination (Figure 3.4). One hybrid, 'King of the North' x Pcap-NY8003-2, had the highest yield, which was significantly more than three of the other hybrids. One of the three 'Keystone Giant' hybrids had a yield significantly greater than 'Keystone Giant' and its PR line parent. None of the 'King of the North' hybrids had a yield significantly greater than 'King of the North,' but all were significantly greater than the PR line parents (Figure 3.4).

Discussion

There is a growing need for northeastern-adapted *P. capsici*-resistant bell pepper cultivars as phytophthora blight becomes more widespread (Dunn et al., 2010). Nine new bell pepper inbred lines with excellent resistance to *P. capsici* have been developed. Four of these lines were previously evaluated for disease resistance and green bell pepper yield (Dunn et al., 2013). In this study, the disease screening data from the full set of PR lines is reported. The PR lines displayed excellent disease resistance to the New York isolate of the pathogen used for screening, which was consistent with results from a study conducted with four Michigan *P. capsici* isolates (Foster and Hausbeck, 2010). In all trials, the PR lines consistently displayed high resistance compared to 'Paladin,' a commonly-available *P. capsici*-resistant cultivar,

Table 3.3 Size of marketable red or yellow bell pepper fruit, measured as fruit weight, fruit length x width, and fruit wall thickness. Bell peppers were evaluated in a conventional, irrigated environment, a conventional, drought-stressed environment (both in Ithaca, NY), and an organic, irrigated environment (in Freeville, NY).

	Mean fruit weight (g)[z]			Mean fruit length x width (cm^2)[y]			Mean fruit wall thickness (mm)[y]		
Cultivar or line	2011 Conventional[x]	2012 Drought-stressed[w]	2012 Organic[x]	2011 Conventional	2012 Drought-stressed	2012 Organic	2011 Conventional	2012 Drought-stressed	2012 Organic
'Aristotle'[v]	196 ab[u]	166.9	-	77.0 ab	70.2	-	6.3 a	6.2 a	-
'Paladin'	148 bc	-	238.1 b	67.8 c	-	106.2 ab	5.7 b	-	6.1 abc
'Revolution'	211 a	-	318.5 a	82.3 a	-	112.8 a	5.9 ab	-	6.8 a
Pcap-NY8001-1	163 abc	-	236.7 b	66.0 c	-	89.8 c	5.9 ab	-	6.2 abc
Pcap-NY8002-3	164 abc	167.5	223.6 b	69.6 bc	65.3	90.7 bc	5.5 bc	5.9 ab	6.4 ab
Pcap-NY8003-2	155 bc	134.5	228.0 b	63.1 c	55.2	93.4 bc	6.0 ab	5.8 ab	6.2 abc
Pcap-NY8006-1	141 c	125.2	205.6 b	63.0 c	54.6	79.9 c	5.2 cd	6.0 a	5.9 bcd
Pcap-NY8006-4	147 bc	148.8	215.2 b	61.2 c	64.9	84.5 c	5.7 b	5.8 ab	6.1 abcd
Pcap-NY8007-1	146 bc	116.7	187.2 b	63.2 c	53.6	79.1 c	5.2 cd	4.8 b	5.4 cd
Pcap-NY8007-2	136 c	156.4	214.7 b	62.7 c	64.2	91.1 bc	5.0 d	5.4 ab	5.7 abcd
Pcap-NY8007-3	135 c	147.5	189.0 b	62.3 c	62.3	82.2 c	5.0 cd	5.6 ab	5.3 d
Pcap-NY8007-4	144 bc	-	194.5 b	61.7 c	-	80.1 c	5.1 cd	-	5.6 bcd
P-value[t]	0.0003	NS	<0.0001	<0.0001	NS	<0.0001	<0.0001	0.0219	<0.0001

[z] 2011 mean fruit weight was calculated from the average weight of 10 arbitrarily selected marketable bell peppers, while 2012 mean fruit weights were calculated from the individual weights of 10 arbitrarily selected marketable bell peppers. 1 g = 0.0353 oz.

[y] 1 cm = 0.3937 inch, 1 mm = 0.0394 inch

[x] "2011 Conventional" treatment is a conventional, irrigated field at the East Ithaca, NY site in 2011 and "2012 Organic" treatment is an organic, irrigated field in 2012 at the Freeville, NY Organic site.

[w] "2012 Drought-stressed" treatment is a conventional, drought-stressed field at the East Ithaca, NY site in 2012. Lines from the 2012 Drought-stressed site with six or fewer analyzed fruit were omitted from the analysis.

[v] 'Aristotle' was omitted from 2012 Organic because untreated seed could not be obtained.

[u] Within a column, means followed by the same letter are not significantly different at $P < 0.05$ using Tukey-Kramer's honestly significant difference test.

[t] *P*-value for analysis of variance testing for significant differences among lines. NS indicates lines were not significantly different at $P < 0.05$.

Table 3.4 Mean percent soluble solids of red and yellow marketable bell peppers. Bell peppers were evaluated in a conventional, irrigated environment, a conventional, drought-stressed environment (both in Ithaca, NY), and an organic, irrigated environment (in Freeville, NY).

	Mean soluble solids (%)		
Cultivar or line	2011 Conventional[z]	2012 Drought-stressed[y]	2012 Organic[z]
'Aristotle'[x]	6.7 bc[w]	9.4 a	-
'Paladin'	6.8 bc	-	7.3 bcd
'Revolution'	6.3 c	-	7.3 bcd
Pcap-NY8001-1	7.1 ab	-	7.5 bcd
Pcap-NY8002-3	6.2 c	7.6 b	6.9 d
Pcap-NY8003-2	6.8 bc	8.7 ab	7.1 cd
Pcap-NY8006-1	7.2 ab	8.3 ab	7.4 bcd
Pcap-NY8006-4	7.1 ab	8.4 ab	7.3 bcd
Pcap-NY8007-1	7.6 a	8.9 ab	8.4 a
Pcap-NY8007-2	7.7 a	8.2 ab	8.0 abc
Pcap-NY8007-3	7.3 ab	8.5 ab	7.6 abcd
Pcap-NY8007-4	7.4 a	-	7.9 ab
P-value[v]	<0.0001	0.0165	<0.0001

[z] "2011 Conventional" treatment is a conventional, irrigated field at the East Ithaca, NY site in 2011 and "2012 Organic" treatment is an organic, irrigated field in 2012 at the Freeville, NY Organic site.
[y] "2012 Drought-stressed" treatment is a conventional, drought-stressed field at the East Ithaca, NY site in 2012. Lines from the 2012 Drought-stressed site with six or fewer analyzed fruit were omitted from the analysis.
[x] 'Aristotle' was omitted from 2012 Organic because untreated seed could not be obtained.
[w] Within a column, means followed by the same letter are not significantly different at $P < 0.05$ using Tukey-Kramer's honestly significant difference test.
[v] *P*-value for analysis of variance testing for significant differences among lines. NS indicates lines were not significantly different at $P < 0.05$.

indicating that the PR lines are a valuable addition to the available *P.capsici*-resistant material.

To fully characterize the PR lines, we conducted a three-site ripe harvest yield trial comparing the nine PR lines to three commercial cultivars. The yield data from this trial were consistent with the green harvest data previously reported (Dunn et al., 2013). A subset of the PR lines had numbers of fruit per plot and percentage of

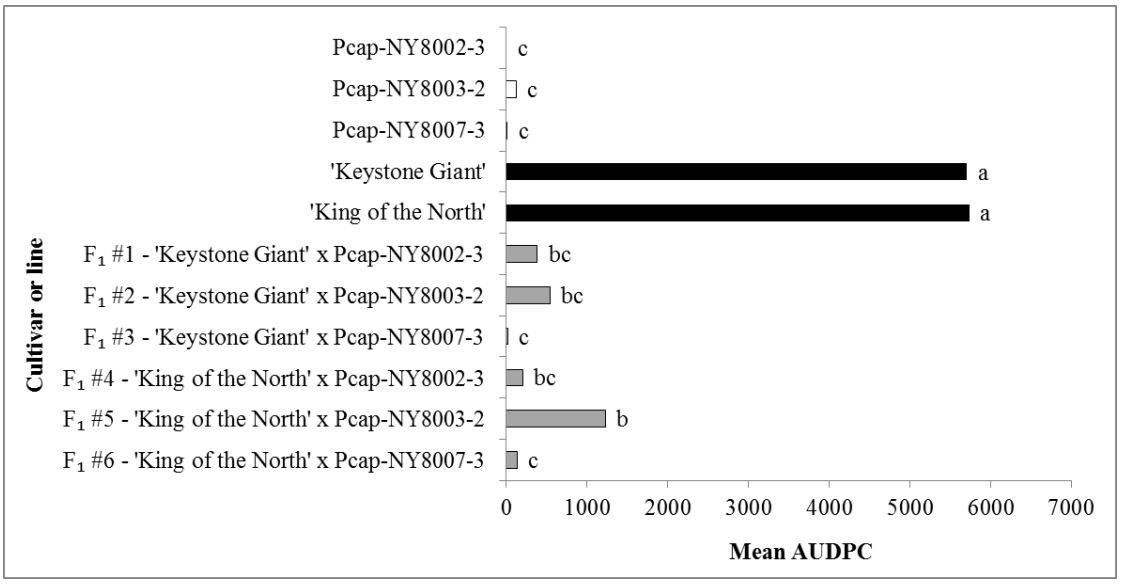


Figure 3.3 Resistance of bell pepper test hybrids to phytophthora blight. Resistance is measured as area under the disease progress curve (AUDPC), calculated by biweekly disease incidence ratings of inoculated plants on the Phytophthora Blight Farm in 2012. The Phytophthora Blight Farm is a research farm in Geneva, NY used annually for field-based inoculation studies because it is quarantined from nearby farms, but does not have an overwintering population of *P. capsici*. White bars indicate Cornell University inbred line parents, black bars indicate commercial parents, and gray bars indicate F_1 hybrids. Means followed by the same letter are not significantly different at $P < 0.05$ using Tukey's honestly significant difference test.

marketable fruit similar to the commercial cultivars, but tended to have smaller fruit, reducing yields on a weight basis.

Yields varied substantially between the three sites, which was expected based on the different environmental variables at each site. The yields were highest at the 2011 Conventional site, which is the environment for which the lines were developed. Yields were lower at the 2012 Organic site, which is likely due to differences between sites in weather, soil properties, and form of nutrition applied. It is common for yields to be lower when lines bred for conventional conditions are grown in organic conditions (Murphy et al., 2007). This indicates a need for further breeding and

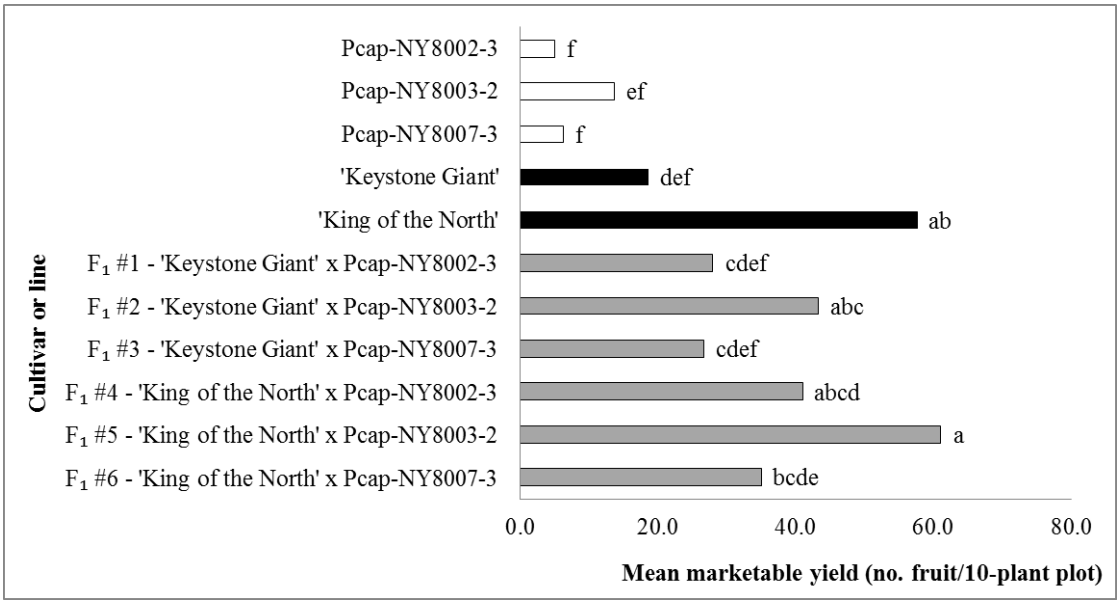


Figure 3.4 Cumulative marketable yield of bell pepper test hybrids at the 2012 Organic site in Freeville, NY. Yield was measured as the total number of red or yellow marketable fruit per 10-plant plot (10-plant plot = 20.625 ft^2 [1 ft^2 = 0.0929 m^2]; 1 fruit/10-plant plot = 0.0485 fruit/ft^2 = 0.5219 fruit/m^2) summed across eight weekly harvests. White bars indicate Cornell University inbred line parents, black bars indicate commercial parents, and gray bars indicate F_1 hybrids. Means followed by the same letter are not significantly different at $P < 0.05$ using Tukey's honestly significant difference test.

selection to adapt the PR lines for organic systems. Yields were lowest at the 2012 Drought-stressed site, which was largely due to an increase in fruit loss through blossom-end rot and other fruit defects. Blossom-end rot is often an issue when bell peppers are grown under stress conditions, including water stress (Silber et al., 2005). Notably, Pcap-NY8006-1 had much less blossom-end rot than the commercial cultivars, suggesting that it could be a good option for areas in which drought stress is common.

The fruit size measurements reported in this study were consistent with those reported in the green harvest yield trial (Dunn et al., 2013), with the PR lines having

smaller fruit than the commercial cultivars. The acceptability of these smaller fruit will vary by market and by the importance growers place on phytophthora blight resistance versus fruit size. In general, fruit size was similar at the two conventional sites, demonstrating that the drought stress reduced the number of fruit per plant, but not the size of the fruit. In contrast, fruit tended to be much bigger at the 2012 Organic site, which was likely related to the lower number of fruit per plant at the site.

The test hybrids evaluated in 2012 demonstrated the potential of using the PR lines as hybrid parents. The majority of commercial bell pepper cultivars are F_1 hybrids. Hybrids provide uniformity and the opportunity to combine different disease resistance traits possessed by each parent. The disease screening results indicated that F_1 hybrids with one of the PR lines as a parent had excellent resistance to phytophthora blight, even when combined with a highly susceptible commercial cultivar as the second parent. This confirmed previous observations that suggested the resistance of the PR lines is predominantly dominant (M. Mazourek, unpublished). The yield evaluation also demonstrated the potential of hybrid combinations to increase yield and improve adaptation, as yield was greater in the hybrids than in the PR lines and one of the commercial parents. For example, hybrids with the organic-adapted 'King of the North' as the commercial parent performed well in organic conditions.

In conclusion, this ripe harvest trial both confirmed previous green harvest results and further characterized the PR lines. The PR lines have excellent resistance to a New York isolate of phytophthora blight both as inbred lines and as F_1 progenitors. Some of the PR lines, such as Pcap-NY8003-2 and Pcap-NY8006-1,

consistently yield numbers of fruit per plot similar to commercial cultivars with low incidence of sunscald and drought-related blossom-end rot. Areas to target in future breeding efforts include increasing fruit size, earliness of fruit set, and speed of green-to-red ripening. The PR lines meet a great need for new *P. capsici*-resistant bell peppers for northeast markets. They can be grown as inbreds for markets where smaller fruit size is acceptable and also have great potential to serve as hybrid parents to develop *P. capsici*-resistant hybrids for the northeast.

Acknowledgments

We thank Maryann Fink-Brodnicki, Michael Glos, Xiaoyun Gong, Yuriy Moshkovskiy, Jenny Moore, and Holly Lange for technical assistance. Support for L. Wyatt was provided by a Cornell University Presidential Life Sciences Fellowship, and USDA National Needs Graduate Fellowship Competitive Grant No. 2008-38420-04755 from the National Institute of Food and Agriculture. Support for A. Dunn was provided by a fellowship from Cornell University College of Agriculture and Life Sciences. Funding for the organic yield trial was provided by the Toward Sustainability Foundation and further funding was provided by USDA National Institute of Food and Agriculture Plant Breeding and Education AFRI competitive grant #2010-85117-20551, USDA Cooperative State Research Education & Extension Service Initiative for Future Agriculture and Food Systems competitive grant #2001-52100-11347, Hatch funds, and the California Pepper Commission/California Pepper Improvement Foundation. New York State Department of Agriculture and Markets

Specialty Crop Block Grant #C200724 provided for seed production and disease resistance trials.

References

Bosland, P.W. and D.L. Lindsey, 1991. A seedling screen for phytophthora root rot of pepper, *Capsicum annuum*. *Plant Disease* 75:1048-1050.

Bowers, J.H. and D.J. Mitchell, 1990. Effect of soil-water matric potential and periodic flooding on mortality of pepper caused by *Phytophthora capsici*. *Phytopathology* 80:1447-1450.

Café-Filho, A.C. and J.M. Duniway, 1995. Effects of furrow irrigation schedules and host genotype on phytophthora root rot of pepper. *Plant Disease* 79:39-43.

Crossan, D.F., F. A. Haasis, and D. E. Ellis, 1954. Phytophthora blight of summer squash. *Plant Disease Reporter* 38:557-559.

Dunn, A.R., M.G. Milgroom, J.C. Meitz, A. McLeod, W.E. Fry, M.T. McGrath, H.R. Dillard, and C.D. Smart, 2010. Population structure and resistance to mefenoxam of *Phytophthora capsici* in New York State. *Plant Disease* 94:1461-1468.

Dunn, A.R., L.E. Wyatt, M. Mazourek, S. Reiners, and C.D. Smart, 2013. Performance and tolerance to phytophthora blight of bell pepper varieties. *HortTechnology* 23:382-390.

Foster, J. and M. Hausbeck, 2010. Resistance of pepper to phytophthora crown, root, and fruit rot is affected by isolate virulence. *Plant Disease* 94:24-30.

Garcia-Rodriguez, M.R., E. Chiquito-Almanza, P.D. Loeza-Lara, H. Godoy-Hernandez, E.V. Pineda, J.L. Pons-Hernandez, M.M. Gonzalez-Chavira, and J.L. Anaya-Lopez, 2010. Production of ancho chili graft on Criollo de Morelos 334 for the control of *Phytophthora capsici*. *Agrociencia* 44:701-709.

Granke, L.L., S.T. Windstam, H.C. Hoch, C.D. Smart, and M.K. Hausbeck, 2009. Dispersal and movement mechanisms of *Phytophthora capsici* sporangia. *Phytopathology* 99:1258-1264.

Greenleaf, W.H, 1986. Pepper breeding, p. 67–134. In: M.J. Basset (ed.), Breeding vegetable crops. AVI Publishing, Westport, CT.

Hausbeck, M. and K. Lamour, 2004. *Phytophthora capsici* on vegetable crops: Research progress and management challenges. *Plant Disease* 88:1292-1303.

Hwang, B.K. and C.H. Kim, 1995. Phytophthora blight of pepper and its control in Korea. *Plant Disease* 79:221-227.

Keinath, A.P, 2012. Managing phytophthora diseases on squash and pepper. Clemson Coop. Ext., Charleston, SC. 15 Feb. 2013. <http://www.clemson.edu/psapublishing/pages/PLNTPATH/IL85.pdf>

Kline, W.L., C.A. Wyenandt, D.L. Ward, J.F. Sudal, and N.L. Maxwell, 2011. Evaluation of six nitrogen fertility programs on marketable yield and development of skin separation in bell pepper fruit. *HortTechnology* 21:323-328.

Lamour, K.H. and M.K. Hausbeck, 2001. The dynamics of mefenoxam insensitivity in a recombining population of *Phytophthora capsici* characterized with amplified fragment length polymorphism markers. *Phytopathology* 91:553-557.

Lamour, K.H. and M.K. Hausbeck, 2003. Effect of crop rotation on the survival of *Phytophthora capsici* in Michigan. *Plant Disease* 87:841-845.

Leonian, L., 1922. Stem and fruit blight of peppers caused by *Phytophthora capsici* sp nov. *Phytopathology* 12:401-408.

Li, Y., 2012. Phytophthora blight of pepper. Conn. Agr. Expt. Sta., New Haven, CT. 15 Feb. 2013. <http://www.ct.gov/caes/lib/caes/documents/publications/fact_sheets/plant_pathology_and_ecology/phytophthora_blight_of_pepper_03-29-12.pdf>

Louws, F.J., G.J. Holmes, and K.L. Ivors, 2008. Pepper - Phytophthora blight fact sheet. NC State Univ., Raleigh, NC. 15 Feb. 2013. <http://www.cals.ncsu.edu/plantpath/extension/fact_sheets/Pepper_-_Phytophthora_blight.htm>

Lu, X.H., S.S. Zhu, Y. Bi, X.L. Liu, and J.J.J. Hao, 2010. Baseline sensitivity and resistance-risk assessment of *Phytophthora capsici* to iprovalicarb. *Phytopathology* 100:1162-1168.

Murphy, K.M., K.G. Campbell, S.R. Lyon, and S.S. Jones, 2007. Evidence of varietal adaptation to organic farming systems. *Field Crops Research* 102:172-177.

Oelke, L., P. Bosland, and R. Steiner, 2003. Differentiation of race specific resistance to phytophthora root rot and foliar blight in *Capsicum annuum*. *Journal of the American Society for Horticultural Science* 128:213-218.

Polach, F. and R. Webster, 1972. Identification of strains and inheritance of pathogenicity in *Phytophthora capsici*. *Phytopathology* 62:20-26.

Ristaino, J.B., 1991. Influence of rainfall, drip irrigation, and inoculum density on the development of phytophthora root and crown rot epidemics and yield in bell pepper. *Phytopathology* 81:922-929.

Ristaino, J.B. and S.A. Johnston, 1999. Ecologically based approaches to management of phytophthora blight on bell pepper. *Plant Disease* 83:1080-1089.

Silber, A., M. Bruner, E. Kenig, G. Reshef, H. Zohar, I. Posalski, H. Yehezkel, D. Shmuel, S. Cohen, M. Dinar, E. Matan, I. Dinkin, Y. Cohen, L. Karni, B. Aloni, and S. Assouline, 2005. High fertigation frequency and phosphorus level: Effects on summer-grown bell pepper growth and blossom-end rot incidence. *Plant Soil* 270:135-146.

Thabuis, A., A. Palloix, B. Servin, A. Daubeze, P. Signoret, F. Hospital, and V. Lefebvre, 2004. Marker-assisted introgression of 4 *Phytophthora capsici* resistance QTL alleles into a bell pepper line: Validation of additive and epistatic effects. *Molecular Breeding* 14:9-20.

U.S. Department of Agriculture, 2013. Vegetables 2012 summary. U.S. Department of Agriculture, Washington, DC.

CHAPTER 4

AN ACORN SQUASH (*CUCURBITA PEPO* SSP. *OVIFERA*) FRUIT AND SEED TRANSCRIPTOME AS A RESOURCE FOR THE STUDY OF FRUIT TRAITS IN *CUCURBITA*[3]

Abstract

Acorn squash (*Cucurbita pepo*) is an iconic fall vegetable in the United States, known for its unique fruit shape and also prized for its culinary properties. Little is known about the metabolism that underlies the development of fruit quality attributes such as color, sweetness, texture, and nutritional qualities in acorn squash, or any other winter squash grown worldwide. To provide insight into winter squash fruit and seed development and add to the genomic resources in the *Cucurbita* genus, RNA sequencing was used to generate an acorn squash fruit and seed transcriptome from the cultivar Sweet REBA at critical points throughout fruit development. 141,838,600 high-quality paired-end Illumina reads were assembled into 55,949 unigenes. 85% of unigenes with predicted open reading frames had homology with previously identified genes and over 62% could be functionally annotated. Comparison with the watermelon and cucumber genomes provided confirmation that the unigenes are full-length and comprehensive, covering an average of 90% of the coding sequence of their homologs and 72% of the cucumber and watermelon exomes. Key candidate genes associated with carotenoid and carbohydrate metabolism were identified toward a resource for winter squash fruit quality trait dissection. This transcriptome represents a major advance in *C. pepo* genomics, providing significant new sequence information

[3] This chapter was originally published in *Horticulture Research* and is re-printed with kind permission from Nature Publishing Group: Wyatt, L., Strickler, S., Mueller, L., and M. Mazourek. 2015. An acorn squash (*Cucurbita pepo* ssp. *ovifera*) fruit and seed transcriptome as a resource for the study of fruit traits in *Cucurbita*. *Horticulture Research* 2: 1-7.

and revealing the repertoire of genes expressed throughout winter squash fruit and seed development. Future studies on the genetic basis of fruit quality and future breeding efforts will be enhanced by tools and insights developed from this resource.

Introduction

Winter squash (*Cucurbita* spp.) is an important vegetable crop known for its nutritional content and long storage life. It is eaten as a fresh market crop, processed to make frozen squash and canned pumpkin, and the seeds are consumed as a snack food and pressed for edible oil. Winter squash is an excellent source of nutrients, including carotenoids, ascorbic acid, and vitamin C (Ferriol and Picó, 2008). It is most notable as a source of carotenoids, primarily β-carotene and lutein (Azevedo-Meleiro and Rodriguez-Amaya, 2007), which are beneficial as a pro-vitamin A compound and for eye health respectively (Cazzonelli and Pogson, 2010). Additionally, squash seeds provide protein, lipids, tocopherols, and phytosterols (Idouraine et al., 1996; Murkovic et al., 2002; Yoshida et al., 2004). Seeds were likely one of the oldest food uses of this crop, as many of the wild accessions have bitter, inedible pericarps (Robinson and Decker-Walters, 1997). Winter squash is a unique source of these nutrients because it can be stored and consumed for many months past the growing season.

To increase the consumption of winter squash, culinary quality is a vital target of improvement through breeding. Fruit quality is determined by several characteristics including color, nutrient content, sweetness, flavor, and texture. Color and nutritional value are dually controlled by carotenoid content (Gajewski et al.,

2008; Itle and Kabelka, 2009), with a deeper yellow or orange color and a higher nutrient level seen as desirable. Sweetness is associated with sugar content, and higher sweetness is desirable for consumer acceptance and contributes to improved squash flavor (Cumarasamy et al., 2002). Winter squash fruit texture is highly correlated with starch content and dry matter, with higher starch content linked to a superior texture (Corrigan et al., 2001; Cumarasamy et al., 2002; Hurst et al., 2006). These quality traits are quantitative and have a complex inheritance (Irving et al., 1997; Lu and Li, 2008), yet are controlled by well-understood metabolic pathways. Understanding the genetic basis of these quality traits will allow for more insight into the breeding of squash for improved fruit quality.

Cucurbita species have a small but growing number of genomic resources that can be employed when studying these quality traits. Other members of the Cucurbitaceae family, namely cucumber, watermelon, and melon, have more extensive genomic resources, including sequenced genomes and many molecular markers (Clepet et al., 2011; Diaz et al., 2011; Garcia-Mas et al., 2012; Guo et al., 2013; Huang et al., 2009; Ren et al., 2012; Zhang et al., 2012). Cucurbita species are diploid (2n = 2x = 40) and have a genome size of approximately 500 Mb (Arumuganathan and Earle, 1991). Their small and numerous chromosomes are a remnant of a polyploidy event with persisting duplications (Weeden, 1984) that complicates their molecular breeding. Nevertheless, several genomic maps have been generated for squash (Gong et al., 2008a; Gong et al., 2008b; Zraidi et al., 2007) including that of Esteras et al. (2012), which was the first squash map to use single nucleotide polymorphism markers. Another major source of genomic information is a

C. pepo transcriptome (Blanca et al., 2011), comprised of 49,610 unigenes derived from flower, leaf, and root tissue. More than 60% of the unigenes were annotated, aiding in the potential identification of genes of interest in the squash genome. Furthermore, the transcriptome was used to identify more than 10,000 potential simple sequence repeats and single nucleotide polymorphisms that could be used to generate molecular markers. In addition, a *C. moschata* transcriptome was sequenced from leaf, stem, and shoot tissue and assembled into 62,480 unigenes (Wu et al., 2014). Sixty-eight percent of the unigenes were annotated and almost 8,000 potential simple sequence repeats were identified.

Transcriptome generation through RNA sequencing is a technology that can be used in the dissection of complex traits; fruit transcriptome analysis has been used to successfully identify transcripts involved in early fruit development of cucumber (Ando and Grumet, 2010; Ando et al., 2012). Assembled transcriptomes also provide valuable sequence resources in species lacking a sequenced genome. However, one limitation of RNA-sequencing data is that it is specific to the plant line, tissue, developmental stage, and physiological condition sequenced. For this reason, it is essential to use transcriptome data relevant to the experimental question of interest. The first *C. pepo* transcriptome (Blanca et al., 2011), assembled from summer squash root, leaf, and flower tissue, constituted a huge advance in *C. pepo* genomics. To study aspects of fruit quality, however, knowledge of fruit-specific genes is important. Fruit development and ripening is a process often associated with the expression of suites of genes at defined stages (Rohrmann et al., 2011; Zenoni et al., 2010). A

comprehensive squash fruit transcriptome must therefore sample these specific tissues and stages.

To create a resource for the study of fruit ripening and development in winter squash, in this study we sequenced and assembled the first *Cucurbita pepo* fruit and seed transcriptome from the acorn squash cultivar Sweet REBA. Acorn squash is a popular winter squash known for its unique ribbed shape. Of the diverse squash types in *C. pepo*, the best-characterized *Cucurbita* species, acorn squash is the longest-storing and has the highest fruit quality. 'Sweet REBA' in particular is an inbred line that is agriculturally successful as a cultivar, making it an ideal candidate for generating a transcriptome that can both build off of the existing *C. pepo* knowledge and be directly applicable to squash breeding. The transcriptome was generated from five different time points throughout fruit development and derived from both mesocarp and seed tissues. The raw sequence data were assembled into 55,949 unigenes, which were then functionally annotated and compared to the cucumber and watermelon genomes. In addition, putative fruit quality-related transcripts were identified.

Materials and Methods

Plant Material

'Sweet REBA,' an acorn squash inbred cultivar (M. Jahn, Cornell University), was grown in the Guterman greenhouse facility at Cornell University, Ithaca, NY using standard horticultural practices. Flowers were manually self-pollinated and three representative, randomly-selected fruits per time point were collected at 5, 10, 15, 20,

and 40 days after pollination. Time points were selected to represent a range of squash fruit and seed developmental stages, including early fruit and seed development (5 and 10 days), seed coat formation (15 and 20 days), and fruit ripening, up until the typical harvest date (40 days) (Loy, 2004). Fruit mesocarp and seed tissues were dissected, flash frozen in liquid nitrogen, and stored at -80 °C until RNA extraction.

RNA Extraction and Sequencing

RNA was extracted from individual fruit and seed samples using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA). Equal amounts of total RNA were pooled together into three independent biological samples, each consisting of RNA from both fruit and seed tissue at all five time points. Library preparation and sequencing were performed by the Genomics Core Facility, Cornell University. The sequencing library was prepared using the TruSeq RNA Sample Preparation Kit v2 (Illumina, San Diego, CA) and paired-end reads were sequenced from the three independent samples on one lane of an Illumina HiSeq 2000.

Sequence Processing and Transcriptome Assembly

Sequencing reads were filtered by removing those flagged as low quality by the Illumina software, then adapter sequences and low-quality bases were trimmed from the end of reads using the software Trimmomatic (http://www.usadellab.org/cms/index.php?page=trimmomatic). Data quality was assessed using the program FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc) and all of the sequence data were combined for transcriptome assembly using the program Trinity (Grabherr et al., 2011). The resulting unigenes were further assembled using the program

iAssembler (Zheng et al., 2011) and then screened using the program DeconSeq (Schmieder and Edwards, 2011) to remove any unigenes of bacterial, viral, or mammalian origin. DeconSeq threshold values for removal were greater than or equal to 95% sequence identity and sequence coverage.

Transcriptome Refinement and Analysis

A BLAST search (Altschul et al., 1997) (July 2012) was performed to compare unigenes to the NCBI non-redundant protein database with a significance level of 1.0E-6 using the blastx algorithm implemented by the program Blast2GO (Conesa et al., 2005). Unigene expression estimates, measured as FPKM values, were calculated with the program RSEM (Li and Dewey, 2011), which aligned the raw reads to the assembled unigenes. Trinity (Grabherr et al., 2011) was used to identify unigenes containing predicted open reading frames. These results were then applied to refine the set of unigenes. As described in the Trinity developers' recommendations, the union of all transcripts with 1) significant homology to a known protein as determined through a blastx search, 2) a predicted open reading frame, and/or 3) an expression level, measured by the FPKM value, of greater than or equal to one were combined to remove likely assembly artifacts from the set of unigenes. To reduce the unigene set to representative transcripts, BLAST (Altschul et al., 1997) was used to query the 99,839 well-supported contigs against themselves. Perl scripts were written to merge self-alignments into groups likely originating from the same gene model.

Candidate coding regions within the transcripts were detected by ESTscan (Iseli et al., 1999) with the Arabidopsis thaliana scoring matrix. The resulting subset of transcripts was screened against the SWISS-PROT and TrEMBL (Apweiler et al.,

2013) protein databases using BLASTx with a cut-off of 1.0E-20 to assign putative functions to unigenes. Predicted peptides were analyzed using InterProScan (Zdobnov and Apweiler, 2001) to identify functional protein domains and assign Gene Ontology (GO) terms. SWISS-PROT manual curation of GO terms was also utilized to assign terms (Camon et al., 2003). Transcripts were queried against the watermelon genome assembly v1 coding sequence (watermelon_v1.cds) (Guo et al., 2013) and the cucumber genome assembly v2 coding sequence (cucumber_v2.cds) (Huang et al., 2009) using BLASTn. Unigenes with a resulting e-value of less than 1.0E-20 were considered to have a significant match. The 55,949 fruit and seed transcripts were compared to the Blanca et al. (Blanca et al., 2011) root, leaf, and flower transcriptome by performing a reciprocal blast search of one against the other using BLASTn with a significance level of 1.0E-25. Bioconductor, using the methods of Sanchez, Salicru and Ocana (Sanchez et al., 2007), was used to determine if there was a statistical difference in GO term representation between the unigenes unique to each transcriptome (Carlson et al.).

Results and Discussion

Transcriptome Sequencing and Assembly

RNA was extracted from fruit mesocarp and seed tissues of self-pollinated ‘Sweet REBA’ acorn squash fruit at 5, 10, 15, 20, and 40 days after pollination (see Figure 4.1). All samples were pooled to create three biological replicates of a comprehensive fruit and seed sequencing library and paired-end reads were sequenced on one lane of an Illumina HiSeq 2000, resulting in a total of 156,540,465 pairs of 100

Figure 4.1 'Sweet REBA' acorn squash fruit at 5 developmental time points. Self-pollinated fruit were harvested at 5, 10, 15, 20, and 40 days after pollinations. Photos are of the interior and exterior of representative fruit at each time point.

base pair (bp) reads (see Table 4.1). After removing low-quality reads and trimming both adapter sequences and low-quality bases from the ends of reads, 141,838,600 paired reads remained, with an average length of 98.46 bp and an average quality score of 36.98, constituting a total of almost 28 Gbp of sequence (reads available in the NCBI Sequence Read Archive repository, run accession numbers SRR747836, SRR747920, SRR747922). All sequence data were combined and assembled using the program Trinity (Grabherr et al., 2011) to yield 161,780 contigs. These contigs were further assembled using the program iAssembler (Zheng et al., 2011), which resulted in a total of 150,044 contigs. The contigs were then screened using the program DeconSeq (Schmieder and Edwards, 2011) to remove 2,081 bacterial and mammalian sequence artifacts, leaving 147,963 unigenes with an average length of 1,311 bp.

The unigenes were filtered to remove any sequences likely resulting from possible sequencing or assembly errors by only retaining unigenes that had homology to a known protein, a predicted open reading frame, and/or a FPKM expression

Table 4.1 Summary of transcriptome sequencing data.

Sample	Total number of paired reads	Average read length	Total sequence length	Average quality score
Raw reads	156,540,465	100 bp	31,308,093,000 bp	34.78
High-quality reads	141,838,600	98.5 bp	27,929,628,643 bp	36.98

estimate value of one or more. There were 83,221 unigenes with homology to a known protein, 57,598 with a predicted open reading frame, and 52,374 with an FPKM expression estimate value of greater than or equal to one. These overlapping subsets of unigenes were combined into a set of 99,839 unigenes, which were then clustered using self-BLAST to group alternatively-spliced transcripts originating from the same gene. An identity of 99.9% was determined to be the optimal value for merging transcripts likely transcribed from the same gene without merging putative paralogous genes. The longest unigene was chosen to represent each cluster in the final transcriptome, resulting in a total of 55,949 unigenes. (This Transcriptome Shotgun Assembly project has been deposited at DDBJ/EMBL/GenBank under the accession GBZI00000000. The version described in this chapter is the first version, GBZI01000000.)

Transcriptome Analysis, Annotation, and Functional Characterization

The final transcriptome consisted of 55,949 unigenes. The unigenes ranged in length from 201 bp to 17,024 bp, with an average length of 1,315 bp, a median length of 876 bp, and a total sequence length of 73,559,618 bp. The high average and median unigene lengths suggested that we were successful in the assembly of many full-length transcripts, although 18,561 unigenes measured 500 bp or less and were

likely only partial transcripts. A BLAST search confirmed that the longest unigenes were homologous to known genes of similar length, indicating that they were likely correctly assembled. To verify that unigenes shorter than 500 bp were indeed mainly partial transcripts, the unigene coverage of matches was considered. A total of 5,921 short unigenes had significant matches to sequences in the NCBI nr database. Of these matches, the average percent of coverage of the subject sequence was 24%. Only 7.6% of these unigenes matched 50% or more of the best nr subject, suggesting that these were truncated transcripts. With the future addition of more sequence data, it is likely that many of these short unigenes will be assembled together into full-length unigenes, reducing the total number of unigenes in the transcriptome.

A total of 41,568 unigenes were predicted to have a single open reading frame and 18 unigenes were split into 2 open reading frames. This subset of unigenes was annotated based on homology to proteins found in SWISS-PROT and TrEMBL, with precedence given to SWISS-PROT annotations. The GenBank NR and TAIR10 databases were also queried, but results were not used in the final annotation. Of the four databases, the highest number of matches was found in GenBank NR (Table 4.2). Unigenes were translated into predicted proteins so that protein domain GO terms could be assigned to the predicted proteins, classifying them based on predicted involvement in biological processes, molecular functions, or cellular components (Ashburner et al., 2000). A total of 153,563 GO terms were assigned to 63% (26,049) of the predicted proteins. The number of GO terms assigned per unigene ranged from 1 to 68 for those proteins with GO annotations, with an average of 6 GO terms per unigene. These GO terms fell into three categories, with 40% assigned to biological

Table 4.2 Results of squash fruit and seed transcriptome annotation based on homology. All unigenes with an open reading frame were blasted against four databases and the number of unigenes with a significant hit are reported. Blast hits were filtered by e-value less than 1.0E-20.

Database	Number of Hits
SWISS-PROT	22,934 (55%)
TrEMBL	33,116 (80%)
TAIR10	30,590 (74%)
GenBank NR	34,794 (84%)

process, 21% assigned to molecular function, and 39% assigned to cellular component. A range of biological processes were represented, with large numbers of unigenes assigned to categories relevant to fruit and seed development, such as "seed development," "embryo development," "regulation of developmental process," "macromolecule metabolic process," "biosynthetic process," "developmental growth," and "multicellular organismal development".

Analysis of Alternative Splicing

Approximately 44% of unigenes were collapsed by self-blast while other overlap-based means of collapsing failed, suggesting that the variants may be alternative isoforms or mis-assemblies. To assess the structure of collapsed unigenes, the clusters were further analyzed; the average number of members in a cluster was 2.5 ± 6.8, ranging from 1 member to 200 members. The clusters with the largest number of constituents were predicted to be DNA glycosylase and MYB genes. Both of these classes of genes are known to undergo alternative splicing (Li et al., 2006; Murphy and Gao, 2001), further reinforcing the notion that the isoforms identified

were alternatively spliced transcripts and supporting our designation of these as such and their subsequent representation by the longest contig in each cluster.

Comparison With Cucumber and Watermelon Genomes

A total of 31,307 unigenes had significant hits to 18,381 cucumber coding sequences, representing approximately 72% of the predicted cucumber coding sequences. The average coding sequence coverage by the homologous unigene was 90% for cucumber coding sequence with hits. A total of 4,573 unigenes with matches to a cucumber coding sequence were shorter than 500 base pairs and 4,351 of these covered less than 90% of their respective match, suggesting that these unigenes are likely partial transcripts. Similarly, 33,588 unigenes significantly matched 16,903 watermelon coding sequences, representing 72% of the watermelon predicted coding sequences. The average coverage of watermelon coding sequence by squash matches was 91%. A total of 4,868 out of 5,110 unigenes shorter than 500 bp had significant matches that covered less than 90% of their respective watermelon coding sequence match. Overall, these comparisons indicated that the majority of our unigenes were fully assembled, with the exception of those shorter than 500 bp, and that the transcriptome as a whole was comprehensive.

Comparison With Root, Leaf, and Flower *C. pepo* Transcriptome

The *C. pepo* fruit transcriptome was compared with the previously published C. pepo transcriptome sequenced from root, leaf, and flower tissue (Blanca et al., 2011). The root, leaf, and flower transcriptome was assembled from sequences derived from a scallop-type and a zucchini-type squash and consisted of 49,610 unigenes with an average length of 626 bp (Blanca et al., 2011). The fruit and seed

transcriptome had a similar number of unigenes as the previous transcriptome, though it had a longer average unigene length. The greater sequencing depth of the fruit and seed transcriptome allowed by the use of Illumina sequencing technology may have enabled the assembly of more full-length transcripts.

To more closely compare the two C. pepo transcriptomes, we conducted a reciprocal BLAST search (Altschul et al., 1997) in which each transcriptome was compared to the other transcriptome using the blastn algorithm with a significance threshold of 1.0E-25 (see Table 4.3). When the root, leaf, and flower transcriptome was compared to the fruit and seed transcriptome, 96% (47,433) of the root, leaf, and flower unigenes were significantly similar to one or more of the unigenes in the fruit and seed transcriptome, with 4% not significantly similar. The matches corresponded to 22,749 fruit and seed unigenes, suggesting redundancy in the root, leaf, and flower transcriptome. From the reciprocal BLAST search, we found that 61% (33,988) of the fruit and seed unigenes were significantly similar to one or more of the unigenes in the root, leaf, and flower transcriptome, with 39% not significantly similar. These matches were to 22,812 of the root, leaf, and flower transcripts. Additionally, 19,831

Table 4.3 Comparison of fruit and seed transcriptome with root, leaf, and flower transcriptome. A reciprocal BLAST search was conducted to compare the two *C. pepo* transcriptomes. Each transcriptome was compared to a database created from the other transcriptome using the blastn algorithm with a significance threshold of 1.0E-25.

Transcriptome	Unigenes with no significant similarity to other transcriptome	Unigenes with significant similarity to other transcriptome
Fruit and seed	21,961 (39%)	33,988 (61%)
Root, leaf, and flower	2,177 (4%)	47,433 (96%)

unigenes had the same reciprocal match in both blast results, suggesting these could be orthologous unigenes between zucchini and acorn squash.

To characterize the unique sequences in the fruit and seed transcriptome, their matches in the SWISS-PROT, TrEMBL, nr, and TAIR10 sequence databases were assessed (e-value 1.0E-20). Of the unique fruit and seed unigenes, 53% (11,671) had a predicted open reading frame and 38% (8,315) had significant similarity to previously identified sequences. Therefore, our transcriptome sequencing identified more than 8,000 novel unigenes in the fruit and seed transcriptome that had not been previously sequenced in the species. To examine gene enrichment in the two datasets, GO terms for the root, leaf, and flower transcriptome were obtained from Blanca et al. (2011). A total of 606 unique root, leaf, and flower unigenes and 5,437 unique fruit and seed unigenes had at least one associated GO term. Statistically significant differences between the GO terms represented in the two samples were found for the level 2 biological process, molecular function, and cellular component groups ($p<0.05$). The unique fruit and seed unigenes were enriched for development-related GO terms, as was expected for the fruit and seed tissues sampled at five different developmental time points. These GO terms included "cellular developmental process," "anatomical structure morphogenesis," "developmental maturation," "developmental process involved in reproduction," "seed development," "embryo development," and "cellular process involved in reproduction", confirming that unigenes unique to fruit and seed development were represented in our transcriptome.

Identification of Fruit Quality Genes

Several metabolic processes that are crucial for winter squash fruit quality occur during fruit development, including the synthesis of carotenoids, sucrose, and starch. To demonstrate the future utility of this transcriptome for the study of fruit quality, we identified genes involved in these processes in our transcriptome. A BLASTn search (Altschul et al., 1997) was conducted to query all biosynthetic genes known to be involved in carotenoid, sugar, and starch metabolism against the fruit and seed transcriptome to identify squash homologs of the genes. Squash unigenes homologous to functionally characterized genes were identified for 18 genes involved in carotenoid synthesis, storage, and degradation (see Table 4.4), 18 genes involved in sucrose metabolism (see Table 4.5), and 14 genes involved in starch metabolism (see Table 4.6), demonstrating that this transcriptome fully captures gene expression related to these metabolic processes of interest.

Structural genes known to control flux or act at pathway branch-points in carotenoid formation will be informative in future analyses of color and nutrient concentration in developing squash fruit. The carotenoid unigenes identified include the gene DOXP synthase, whose product is predicted to synthesize a regulatory step of carotenoid biosynthesis in tomato fruit (Lois et al., 2000). DOXP synthase is the first enzyme in the DOXP pathway that produces isopentenyl diphosphate, the precursor for carotenoids (Hirschberg, 2001). The unigene for phytoene synthase, which performs the first committed step in carotenoid synthesis and has been shown to be rate-limiting in marigolds, canola, and tomato (Hirschberg, 2001), was also identified. An important branch-point in the squash carotenoid pathway involves lycopene ε-

Table 4.4 Carotenoid metabolism gene homologs found in fruit transcriptome. Homologs to functionally characterized carotenoid metabolism genes were identified in the fruit and seed transcriptome through a BLAST search. The best candidate(s) for each gene are listed, identified by unigene number.

Carotenoid metabolism gene	***C. pepo* unigene homolog(s)**
β-carotene hydroxylase	CP112262
carotenoid cleavage dioxygenase	CP134539
carotenoid isomerase	CP040968
DOXP reductoisomerase	CP135419
DOXP synthase	CP135924, CP060201, CP060202
ε-hydroxylase	CP029795
geranylgeranyl hydrogenase	CP056527, CP056528, CP072215, CP072216
GGPP synthase	CP135982, CP031308
IPP isomerase	CP096828, CP006070, CP135789
lycopene β-cyclase	CP073600
lycopene ε-cyclase	CP120024
orange gene	CP135849
phytoene desaturase	CP120142, CP120144
phytoene synthase	CP005213, CP097163
violaxanthin de-epoxidase[z]	CP068838
ζ-carotene desaturase	CP071571
ζ-carotene isomerase	CP003737
zeaxanthin epoxidase	CP041038

[z]Homology found to genes annotated with only putative function.

Table 4.5 Sugar metabolism gene homologs found in fruit transcriptome. Homologs to functionally characterized simple sugar metabolism genes were identified in the fruit and seed transcriptome through a BLAST search. The best candidate(s) for each gene are listed, identified by unigene number.

Simple sugar metabolism gene	***C. pepo* unigene homolog(s)**
acid α-galactosidase	CP031132, CP041982
acid invertase	CP078022, CP039349
alkaline α-galactosidase	CP135438, CP082553, CP082552, CP097985
fructokinase	CP138498, CP138398
galactokinase	CP100962, CP100963
galactose-1-phosphate uridyltransferase	CP115252, CP115240
hexokinase	CP081918
hexose transporter	CP139122, CP032998, CP047946
neutral invertase	CP120038, CP120034, CP140685
phosphofructokinase	CP076642
phosphoglucomutase	CP124387, CP005960
sucrose synthase	CP039490, CP036200, CP033875
sucrose transporter	CP123643, CP123649, CP008293, CP009234
sucrose-phosphate phosphatase	CP033610, CP078010
sucrose-phosphate synthase	CP125836
UDPglucose epimerase	CP083217, CP083218, CP040136
UDPglucose/galactose pyrophosphorylase	CP134768
UDPglucose pyrophosphorylase	CP062357

Table 4.6 Starch metabolism gene homologs found in fruit transcriptome. Homologs to functionally characterized starch metabolism genes were identified in the fruit and seed transcriptome through a BLAST search. The best candidate(s) for each gene are listed, identified by unigene number.

Starch metabolism gene	***C. pepo* unigene homolog(s)**
ADP-glucose pyrophosphorylase L	CP101035, CP101034
ADP-glucose pyrophosphorylase S	CP126587
α-amylase	CP135677
α-glucosidase	CP144486, CP116493, CP116492, CP006281
amyloplastidial ATP/ADP translocator	CP098764, CP001903, CP098765, CP001904
β-amylase	CP070598, CP070601, CP070602, CP074740
cell wall invertase	CP052558
debranching enzyme	CP039826
glucose phosphate transporter	CP086414, CP081420, CP034210, CP034209, CP037833
phosphoglucose isomerase	CP109433
starch phosphorylase	CP033924, CP004861, CP136313
starch synthase - granule bound	CP118399, CP118397, CP118398, CP118396
starch synthase - soluble	CP113496, CP113492
starch-branching enzyme	CP033816, CP120408, CP120407

cyclase and lycopene β-cyclase, also identified in the transcriptome, which both encode enzymes that act on lycopene to direct flux towards either lutein or β-carotene (Lu and Li, 2008), the two primary carotenoids in *C. pepo* (Azevedo-Meleiro and Rodriguez-Amaya, 2007). Natural variation in lycopene ε-cyclase in maize influences the partition of carotenoids between the two branches of the pathway (Harjes et al.,

2008), so these genes are also likely to impact the ratio of lutein to β-carotene in squash. The orange gene (*Or*) was also identified, which encodes an enzyme that is important for regulating the accumulation of carotenoids through the differentiation of chromoplasts and enables high levels of carotenoid accumulation in cauliflower and potato (Lu et al., 2006).

The unigenes also included key genes involved in carbohydrate metabolism. Sucrose-phosphate synthase and sucrose-phosphate phosphatase were identified, which encode enzymes that act sequentially to synthesize sucrose from UDP-glucose and fructose 6-phosphate. Sucrose-phosphate synthase is a regulatory step for sucrose biosynthesis (Huber and Huber, 1996) and its higher enzyme activity coincided with higher sucrose accumulation in pear (Zhang et al., 2014) and in muskmelon (Hubbard et al., 1989). Starch synthesis genes identified included ADP-glucose pyrophosphorylase and AATPT (amyloplastidial ATP/ADP translocator). ADP-glucose pyrophosphorylase synthesizes ADP-glucose from glucose-1-phosphate and is the first committed step in starch biosynthesis (Geigenberger et al., 2004). AATPT transfers ATP and ADP between the cytosol and the amyloplast, providing the ATP needed for starch biosynthesis (Geigenberger et al., 2004). Also identified was a unigene encoding α-amylase, which is the primary enzyme in squash that breaks down starch, yielding simple sugars that confer sweetness (Irving et al., 1999). Together, these genes form the core of hypotheses that will be tested in future comparative studies of winter squash quality.

Conclusion

In this study, we report the first *Cucurbita pepo* winter squash fruit and seed transcriptome, with more than 141 million high-quality paired-end sequencing reads compiled over five developmental time points and assembled into a final transcriptome of 55,949 unigenes. Approximately 85% of the unigenes with open reading frames shared homology with known proteins and 62% could be functionally annotated. This transcriptome was compared to the cucumber and watermelon genomes, as well as the previously sequenced *C. pepo* root, leaf, and flower transcriptome, which confirmed that it was comprehensive, had a majority of full-length unigenes, and contained unigenes unique to fruit and seed development. This fruit and seed transcriptome represents a major contribution to *C. pepo* genomic resources, with more than 8,000 *C. pepo* unigenes homologous to known genes that are new to the sequenced exome and will be useful in future genome annotation efforts. Further, the identification of likely candidates for carotenoid and carbohydrate metabolism genes suggests that this novel resource will enable further study of fruit quality and development to enhance future squash breeding efforts that seek to produce higher quality fruit with greater nutritional and culinary value.

Acknowledgments

We thank the Cornell University Biotechnology Resource Center and Giovanna Danies for bioinformatics support, Li Li and Zhangjun Fei for experimental design advice, Scott Anthony for providing plant care, and Paige Roosa for comments on the manuscript. This research was supported through funds from Cornell University

and the Vegetable Breeding Institute. Support for Lindsay Wyatt was provided by a Cornell University Presidential Life Sciences Fellowship, USDA National Needs Graduate Fellowship Competitive Grant No. 2008-38420-04755 from the National Institute of Food and Agriculture, and the Agriculture and Food Research Initiative Competitive Grant No. 2013-67011-21122 from the USDA National Institute of Food and Agriculture.

References

Altschul, S.F., T.L. Madden, A.A. Schaffer, J.H. Zhang, Z. Zhang, W. Miller, and D.J. Lipman, 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25:3389-3402.

Ando, K., K.M. Carr, and R. Grumet, 2012. Transcriptome analyses of early cucumber fruit growth identifies distinct gene modules associated with phases of development. *BMC Genomics* 13.

Ando, K. and R. Grumet, 2010. Transcriptional Profiling of Rapidly Growing Cucumber Fruit by 454-Pyrosequencing Analysis. *Journal of the American Society for Horticultural Science* 135:291-302.

Apweiler, R., M.J. Martin, C. O'Donovan, M. Magrane, Y. Alam-Faruque, E. Alpi, R. Antunes, J. Arganiska, E.B. Casanova, B. Bely, M. Bingley, C. Bonilla, R. Britto, B. Bursteinas, W.M. Chan, G. Chavali, E. Cibrian-Uhalte, A. Da Silva, M. De Giorgi, E. Dimmer, F. Fazzini, P. Gane, A. Fedotov, L.G. Castro, P. Garmiri, E. Hatton-Ellis, R. Hieta, R. Huntley, J. Jacobsen, R. Jones, D. Legge, W.D. Liu, J. Luo, A. MacDougall, P. Mutowo, A. Nightingale, S. Orchard, S. Patient, K. Pichler, D. Poggioli, S. Pundir, L. Pureza, G.Y. Qi, S. Rosanoff, T. Sawford, H. Sehra, E. Turner, V. Volynkin, T. Wardell, X. Watkins, H. Zellner, M. Corbett, M. Donnelly, P. van Rensburg, M. Goujon, H. McWilliam, R. Lopez, I. Xenarios, L. Bougueleret, A. Bridge, S. Poux, N. Redaschi, A. Auchincloss, K. Axelsen, P. Bansal, D. Baratin, P.A. Binz, M.C. Blatter, B. Boeckmann, J. Bolleman, E. Boutet, L. Breuza, E. de Castro, L. Cerutti, E. Coudert, B. Cuche, M. Doche, D. Dornevil, S. Duvaud, A. Estreicher, L. Famiglietti, M. Feuermann, E. Gasteiger, S. Gehant, V. Gerritsen, A. Gos, N. Gruaz-Gumowski, U. Hinz, C. Hulo, J. James, F. Jungo, G. Keller, V. Lara, P. Lemercier, J. Lew, D. Lieberherr, X. Martin, P. Masson, A. Morgat, T. Neto, et al., 2013. Update on activities at the Universal Protein Resource (UniProt) in 2013. *Nucleic Acids Research* 41:D43-D47.

Arumuganathan, K. and E.D. Earle, 1991. Nuclear DNA Content of Some Important Plant Species. *Plant Molecular Biology Reporter* 9:208-218.

Ashburner, M., C.A. Ball, J.A. Blake, D. Botstein, H. Butler, J.M. Cherry, A.P. Davis, K. Dolinski, S.S. Dwight, J.T. Eppig, M.A. Harris, D.P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J.C. Matese, J.E. Richardson, M. Ringwald, G.M. Rubin, G. Sherlock, and C. Gene Ontology, 2000. Gene Ontology: tool for the unification of biology. *Nature Genetics* 25:25-29.

Azevedo-Meleiro, C.H. and D.B. Rodriguez-Amaya, 2007. Qualitative and quantitative differences in carotenoid composition among *Cucurbita moschata*,

Cucurbita maxima, and *Cucurbita pepo*. *Journal of Agricultural and Food Chemistry* 55:4027-4033.

Blanca, J., J. Canizares, C. Roig, P. Ziarsolo, F. Nuez, and B. Pico, 2011. Transcriptome characterization and high throughput SSRs and SNPs discovery in *Cucurbita pepo* (Cucurbitaceae). *BMC Genomics* 12.

Camon, E., M. Magrane, D. Barrell, D. Binns, W. Fleischmann, P. Kersey, N. Mulder, T. Oinn, J. Maslen, A. Cox, and R. Apweiler, 2003. The gene ontology annotation (GOA) project: Implementation of GO in SWISS-PROT, TrEMBL, and InterPro. *Genome Research* 13:662-672.

Carlson, M., S. Falcon, H. Pages, and N. Li, GO.db: A set of annotation maps describing the entire Gene Ontology. R package version 2.6.1 ed.

Cazzonelli, C.I. and B.J. Pogson, 2010. Source to sink: regulation of carotenoid biosynthesis in plants. *Trends in Plant Science* 15:266-274.

Clepet, C., T. Joobeur, Y. Zheng, D. Jublot, M.Y. Huang, V. Truniger, A. Boualem, M.E. Hernandez-Gonzalez, R. Dolcet-Sanjuan, V. Portnoy, A. Mascarell-Creus, A.I. Cano-Delgado, N. Katzir, A. Bendahmane, J.J. Giovannoni, M.A. Aranda, J. Garcia-Mas, and Z.J. Fei, 2011. Analysis of expressed sequence tags generated from full-length enriched cDNA libraries of melon. *BMC Genomics* 12.

Conesa, A., S. Gotz, J.M. Garcia-Gomez, J. Terol, M. Talon, and M. Robles, 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21:3674-3676.

Corrigan, V.K., P.L. Hurst, and J.F. Potter, 2001. Winter squash (*Cucurbita maxima*) texture: sensory, chemical, and physical measures. *New Zealand Journal of Crop and Horticultural Science* 29:111-124.

Cumarasamy, R., V. Corrigan, P. Hurst, and M. Bendall, 2002. Cultivar differences in New Zealand "Kabocha" (buttercup squash, *Cucurbita maxima*). *New Zealand Journal of Crop and Horticultural Science* 30:197-208.

Diaz, A., M. Fergany, G. Formisano, P. Ziarsolo, J. Blanca, Z.J. Fei, J.E. Staub, J.E. Zalapa, H.E. Cuevas, G. Dace, M. Oliver, N. Boissot, C. Dogimont, M. Pitrat, R. Hofstede, P. van Koert, R. Harel-Beja, G. Tzuri, V. Portnoy, S. Cohen, A. Schaffer, N. Katzir, Y. Xu, H.Y. Zhang, N. Fukino, S. Matsumoto, J. Garcia-Mas, and A.J. Monforte, 2011. A consensus linkage map for molecular markers and Quantitative Trait Loci associated with economically important traits in melon (*Cucumis melo* L.). *BMC Plant Biology* 11.

Esteras, C., P. Gomez, A.J. Monforte, J. Blanca, N. Vicente-Dolera, C. Roig, F. Nuez, and B. Pico, 2012. High-throughput SNP genotyping in *Cucurbita pepo* for map construction and quantitative trait loci mapping. *BMC Genomics* 13.

Ferriol, M. and B. Picó, 2008. Pumpkin and Winter Squash, p. 317-349. In: J. Prohens and F. Nuez (eds.), Vegetables I. Springer New York, New York.

Gajewski, M., J. Radzanowska, H. Danilcenko, E. Jariene, and J. Cerniauskiene, 2008. Quality of Pumpkin Cultivars in Relation to Sensory Characteristics. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca* 36:73-79.

Garcia-Mas, J., A. Benjak, W. Sanseverino, M. Bourgeois, G. Mir, V.M. Gonzalez, E. Henaff, F. Camara, L. Cozzuto, E. Lowy, T. Alioto, S. Capella-Gutierrez, J. Blanca, J. Canizares, P. Ziarsolo, D. Gonzalez-Ibeas, L. Rodriguez-Moreno, M. Droege, L. Du, M. Alvarez-Tejado, B. Lorente-Galdos, M. Mele, L.M. Yang, Y.Q. Weng, A. Navarro, T. Marques-Bonet, M.A. Aranda, F. Nuez, B. Pico, T. Gabaldon, G. Roma, R. Guigo, J.M. Casacuberta, P. Arus, and P. Puigdomenech, 2012. The genome of melon (*Cucumis melo* L.). *Proceedings of the National Academy of Sciences of the United States of America* 109:11872-11877.

Geigenberger, P., M. Stitt, and A.R. Fernie, 2004. Metabolic control analysis and regulation of the conversion of sucrose to starch in growing potato tubers. *Plant Cell and Environment* 27:655-673.

Gong, L., M. Pachner, K. Kalai, and T. Lelley, 2008a. SSR-based genetic linkage map of *Cucurbita moschata* and its synteny with *Cucurbita pepo*. *Genome* 51:878-887.

Gong, L., G. Stift, R. Kofler, M. Pachner, and T. Lelley, 2008b. Microsatellites for the genus *Cucurbita* and an SSR-based genetic linkage map of *Cucurbita pepo* L. *Theoretical and Applied Genetics* 117:37-48.

Grabherr, M.G., B.J. Haas, M. Yassour, J.Z. Levin, D.A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q.D. Zeng, Z.H. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B.W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman, and A. Regev, 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29:644-652.

Guo, S.G., J.G. Zhang, H.H. Sun, J. Salse, W.J. Lucas, H.Y. Zhang, Y. Zheng, L.Y. Mao, Y. Ren, Z.W. Wang, J.M. Min, X.S. Guo, F. Murat, B.K. Ham, Z.L. Zhang, S. Gao, M.Y. Huang, Y.M. Xu, S.L. Zhong, A. Bombarely, L.A. Mueller, H. Zhao, H.J. He, Y. Zhang, Z.H. Zhang, S.W. Huang, T. Tan, E.L. Pang, K. Lin, Q. Hu, H.H. Kuang, P.X. Ni, B. Wang, J.A. Liu, Q.H. Kou, W.J.

Hou, X.H. Zou, J. Jiang, G.Y. Gong, K. Klee, H. Schoof, Y. Huang, X.S. Hu, S.S. Dong, D.Q. Liang, J. Wang, K. Wu, Y. Xia, X. Zhao, Z.Q. Zheng, M. Xing, X.M. Liang, B.Q. Huang, T. Lv, J.Y. Wang, Y. Yin, H.P. Yi, R.Q. Li, M.Z. Wu, A. Levi, X.P. Zhang, J.J. Giovannoni, Y.F. Li, Z.J. Fei, and Y. Xu, 2013. The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions. *Nature Genetics* 45:51-58.

Harjes, C.E., T.R. Rocheford, L. Bai, T.P. Brutnell, C.B. Kandianis, S.G. Sowinski, A.E. Stapleton, R. Vallabhaneni, M. Williams, E.T. Wurtzel, J.B. Yan, and E.S. Buckler, 2008. Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. *Science* 319:330-333.

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

Huang, S.W., R.Q. Li, Z.H. Zhang, L. Li, X.F. Gu, W. Fan, W.J. Lucas, X.W. Wang, B.Y. Xie, P.X. Ni, Y.Y. Ren, H.M. Zhu, J. Li, K. Lin, W.W. Jin, Z.J. Fei, G.C. Li, J. Staub, A. Kilian, E.A.G. van der Vossen, Y. Wu, J. Guo, J. He, Z.Q. Jia, Y. Ren, G. Tian, Y. Lu, J. Ruan, W.B. Qian, M.W. Wang, Q.F. Huang, B. Li, Z.L. Xuan, J.J. Cao, Asan, Z.G. Wu, J.B. Zhang, Q.L. Cai, Y.Q. Bai, B.W. Zhao, Y.H. Han, Y. Li, X.F. Li, S.H. Wang, Q.X. Shi, S.Q. Liu, W.K. Cho, J.Y. Kim, Y. Xu, K. Heller-Uszynska, H. Miao, Z.C. Cheng, S.P. Zhang, J. Wu, Y.H. Yang, H.X. Kang, M. Li, H.Q. Liang, X.L. Ren, Z.B. Shi, M. Wen, M. Jian, H.L. Yang, G.J. Zhang, Z.T. Yang, R. Chen, S.F. Liu, J.W. Li, L.J. Ma, H. Liu, Y. Zhou, J. Zhao, X.D. Fang, G.Q. Li, L. Fang, Y.R. Li, D.Y. Liu, H.K. Zheng, Y. Zhang, N. Qin, Z. Li, G.H. Yang, S. Yang, L. Bolund, K. Kristiansen, H.C. Zheng, S.C. Li, X.Q. Zhang, H.M. Yang, J. Wang, R.F. Sun, B.X. Zhang, S.Z. Jiang, Y.C. Du, and S.G. Li, 2009. The genome of the cucumber, *Cucumis sativus* L. *Nature Genetics* 41:1275-1281.

Hubbard, N.L., S.C. Huber, and D.M. Pharr, 1989. Sucrose phosphate synthase and acid invertase as determinants of sucrose concentration in developing muskmelon (*Cucumis melo* L.) fruits. *Plant Physiology* 91:1527-1534.

Huber, S.C. and J.L. Huber, 1996. Role and regulation of sucrose-phosphate synthase in higher plants. *Annual Review of Plant Physiology and Plant Molecular Biology* 47:431-444.

Hurst, P.L., V.K. Corrigan, and J. Koolaard, 2006. Genetic analysis of sweetness and textural attributes in winter squash (*Cucurbita maxima*). *New Zealand Journal of Crop and Horticultural Science* 34:359-367.

Idouraine, A., E.A. Kohlhepp, C.W. Weber, W.A. Warid, and J.J. Martinez Tellez, 1996. Nutrient constituents from eight lines of naked seed squash (*Cucurbita pepo* L). *Journal of Agricultural and Food Chemistry* 44:721-724.

Irving, D.E., P.L. Hurst, and J.S. Ragg, 1997. Changes in carbohydrates and carbohydrate metabolizing enzymes during the development, maturation, and ripening of buttercup squash (*Cucurbita maxima* D 'Delica'). *Journal of the American Society for Horticultural Science* 122:310-314.

Irving, D.E., G.J. Shingleton, and P.L. Hurst, 1999. Starch degradation in buttercup squash (*Cucurbita maxima*). *Journal of the American Society for Horticultural Science* 124:587-590.

Iseli, C., C. Jongeneel, and P. Bucher. 1999. ESTScan: a program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. *Proceedings of the International Conference on Intelligent Systems for Molecular Biology* 138-148.

Itle, R.A. and E.A. Kabelka, 2009. Correlation Between L*a*b* Color Space Values and Carotenoid Content in Pumpkins and Squash (*Cucurbita* spp.). *Hortscience* 44:633-637.

Li, B. and C.N. Dewey, 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12.

Li, J.G., X.J. Li, L. Guo, F. Lu, X.J. Feng, K. He, L.P. Wei, Z.L. Chen, L.J. Qu, and H.Y. Gu, 2006. A subgroup of MYB transcription factor genes undergoes highly conserved alternative splicing in Arabidopsis and rice. *Journal of Experimental Botany* 57:1263-1273.

Lois, L.M., M. Rodriguez-Concepcion, F. Gallego, N. Campos, and A. Boronat, 2000. Carotenoid biosynthesis during tomato fruit development: regulatory role of 1-deoxy-D-xylulose 5-phosphate synthase. *Plant Journal* 22:503-513.

Loy, J.B., 2004. Morpho-physiological aspects of productivity and quality in squash and pumpkins (*Cucurbita* spp.). *Critical Reviews in Plant Sciences* 23:337-363.

Lu, S. and L. Li, 2008. Carotenoid metabolism: Biosynthesis, regulation, and beyond. *Journal of Integrative Plant Biology* 50:778-785.

Lu, S., J. Van Eck, X. Zhou, A.B. Lopez, D.M. O'Halloran, K.M. Cosman, B.J. Conlin, D.J. Paolillo, D.F. Garvin, J. Vrebalov, L.V. Kochian, H. Kupper, E.D. Earle, J. Cao, and L. Li, 2006. The cauliflower or gene encodes a DnaJ cysteine-rich domain-containing protein that mediates high levels of beta-carotene accumulation. *Plant Cell* 18:3594-3605.

Murkovic, M., U. Mulleder, and H. Neunteufl, 2002. Carotenoid content in different varieties of pumpkins. *Journal of Food Composition and Analysis* 15:633-638.

Murphy, T.M. and M. Gao, 2001. Multiple forms of formamidopyrimidine-DNA glycosylase produced by alternative splicing in *Arabidopsis thaliana*. *Journal of Photochemistry and Photobiology B-Biology* 61:87-93.

Ren, Y., H. Zhao, Q. Kou, J. Jiang, S. Guo, H. Zhang, W. Hou, X. Zou, H. Sun, G. Gong, A. Levi, and Y. Xu, 2012. A high resolution genetic map anchoring scaffolds of the sequenced watermelon genome. *Plos One* 7:e29453.

Robinson, R.W. and D.S. Decker-Walters, 1997. Cucurbits. CAB International, New York.

Rohrmann, J., T. Tohge, R. Alba, S. Osorio, C. Caldana, R. McQuinn, S. Arvidsson, M.J. van der Merwe, D.M. Riano-Pachon, B. Mueller-Roeber, Z.J. Fei, A.N. Nesi, J.J. Giovannoni, and A.R. Fernie, 2011. Combined transcription factor profiling, microarray analysis and metabolite profiling reveals the transcriptional control of metabolic shifts occurring during tomato fruit development. *Plant Journal* 68:999-1013.

Sanchez, A., M. Salicru, and J. Ocana, 2007. Statistical methods for the analysis of high-throughput data based on functional profiles derived from the Gene Ontology. *Journal of Statistical Planning and Inference* 137:3975-3989.

Schmieder, R. and R. Edwards, 2011. Fast Identification and Removal of Sequence Contamination from Genomic and Metagenomic Datasets. *Plos One* 6.

Weeden, F., 1984. Isozyme studies indicate that the genus *Cucurbita* is an ancient tetraploid. *Cucurbit Genetics Cooperative Report* 7:84-85.

Wu, T., S. Luo, R. Wang, Y. Zhong, X. Xu, Y.e. Lin, X. He, B. Sun, and H. Huang, 2014. The first Illumina-based de novo transcriptome sequencing and analysis of pumpkin (*Cucurbita moschata* Duch.) and SSR marker development. *Molecular Breeding* 34:1437-1447.

Yoshida, H., Y. Shougaki, Y. Hirakawa, Y. Tomiyama, and Y. Mizushina, 2004. Lipid classes, fatty acid composition and triacylglycerol molecular species in the kernels of pumpkin (*Cucurbita* spp) seeds. *Journal of the Science of Food and Agriculture* 84:158-163.

Zdobnov, E.M. and R. Apweiler, 2001. InterProScan - an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* 17:847-848.

Zenoni, S., A. Ferrarini, E. Giacomelli, L. Xumerle, M. Fasoli, G. Malerba, D. Bellin, M. Pezzotti, and M. Delledonne, 2010. Characterization of Transcriptional Complexity during Berry Development in *Vitis vinifera* Using RNA-Seq. *Plant Physiology* 152:1787-1795.

Zhang, H.P., J.Y. Wu, G.H. Qin, G.F. Yao, K.J. Qi, L.F. Wang, and S.L. Zhang, 2014. The role of sucrose-metabolizing enzymes in pear fruit that differ in sucrose accumulation. *Acta Physiologiae Plantarum* 36:71-77.

Zhang, W.W., J.S. Pan, H.L. He, C. Zhang, Z. Li, J.L. Zhao, X.J. Yuan, L.H. Zhu, S.W. Huang, and R. Cai, 2012. Construction of a high density integrated genetic map for cucumber (*Cucumis sativus* L.). *Theoretical and Applied Genetics* 124:249-259.

Zheng, Y., L.J. Zhao, J.P. Gao, and Z.J. Fei, 2011. iAssembler: a package for de novo assembly of Roche-454/Sanger transcriptome sequences. *BMC Bioinformatics* 12.

Zraidi, A., G. Stift, M. Pachner, A. Shojaeiyan, L. Gong, and T. Lelley, 2007. A consensus map for *Cucurbita pepo*. *Molecular Breeding* 20:375-388.

CHAPTER 5

COMPARATIVE ANALYSIS OF *CUCURBITA PEPO* METABOLISM THROUGHOUT FRUIT DEVELOPMENT IN ACORN SQUASH AND OILSEED PUMPKIN

Abstract

Winter squash can be used for consumption of both fruit mesocarp and seeds. Cultivars bred for fruit consumption are selected for fruit quality traits such as carotenoid content, percent dry matter, and percent soluble solids, while these traits are virtually ignored in oilseed pumpkins. We sequenced the fruit and seed transcriptome of two cultivars developed for different purposes: an acorn squash, 'Sweet REBA', and an oilseed pumpkin, 'Lady Godiva'. Putative metabolic pathways were developed for carotenoid, starch, and sucrose synthesis in winter squash fruit and squash homologs were identified for each of the structural genes in the pathways. Gene expression, especially of known rate limiting and branch point genes, corresponded with metabolite accumulation both across development and between the two cultivars. Thus, transcriptional regulation of metabolite genes is an important factor in winter squash fruit quality.

Introduction

Cucurbita pepo is the most genetically and morphologically diverse species within *Cucurbita*, spanning two subspecies, *C. pepo* ssp. *pepo* and *C. pepo* ssp. *texana*. Market classes of *C. pepo* include summer squash and zucchini that are eaten at an immature stage, decorative gourds and pumpkins, oilseed pumpkins, and acorn and delicata winter squash that are consumed when ripe and have a highly refined

eating quality. Often in the breeding process, certain characteristics are prioritized at the expense of others in each market class. This contrast is particularly pronounced with oilseed pumpkins, where the fruit is essentially inedible because efforts have focused primarily on improving seed characteristics. These extremes in fruit phenotypes resulting from divergent breeding for fruit or seed consumption provide an excellent opportunity to study the genetic basis of fruit quality.

Fruit quality is one of the most important breeding goals for winter squash. While fruit quality involves the complex interplay of flavor, texture, and appearance, there are three easily measured metabolites that have a major impact on quality. First, carotenoids are the source of winter squash fruit color and provide some of the greatest nutritional benefits of eating squash. The primary carotenoids in *C. pepo* are β-carotene and lutein (Azevedo-Meleiro and Rodriguez-Amaya, 2007; Murkovic et al., 2002). Second, starch content is important for winter squash fruit texture. Ideally squash have a smooth, slightly dry consistency, which is correlated with higher starch and dry matter content (Corrigan et al., 2001). Two negative texture attributes, fibrousness and wateriness, are correlated with low starch and dry matter (Corrigan et al., 2001). Third, sugar content is correlated with perceived sweetness of winter squash, with sucrose as the most important predictor (Corrigan et al., 2000). Sweetness is thought to be important for consumer acceptability and to influence perception of overall squash flavor, as squash flavor was positively correlated with perceived sweetness in taste panels (Cumarasamy et al., 2002). Percent soluble solids, an efficient trait to phenotype, is also correlated with sucrose levels and perceived

sweetness (Harvey et al., 1997). Therefore, by measuring carotenoids, starch, and sugar, it is possible to capture many of the important aspects of fruit quality.

These three fruit quality metabolites are synthesized and accumulate in predictable patterns throughout fruit development. The first stage in fruit development is fruit expansion, where the new fruit quickly grows to reach its maximum size 15-24 days after pollination (DAP). At the same time as fruit expansion, starch accumulates in the growing fruit, reaching maximum dry matter around 30 DAP. At this time, sugars start to accumulate, and eventually, especially after harvest, starch is degraded and sugars continue to accumulate (Irving et al., 1997; Loy, 2004). Carotenoids also steadily accumulate after 30 DAP and in storage (Harvey et al., 1997; Zhang et al., 2014).

Transcriptome sequencing can provide valuable insight into the gene expression underlying developmental changes during fruit development. It has been used to study fruit development and metabolite accumulation in pepper (Liu et al., 2013; Martinez-Lopez et al., 2014), grape (Degu et al., 2014), banana (Asif et al., 2014), and more. Such studies have successfully captured the expression of metabolic genes involved in the biosynthesis of the metabolites of interest. This technique is especially useful in squash ($2n = 2x = 40$), because a genome sequence is not yet available for its 500 Mb genome (Arumuganathan and Earle, 1991). Previous squash transcriptome studies have included analyses of summer squash (*C. pepo*) root, leaf, and flower tissue (Blanca et al., 2011), *C. moschata* leaf, stem, and shoot tissue (Wu et al., 2014), and acorn squash (*C. pepo*) fruit and seed tissue (Wyatt et al., 2015). No

study to date, however, has used transcriptome sequencing to look at differential gene expression throughout winter squash fruit development.

In this study, we describe comparative transcriptome sequencing throughout fruit development of two winter squash bred for different purposes. The first cultivar we sequenced was bred for consumption of the fruit mesocarp: 'Sweet REBA'. 'Sweet REBA' is an inbred acorn squash cultivar whose fruit and seed transcriptome has been sequenced (Wyatt et al., 2015). We also sequenced 'Lady Godiva', an oilseed pumpkin that was developed for the consumption of its hull-less seeds and whose mesocarp is not intended for consumption and is discarded after seed extraction. These two cultivars were sampled at five developmental time points in order to examine how metabolic processes act throughout fruit development to create squash with contrasting fruit quality phenotypes.

Materials and Methods

Plant Material

'Sweet REBA' acorn squash and an inbred line derived from the OP oilseed pumpkin 'Lady Godiva' were grown at the Homer C. Thompson Vegetable Research Farm (Freeville, NY) in summer 2012 using standard horticultural practices. Flowers were manually self-pollinated and mesocarp and seed tissue were collected from fruits at 5, 10, 15, 20, and 40 DAP. Mesocarp samples were taken for carotenoid, soluble solids, and dry matter analysis, and mesocarp and seed samples for RNA-sequencing were flash-frozen in liquid nitrogen and stored at -80°C. Three biological replicates were collected for each genotype per time point, each consisting of two fruits.

Fruit Metabolite Phenotyping

Fruit quality traits were phenotyped using mesocarp samples from the same fruits used for RNA sequencing. Carotenoids were measured with high-performance liquid chromatography using the protocol described in Van Eck et al. (2010). Percent soluble solids (degrees Brix) was determined by freezing samples, allowing them to thaw, and analyzing their juice using a digital refractometer. Percent dry matter was calculated by drying mesocarp samples at 60°C until they reached a constant weight and then dividing the final weight by the starting weight.

RNA Extraction and Sequencing

Total RNA was extracted from mesocarp and seed samples using a Qiagen RNeasy Plant Mini Kit (Qiagen, Valencia, CA). For each biological replicate, equal amounts of total RNA from two fruit were pooled. The Indiana University Genomics Facility prepared barcoded sequencing libraries using the Illumina TruSeq Sample Prep Kit (Illumina, San Diego, CA) and samples were sequenced on three lanes of an Illumina HiSeq 2000.

Transcriptome Assembly

Sequence processing and initial assembly was conducted by the Indiana University Genomics Facility. Trimmomatic (Bolger et al., 2014) was used to remove low-quality reads and trim off adapter sequences. Reads were then normalized using the *in silico* normalization utility in Trinity (Grabherr et al., 2011) and assembled using Velvet (Zerbino and Birney, 2008) and Oases (Schulz et al., 2012). The longest isotig was chosen to represent each isogroup, and the resulting set of unigenes was combined with the published fruit and seed transcriptome (Wyatt et al., 2015) for

additional assembly. The combined set of unigenes was clustered using CD-HIT (Fu et al., 2012) with an identity of 99%, further assembled using CAP3 (Huang and Madan, 1999) with an identity of 99%, and then clustered again using self-blast with an identity of 99%. The longest unigene was chosen to represent each resulting cluster in the final transcriptome.

Transcription Annotation and Analysis

For the unigenes derived from the published fruit and seed transcriptome, annotations were obtained from the published data set (Wyatt et al., 2015). The unigenes from the new fruit and seed sequence data were annotated using ISGA (Hemmerich et al., 2010) to run BLASTx (Altschul et al., 1997) against the NCBI non-redundant protein database and to run BLAST2GO (Conesa et al., 2005) to assign GO terms. The unigenes created by further assembly of unigenes from the two sets were annotated by BLASTx against the SWISS-PROT and TrEMBL databases (Apweiler et al., 2013), keeping hits with an e-value less than 1.0E-20 and InterProScan (Zdobnov and Apweiler, 2001) was run to identify protein domains and assign GO terms.

Gene expression was analyzed by mapping reads to the collapsed transcriptome using TopHat2 (Kim et al., 2013), filtering reads to retain only those with a mapping quality of more than 20, and then using Cuffdiff (Trapnell et al., 2012) to calculate FPKM (fragments per kilobase of exon per million reads mapped) values and identify differentially expressed genes. Carotenoid, sugar, and starch metabolic gene homologs were identified based on their annotations and using BLAST (Altschul

et al., 1997). FPKM values were log transformed using log2+1 and plotted using the R heatmap.2 function in the gplots package (Warnes et al., 2015).

Results

Fruit Metabolite Phenotypes

Self-pollinated fruits were collected from the inbred acorn squash cultivar 'Sweet REBA' and an inbred line derived from the OP oilseed pumpkin 'Lady Godiva' at 5, 10, 15, 20, and 40 DAP (Figure 5.1) and mesocarp tissue was analyzed to determine the concentration of important fruit quality metabolites at each time point. The two primary carotenoids detected in the mesocarp were lutein and β-carotene. The lutein content of 'Lady Godiva' was much higher than 'Sweet REBA', with more than 5 μg/g fresh weight by 40 DAP (Figure 5.2a). Conversely, 'Sweet REBA' accumulated approximately twice as much β-carotene as 'Lady Godiva' (Figure 5.2b). Overall, 'Lady Godiva' accumulated a greater amount of carotenoids. In contrast, 'Sweet REBA' had higher sugar, measured as percent soluble solids,

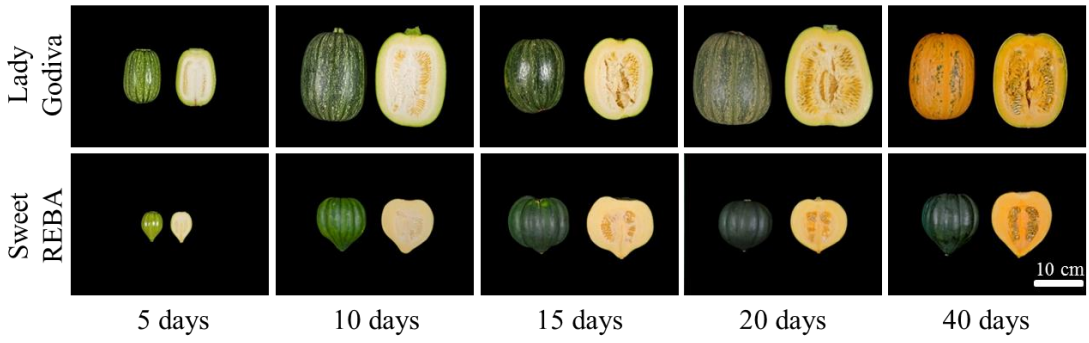


Figure 5.1 'Lady Godiva' oilseed pumpkin and 'Sweet REBA' acorn squash at five developmental time points. Self-pollinated fruit were harvested at 5, 10, 15, 20, and 40 days after pollination. Photos are of the interior and exterior of representative fruit at each time point.

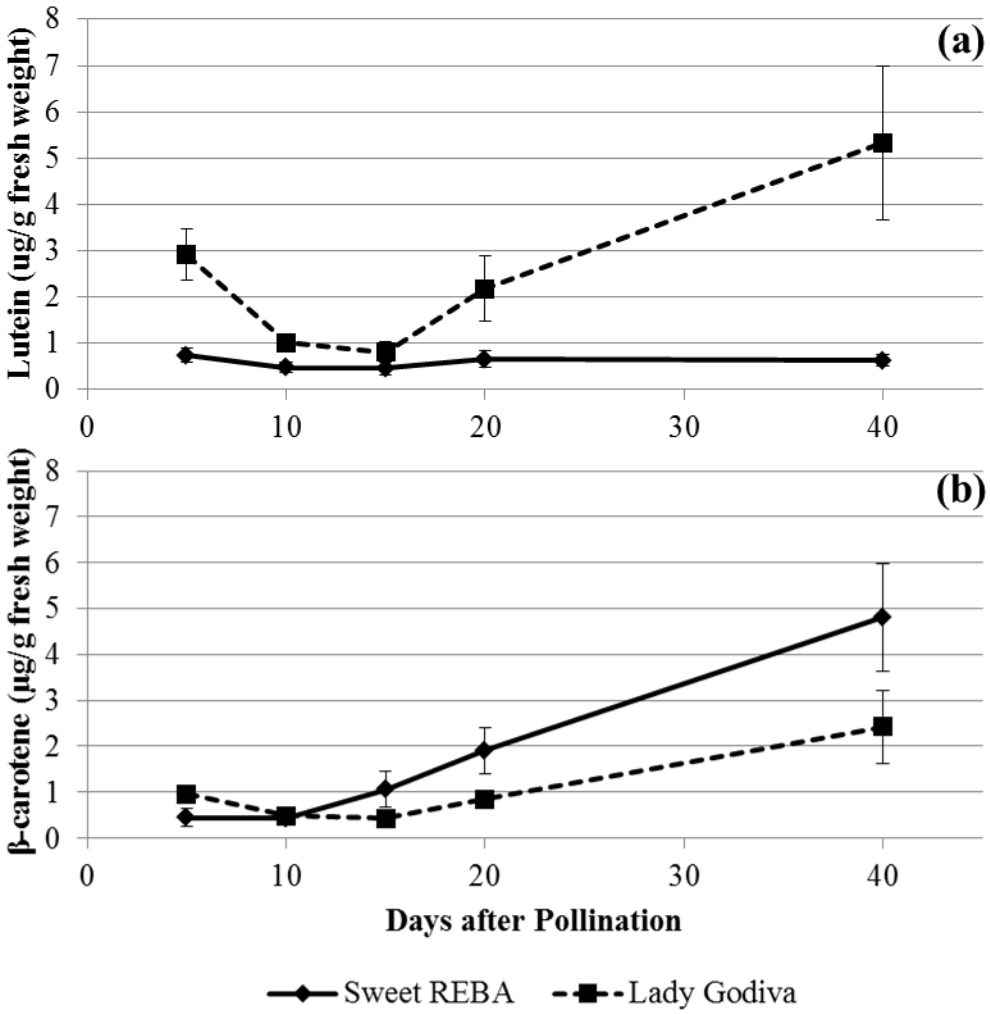


Figure 5.2 Carotenoid concentrations of 'Sweet REBA' and 'Lady Godiva' squash throughout fruit development. Self-pollinated fruit were collected at 5, 10, 15, 20, and 40 days after pollination. All measurements were performed on fruit mesocarp and are averages of six fruit. Error bars indicate the standard deviation. Lutein **(a)** and β-carotene **(b)**, measured using HPLC.

at all five developmental time points (Figure 5.3a). Both cultivars had soluble solids that ranged between 4-6% for the first 20 days of fruit development, but then by 40 DAP, 'Sweet REBA' had more than 9% soluble solids while 'Lady Godiva' stayed near 6%. A similar pattern was seen in starch levels, measured as percent dry matter

(Figure 5.3b). 'Sweet REBA' had a higher dry matter at all time points, with an especially large contrast with 'Lady Godiva' by 40 DAP.

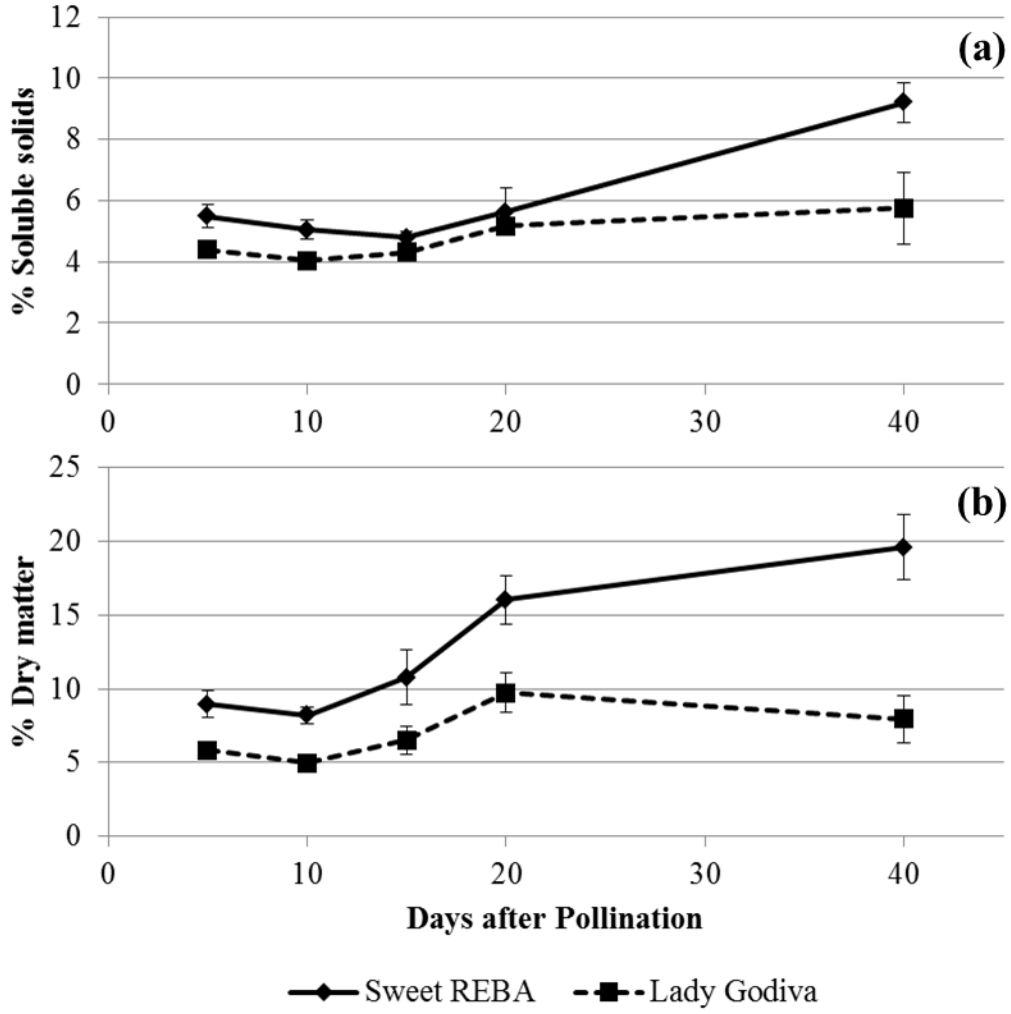


Figure 5.3 Percent soluble solids and percent dry matter of 'Sweet REBA' and 'Lady Godiva' squash throughout fruit development. Self-pollinated fruit were collected at 5, 10, 15, 20, and 40 days after pollination. All measurements were performed on fruit mesocarp and are averages of six fruit. Error bars indicate the standard deviation. Percent soluble solids **(a)**, measured using a refractometer. Percent dry matter **(b)**, measured through comparison of wet and dry weights.

Transcriptome Sequencing and Assembly

RNA was extracted from the mesocarp samples used for phenotypic analysis, as well as from the seeds of the same fruit. Three biological replicates each were sequenced for both mesocarp and seed tissue of each genotype at each of the five developmental time points. Each biological replicate was made up of pooled RNA from two fruits. Barcoded RNA-seq libraries were sequenced on three lanes of an Illumina HiSeq 2000. Using Velvet (Zerbino and Birney, 2008) and Oases (Schulz et al., 2012), the combined set of all the reads was assembled into 159,714 isotigs, and then the longest isotig was chosen to represent each isogroup, yielding 51,978 unigenes. These were combined with 55,949 acorn squash fruit and seed unigenes from Wyatt et al. (2015) to create a comprehensive *C. pepo* fruit and seed transcriptome. The 107,927 combined unigenes were clustered using CD-HIT (Fu et al., 2012), reducing the number slightly to 107,848. Some of the unigenes from the two sets were then found to overlap and were assembled together with CAP3 (Huang and Madan, 1999) to yield 94,535 unigenes. These were clustered using self-blast with an identity threshold of 99% to group transcripts derived from the same gene. The longest unigene was chosen to represent each cluster, resulting in a total of 68,646 unigenes in the transcriptome with an average unigene length of 1411 base pairs.

Metabolite Pathway Analysis

Proposed pathways for carotenoid, starch, and sugar biosynthesis during winter squash fruit development were assembled from the literature and one or more homologs for all of the structural genes involved were identified using annotation data and BLAST. For instances where multiple homologs were identified, the homolog

that was the best fit for our proposed biosynthetic pathway and most consistent with the observed phenotypic differences between the two cultivars was selected (Figures 5.4-5.6).

Homologs were found for all genes in the well-characterized carotenoid pathway (Cazzonelli and Pogson, 2010; Hirschberg, 2001; Lu and Li, 2008). Most of the genes were either more highly expressed in 'Lady Godiva' or had approximately equal expression levels in the two cultivars (Figure 5.4). Three exceptions were lycopene β-cyclase, ε-hydroxylase, and violaxanthin de-epoxidase, which were more highly expressed in 'Sweet REBA'. In both cultivars, GGPP synthase, phytoene synthase, and carotenoid cleavage dioxygenase were more highly expressed than the other genes in the pathway, with FPKM values exceeding 200, while the other genes averaged an overall FPKM of approximately 25. In addition, many genes in the pathway were significantly differentially expressed between 'Sweet REBA' and 'Lady Godiva' at most time points through fruit development, including DOXP synthase, 4-hydroxy-3-methylbut-2-enyl diphosphate reductase, GGPP synthase, phytoene synthase, lycopene β-cyclase, lycopene ε-cyclase, β-carotene hydroxylase, and zeaxanthin epoxidase.

A putative starch biosynthesis pathway for winter squash fruit was assembled from the literature (Dai et al., 2011; Irving et al., 1997; Keeling and Myers, 2010; Tjaden et al., 1998; Weber, 2004). Imported sugars stachyose, raffinose, sucrose, galactose, glucose, and fructose (Richardson et al., 1982) are converted into glucose-6-P, which is imported into the amyloplast and used to synthesize starch. Homologs were found for all of the genes in our proposed pathway and in 'Sweet REBA', most

Figure 5.4 Carotenoid metabolism and associated gene expression. Proposed pathway of carotenoid metabolism in developing winter squash fruit, as derived from the literature (Cazzonelli and Pogson, 2010; Hirschberg, 2001; Lu and Li, 2008). For each step, gene expression throughout fruit development is displayed for the best candidate homolog in our fruit and seed transcriptome. Gene expression is displayed as heat maps depicting the log2+1 transformation of FPKM (fragments per kilobase of exon per million reads mapped) values for 'Sweet REBA' (top heat map) and 'Lady Godiva' (bottom heat map) at 5, 10, 15, 20, and 40 DAP (days after pollination). Pathway genes, abbreviations, and best candidates are as follows: DOXP synthase, DXS, CP224205; DOXP reductoisomerase, DXR, CP250482; 4-hydroxy-3-methylbut-2-enyl diphosphate reductase, HDR, CP106522; IPP isomerase, IPI, CP135789; GGPP synthase, GGPS, CP135982; phytoene synthase, PSY, CP300096; phytoene desaturase, PDS, CP120144; z-carotene isomerase, Z-ISO, CP003737; z-carotene desaturase, ZDS, CP071571; carotenoid isomerase, CRTISO, CP249153; lycopene β-cyclase, LYC-B, CP073600; lycopene ε-cyclase, LYC-E, CP225081; ε-hydroxylase, EOH, CP210319; β-carotene hydroxylase, BOH, CP206844; zeaxanthin epoxidase, ZEP, CP302825; violaxanthin de-epoxidase, VDE, CP300361; carotenoid cleavage dioxygenase, CCD, CP134539. Metabolite abbreviations and names are as follows: GAP, glyceraldehyde-3-phosphate; DOXP, 1-deoxyxylulose 5-phosphate; MEP, 2-C-methyl-D-erythritol 4-phosphate; IPP, isopentenyl diphosphate; DMAPP, dimethylallyl diphosphate; GGPP, geranylgeranyl diphosphate.

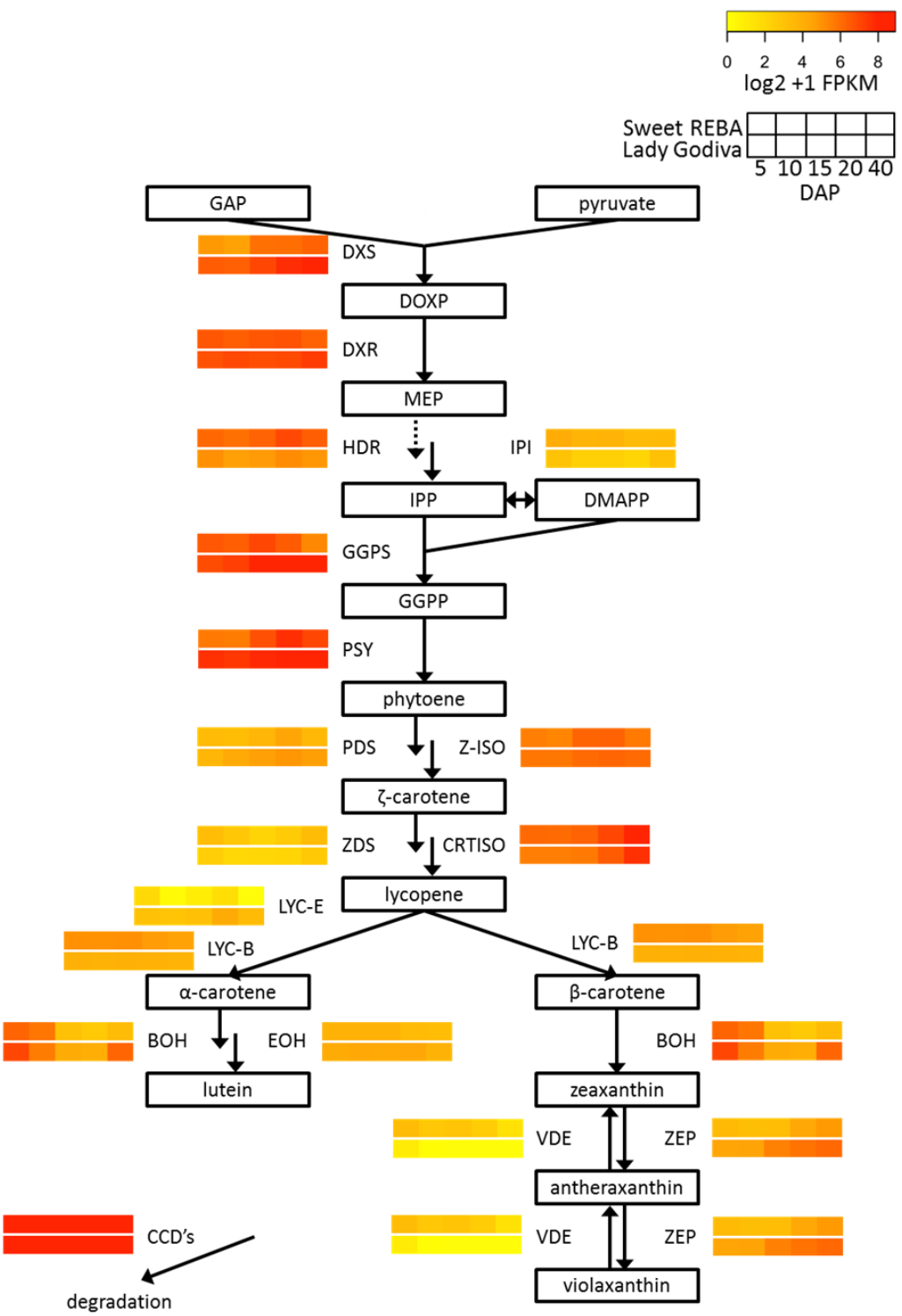

0 2 4 6 8
log2 +1 FPKM
Sweet REBA
Lady Godiva
5 10 15 20 40
DAP
GAP
pyruvate
DXS
DOXP
DXR
MEP
HDR
IPI
IPP
DMAPP
GGPS
GGPP
PSY
phytoene
PDS
Z-ISO
ζ-carotene
ZDS
CRTISO
lycopene
LYC-E
LYC-B
LYC-B
α-carotene
β-carotene
BOH
EOH
BOH
lutein
zeaxanthin
VDE
ZEP
antheraxanthin
VDE
ZEP
violaxanthin
CCD's
degradation

Figure 5.5 Starch biosynthesis and associated gene expression. Proposed pathway of starch biosynthesis in developing winter squash fruit, as derived from the literature (Dai et al., 2011; Irving et al., 1997; Keeling and Myers, 2010; Tjaden et al., 1998; Weber, 2004). For each step, gene expression throughout fruit development is displayed for the best candidate homolog in our fruit and seed transcriptome. Gene expression is displayed as heat maps depicting the log2+1 transformation of FPKM (fragments per kilobase of exon per million reads mapped) values for 'Sweet REBA' (top heat map) and 'Lady Godiva' (bottom heat map) at 5, 10, 15, 20, and 40 DAP (days after pollination).
Metabolites in bold are those imported into the fruit (Richardson et al., 1982). Pathway genes, abbreviations, and best candidates are as follows: acid α-galactosidase, AAG, CP041982; alkaline α-galactosidase, NAG, CP135438; galactokinase, GK, CP100962; UDP-glucose/galactose pyrophosphorylase, UGGP, CP134768; UDP-glucose epimerase, UGE, CP083218; acid invertase, AIN, CP208640; neutral invertase, NIN, CP203652; sucrose synthase, SUS, CP207904; hexokinase, HK, CP206102; fructokinase, FK, CP305181; phosphoglucose isomerase, PG, CP248045; UDP-glucose pyrophosphorylase, UGPase, CP304876; phosphoglucomutase-cytoplasmic, PGM, CP005960; glucose phosphate transporter, GPT, CP081420; amyloplastidial ATP/ADP translocator, AATP, CP251332; phosphoglucomutase-plastidic, PGM, CP111033; ADP-glucose pyrophosphorylase, AGPase, CP101034 (large subunit, top heat map), CP212743 (small subunit, bottom heatmap); soluble starch synthase, SS, CP113492; granule-bound starch synthase, GBSS, CP202177; starch-branching enzyme, SBE, CP304556; debranching enzyme, DBE, CP066452; starch phosphorylase, SP, CP306266.

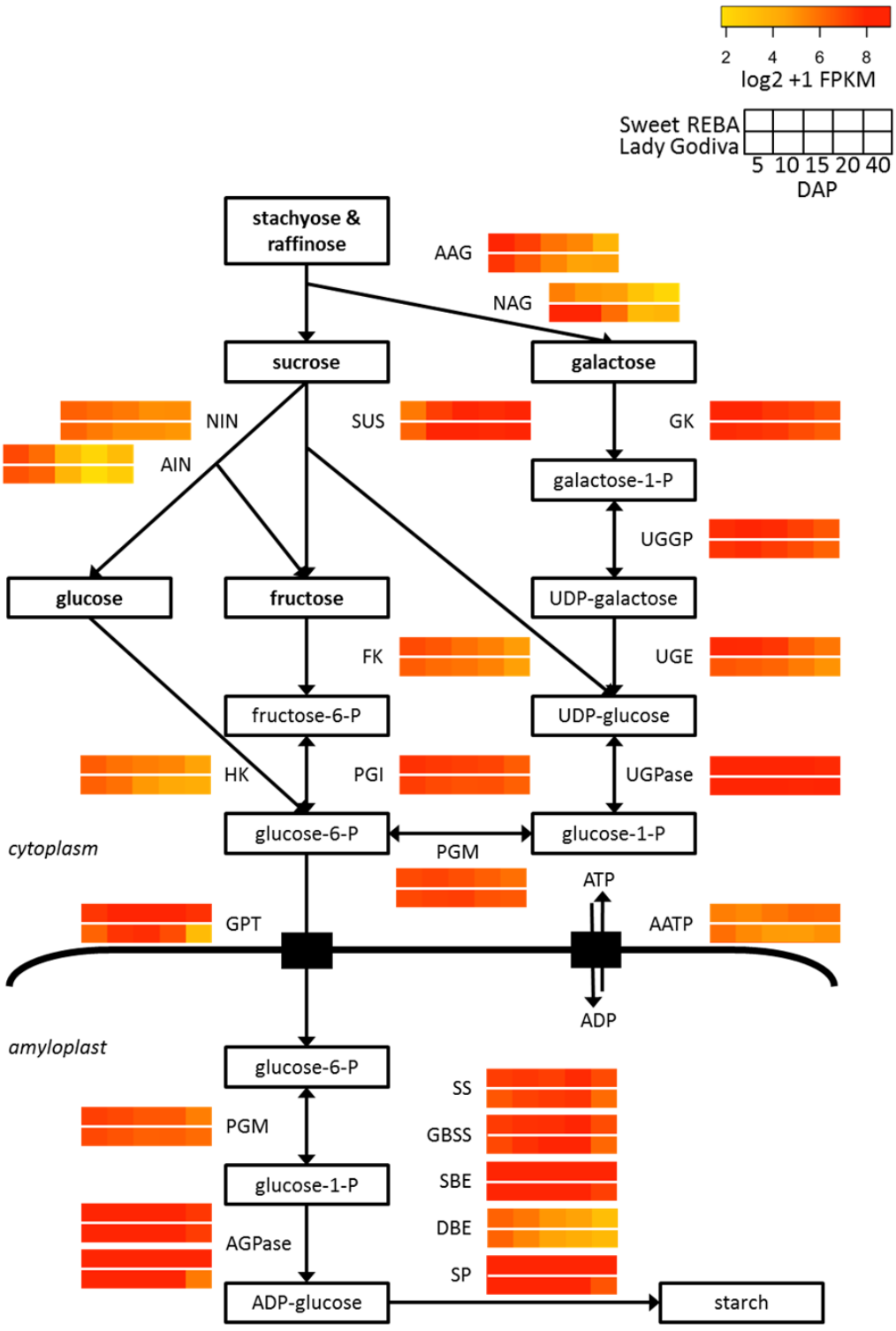
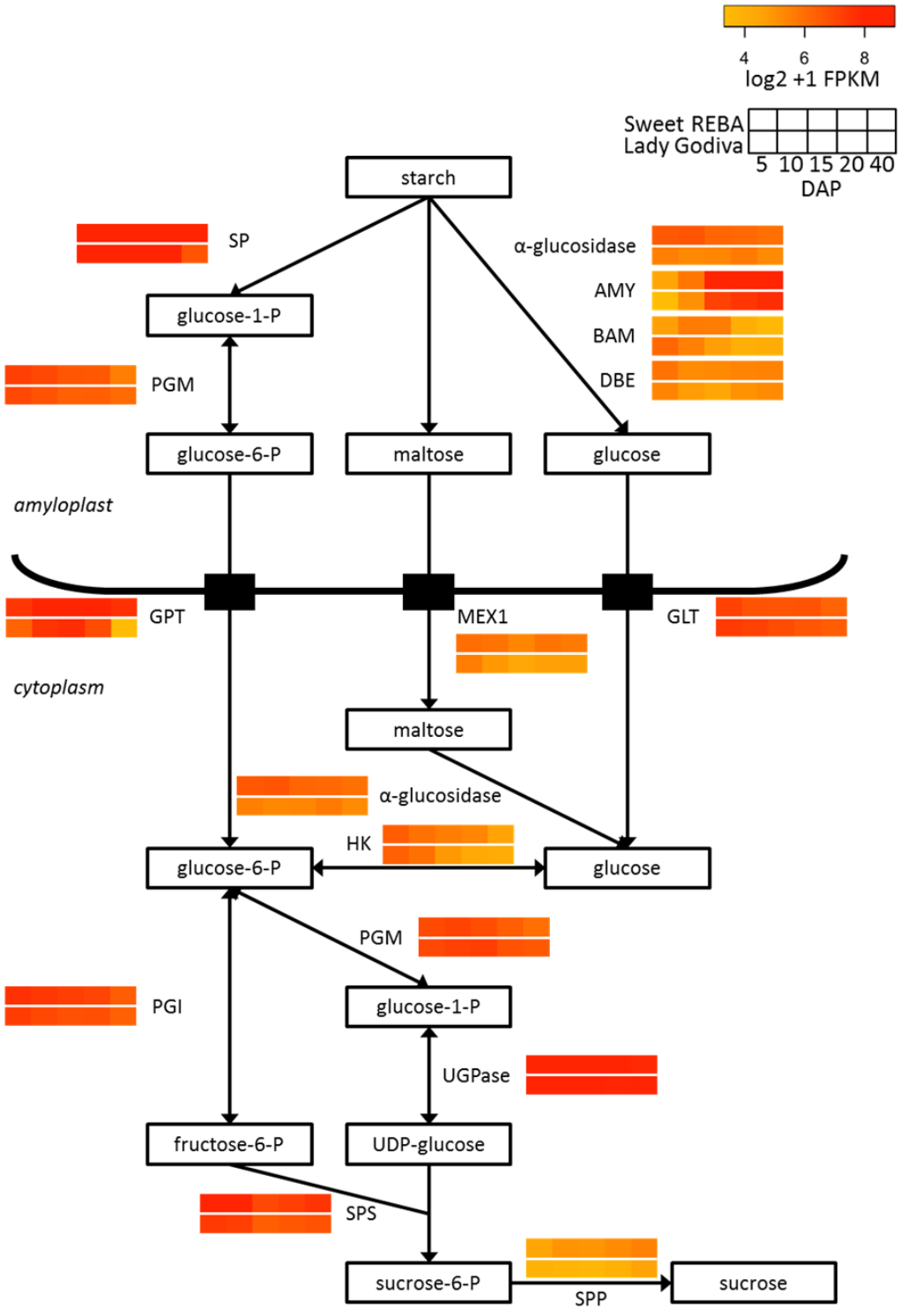

2 4 6 8
log2 +1 FPKM
Sweet REBA
Lady Godiva
5 10 15 20 40
DAP
stachyose & raffinose
AAG
NAG
sucrose
galactose
NIN
AIN
SUS
GK
galactose-1-P
UGGP
glucose
fructose
UDP-galactose
FK
UGE
fructose-6-P
UDP-glucose
HK
PGI
UGPase
glucose-6-P
glucose-1-P
PGM
cytoplasm
ATP
GPT
AATP
ADP
amyloplast
glucose-6-P
SS
PGM
GBSS
glucose-1-P
SBE
DBE
AGPase
SP
ADP-glucose
starch

Figure 5.6 Sucrose biosynthesis and associated gene expression. Proposed pathway of starch breakdown and sucrose biosynthesis in developing winter squash fruit, as derived from the literature (Irving et al., 1999; Keeling and Myers, 2010; Malone et al., 2006; Sharkey et al., 2004; Smith, 2012). For each step, gene expression throughout fruit development is displayed for the best candidate homolog in our fruit and seed transcriptome. Gene expression is displayed as heat maps depicting the log2+1 transformation of FPKM (fragments per kilobase of exon per million reads mapped) values for 'Sweet REBA' (top heat map) and 'Lady Godiva' (bottom heat map) at 5, 10, 15, 20, and 40 DAP (days after pollination). Pathway genes, abbreviations, and best candidates are as follows: α-amylase, AMY, CP111961; β-amylase, BAM, CP304060; debranching enzyme, DBE, CP302373; α-glucosidase, α-glucosidase, CP200748; starch phosphorylase, SP, CP306266; phosphoglucomutase-plastidic, PGM, CP111033; glucose-phosphate transporter, GPT, CP081420; MEX1, MEX1, CP201523; plastidic glucose transporter, GLT, CP235746; hexokinase, HK, CP206102; phosphoglucomutase-cytoplasmic, PGM, CP005960; UDP-glucose pyrophosphorylase, UGPase, CP304876; phosphoglucose isomerase, PG, CP248045; sucrose-phosphate synthase, SPS, CP125836; sucrose-phosphate phosphatase, SPP, CP033610.


4
6
8
log2 +1 FPKM
Sweet REBA
Lady Godiva
5 10 15 20 40
DAP
starch
SP
α-glucosidase
AMY
BAM
DBE
glucose-1-P
PGM
glucose-6-P
maltose
glucose
amyloplast
GPT
MEX1
GLT
cytoplasm
maltose
α-glucosidase
HK
glucose-6-P
glucose
PGM
PGI
glucose-1-P
UGPase
fructose-6-P
UDP-glucose
SPS
sucrose-6-P
SPP
sucrose

of the genes were expressed at a level equal to or higher than in 'Lady Godiva' (Figure 5.5). Three exceptions were alkaline α-galactosidase, sucrose synthase, and granule-bound starch synthase, which were more highly expressed in 'Lady Godiva'. Three genes with much higher expression than the others in both cultivars were ADP-glucose pyrophosphorylase, starch-branching enzyme, and starch phosphorylase. Starch-branching enzyme in particular had an average FPKM value of more than 1800, which was more than 6-fold higher than the overall average of all of the starch biosynthesis genes. Several genes were differentially expressed between the two cultivars: acid α-galactosidase, alkaline α-galactosidase, UDP-glucose epimerase, glucose phosphate transporter, and the large subunit of ADP-glucose pyrophosphorylase.

After squash reach a maximum dry weight, starch is hydrolyzed over the remainder of fruit development and storage and the resulting sugars are used to synthesize sucrose. We assembled a putative pathway for this process from the literature (Irving et al., 1999; Keeling and Myers, 2010; Malone et al., 2006; Sharkey et al., 2004; Smith, 2012) and found homologs in our fruit and seed transcriptome for all of the genes involved. As with the proposed starch synthesis pathway, most of the genes in the sucrose synthesis pathway were either more highly expressed in 'Sweet REBA' or were equally expressed in both cultivars (Figure 5.6). The only exception was the plastidic glucose transporter, which was more highly expressed in 'Lady Godiva'. Starch phosphorylase, as in the starch biosynthesis pathway, was very highly expressed in both cultivars. Six of the genes involved in starch breakdown and sucrose synthesis were differentially expressed at four or five of the five

developmental time points: α-amylase, α-glucosidase, glucose phosphate transporter, MEX1, sucrose-phosphate synthase, and sucrose-phosphate phosphatase.

Discussion

C. pepo Fruit and Seed Transcriptome

In this study, we successfully sequenced and assembled new transcriptome data from *C. pepo* fruit and seeds throughout development. Our updated fruit and seed transcriptome now contains data from the oilseed pumpkin 'Lady Godiva' and has more full-length unigenes created by assembling partial unigenes from the new data set with partial unigenes from the first fruit and seed transcriptome (Wyatt et al., 2015). This updated transcriptome has a greater number of unigenes than the two published *C. pepo* transcriptomes, which could be due to an inability to collapse the different alleles sequenced from the two genotypes for some genes. When the in-progress *C. pepo* genome sequencing effort is completed, use of the genome will enable us reduce the number of unigenes to accurately reflect the number of genetic loci, removing redundancy. Based on our identification of homologs of all the metabolic genes of interest, the transcriptome successfully captured gene expression throughout winter squash fruit and seed development.

Fruit Metabolite Phenotypes

Our phenotypic analysis revealed the expected differences in the fruit quality metabolites of 'Sweet REBA' and 'Lady Godiva', which were bred for different purposes. The two primary carotenoids detected were lutein and β-carotene, as found in previous *C. pepo* carotenoid analyses (Azevedo-Meleiro and Rodriguez-Amaya,

2007; Itle and Kabelka, 2009; Murkovic et al., 2002; Obrero et al., 2013). The increase in carotenoid levels over fruit development was also found in studies of carotenoid accumulation in *C. pepo* (Obrero et al., 2013) and *C. moschata* and *C. maxima* (Nakkanong et al., 2012a). The final lutein and β-carotene concentration in our two cultivars was within the range reported for *C. pepo* in the literature: 0-23 ug/g β-carotene and 0-10 ug/g lutein (Azevedo-Meleiro and Rodriguez-Amaya, 2007; Itle and Kabelka, 2009; Murkovic et al., 2002). We were surprised to find that 'Lady Godiva' accumulated more total carotenoids than 'Sweet REBA' since the flesh of this fruit is not consumed, but it is still well within the range of normal variation for *C. pepo*. The higher carotenoid levels of 'Lady Godiva' could be a pleiotropic effect of selecting for dark green seed and seed oil color, traits that are prized in oilseed pumpkins.

The two carbohydrate metabolite analyses we conducted revealed differences in fruit quality traits between the two cultivars. 'Sweet REBA,' which was bred for consumption of fruit mesocarp, had a higher percent dry matter and percent soluble solids than 'Lady Godiva', which was selected for seed traits. Dry matter generally increased over fruit development, with a slight dip at 10 DAP, which is during a period of rapid fruit expansion. This increase in dry matter and starch content over fruit development is consistent with previous studies of starch accumulation in winter squash (Culpepper and Moon, 1945; Harvey et al., 1997; Irving et al., 1997; Nakkanong et al., 2012b), and the higher dry matter in 'Sweet REBA' underlies its smoother texture. The percent soluble solids measured across fruit development is also in line with previous studies, in which soluble solids generally stayed at a low

level and then started increasing around 40 DAP (Culpepper and Moon, 1945; Harvey et al., 1997; Irving et al., 1997).

Comparative Gene Expression in the Carotenoid Pathway

Carotenoid structural gene expression levels corresponded with carotenoid accumulation observed across fruit development. This pattern is shared in other cucurbits, as reported in summer squash (Obrero et al., 2013) and watermelon (Grassi et al., 2013). 'Lady Godiva' had both higher total carotenoid concentration and higher expression of carotenoid biosynthesis genes, with eight out of the sixteen genes significantly differentially expressed at most or all of the time points sampled. This correlation between carotenoid gene expression and levels of carotenoid accumulation has also been observed when comparing both between cultivars of summer squash and between different fruit tissues within a cultivar (Obrero et al., 2013), between *C. moschata* and *C. maxima* (Nakkanong et al., 2012a), and between watermelon cultivars with different fruit flesh colors (Kang et al., 2010; Lv et al., 2015). Taken together, these patterns suggest that transcriptional regulation is an important determinant of carotenoid accumulation in *C. pepo* winter squash, as it is in other squash market classes and other species (Kang et al., 2010; Lv et al., 2015; Obrero et al., 2013).

The differential carotenoid gene expression between 'Lady Godiva' and 'Sweet REBA' was especially pronounced for the genes known in other species to be important determinants of carotenoid synthesis and degradation. DOXP synthase, which is the first enzyme in the DOXP pathway that produces the precursors for carotenoids, is predicted to be a regulatory step for carotenoid synthesis in tomatoes

(Lois et al., 2000). DOXP synthase had up to a 3.7-fold higher expression in 'Lady Godiva', especially later in fruit development when carotenoids were accumulating to higher concentrations. Differential expression was also seen for GGPP synthase. Another important step in carotenoid synthesis is phytoene synthase, which is a rate-limiting step in the carotenoid biosynthetic pathway in marigolds, canola, and tomato (Hirschberg, 2001). Phytoene synthase was more highly expressed in 'Lady Godiva', with 3.5-fold higher expression at 40 DAP, similar to what was observed in summer squash, where phytoene synthase was much more highly expressed in an orange flesh cultivar as compared to a white flesh cultivar (Obrero et al., 2013) and in *C. maxima*, which had a higher carotenoid concentration and also higher phytoene synthase expression than *C. moschata* (Nakkanong et al., 2012a). Phytoene synthase has also been suggested to be a central enzyme in controlling carotenoid concentration in watermelon (Guo et al., 2011). Finally, carotenoid breakdown is also an important determinant of carotenoid levels in fruit tissue. Carotenoid cleavage dioxygenases were not differentially expressed in the two cultivars, and thus likely did not influence their differential carotenoid concentrations.

From lycopene, either both lycopene ε-cyclase and lycopene β-cyclase can act to make α-carotene, which is then converted to lutein, or solely lycopene β-cyclase can act to make β-carotene. In maize, natural variation in lycopene ε-cyclase influences the division of the carotenoid flux through the two branches of the pathway (Harjes et al., 2008). In our study, the higher expression of lycopene β-cyclase in 'Sweet REBA' is consistent with its higher accumulation of β-carotene and the higher expression of

lycopene ε-cyclase in ‘Lady Godiva’ is consistent with its higher accumulation of lutein.

We observed a relatively high level of expression of β-carotene hydroxylase and zeaxanthin epoxidase, even though zeaxanthin and violaxanthin did not accumulate to appreciable levels in our squash. This is likely explained by the high levels of carotenoid cleavage dioxygenase in both cultivars, which degrade carotenoids and could have kept the levels of zeaxanthin and violaxanthin low by degrading them quickly. This phenomenon was described previously in watermelon (Grassi et al., 2013). The high levels of β-carotene hydroxylase, especially in ‘Lady Godiva’ near the end of fruit development, could also explain why ‘Lady Godiva’ accumulated little β-carotene, as any β-carotene would have been made into zeaxanthin and quickly hydrolyzed by carotenoid cleavage dioxygenases.

Comparative Gene Expression in the Starch Synthesis Pathway

In early fruit development, sugars must be transported to the fruit to provide the precursors for metabolite biosynthesis as well as other developmental processes. The two primary transport sugars in squash are stachyose and raffinose. Sucrose, galactose, glucose, and fructose are also transported (Richardson et al., 1982). Alkaline and acid α-galactosidase both act to break down stachyose and raffinose, and they had high expression levels early in fruit development in our study. Similar patterns of enzyme activity were observed for alkaline and acid α-galactosidase in developing buttercup squash (Irving et al., 1997), but while that study observed a higher expression of alkaline α-galactosidase and hypothesized that it was the primary enzyme working to break down stachyose and raffinose, we observed high levels of

gene expression for both α-galactosidases with differential expression between the genotypes.

Stachyose and raffinose are broken down into sucrose and galactose, which then must be converted to glucose-6-P in order to synthesize starch. Sucrose can be broken down by either acid and neutral invertases or sucrose synthase. We observed higher levels of invertase expression in very early fruit development and higher levels of sucrose synthase expression for the remainder of development. In buttercup squash, invertase activity was also high for the beginning of fruit development and then decreased, which was suggested to provide increased osmotic potential early in development for the rapidly expanding fruit (Irving et al., 1997). In the same study, sucrose synthase activity started high and then declined, but then increased again during the ripening stage. This continued sucrose synthase activity later in fruit development was similar to what we observed. Galactose is broken down by the sequential activity of galactokinase, UDP-glucose/galactose pyrophosphorylase, and UDP-glucose epimerase, which were all expressed at a high level, especially in the early stages of fruit development. UDP-glucose pyrophosphorylase and phosphoglucomutase were also expressed at a high level, which is in accord with their role in synthesizing glucose-6-P, which is then imported into the amyloplast as the substrate for starch synthesis.

Overall, the expression of these sugar metabolism genes was generally higher in ‘Sweet REBA’, which likely acted to provide the substrate needed for ‘Sweet REBA’ to synthesize more starch than ‘Lady Godiva’. One exception was alkaline α-galactosidase, which was more highly expressed in ‘Lady Godiva’ and was

particularly highly expressed at 5 and 10 DAP. Because 'Lady Godiva' was not accumulating as much starch as 'Sweet REBA', it may have required higher levels of imported sugars for other developmental processes, such as cellulose synthesis. Sucrose synthase was more highly expressed in 'Lady Godiva', and could have also been working to supply substrate for cellulose synthesis, as previously suggested in winter squash (Irving et al., 1997).

To synthesize starch, glucose-6-P and ATP must be imported to the amyloplast by the glucose phosphate transporter and the amyloplastidial ATP/ADP translocator (AATP) respectively. AATP is a major rate-limiting enzyme for starch synthesis in potato (Ferreira et al., 2010; Geigenberger et al., 2004) and had a higher expression in 'Sweet REBA', which could contribute to the greater starch accumulation in 'Sweet REBA'. Other important regulatory steps in potato starch synthesis are plastidic phosphoglucomutase, which converts glucose-6-P imported into the amyloplast into glucose-1-P, and ADP-glucose pyrophosphorylase, which makes ADP-glucose, the first committed reaction in starch synthesis (Ferreira et al., 2010; Geigenberger et al., 2004). Both plastidic phosphoglucomutase and the large subunit of ADP-glucose pyrophosphorylase had a higher expression in 'Sweet REBA' and in another study of winter squash, ADP-glucose phosphorylase was thought to be important in determining the level of starch accumulation (Nakkanong et al., 2012b).

Several different enzymes work together to synthesize starch from ADP-glucose. Granule-bound starch synthase makes amylose, which is the un-branched form of starch. Soluble starch synthase, starch branching enzyme, and debranching enzyme work together to synthesize amylopectin, the branched form of starch. In *C.*

moschata as compared to *C. maxima*, expression of granule-bound starch synthase was positively correlated with amylose content (Nakkanong et al., 2012b). We observed higher expression of amylose-synthesizing granule-bound starch synthase in 'Lady Godiva' and higher expression of amylopectin-synthesizing soluble starch synthase and starch branching enzyme in 'Sweet REBA', which may correspond to differential ratios of amylose and amylopectin in these two cultivars.

Comparative Gene Expression in the Sucrose Synthesis Pathway

After winter squash fruit reach full size and maximum dry matter, they shift from accumulating starch to degrading starch and accumulating sucrose. During this shift, starch degradation genes increased in expression. We noted a particularly large increase in α-amylase expression and a relatively constant level of β-amylase expression, similar to enzyme activity patterns observed in buttercup squash (Irving et al., 1999). A similar developmental transition occurs in kiwifruit, where α-amylase activity increases after peak starch accumulation and soluble solid levels start to rise (Wegrzyn and Macrae, 1995). In banana, which also accumulates starch and converts it to sugar during ripening, α-amylase activity, and then α-glucosidase and β-amylase activity, coincided with the onset of ripening and starch degradation (Garcia and Lajolo, 1988) and β-amylase activity was correlated with a decrease in starch and an increase in total sugars (do Nascimento et al., 2006). These starch degradation genes were more highly expressed in 'Sweet REBA', providing more substrate for sucrose synthesis.

Starch breakdown can be phosphorolytic or hydrolytic. Hydrolytic breakdown via α-amylase and β-amylase was thought to be the primary method in squash due to

measurable α-amylase and β-amylase activity and the detection of maltose, a product of hydrolytic breakdown (Irving et al., 1999). We observed high α-amylase expression, but also high expression of starch phosphorylase, which is responsible for phosphorolytic starch breakdown. Starch phosphorylase, however, is also involved in starch synthesis, so because its expression was especially high during early fruit development, perhaps it was primarily involved in starch synthesis, leaving hydrolytic breakdown as the predominant starch breakdown pathway in squash.

The products of starch breakdown are exported from the amyloplast and then converted into fructose-6-P and UDP-glucose, the two substrates for sucrose synthesis. The sugar transporters glucose-phosphate transporter and MEX1 were more highly expressed in 'Sweet REBA', indicating that more starch breakdown products may have been leaving the amyloplast. In addition, some of the steps in the conversion to fructose-6-P and UDP-glucose were upregulated in 'Sweet REBA': α-glucosidase, hexokinase, UDP-glucose pyrophosphorylase, and phosphoglucose isomerase.

The final steps of sucrose synthesis are the sequential action of sucrose-phosphate synthase and sucrose-phosphate phosphatase to yield sucrose. The same final steps in starch-to-sugar conversion are seen in banana, where sucrose-phosphate synthase activity increases as starch is being degraded and sucrose starts to accumulate. Sucrose-phosphate synthase is a major enzyme that is associated with sugar accumulation (Huber and Huber, 1996). In melon, sucrose-phosphate synthase activity was correlated with sucrose accumulation between cultivars that accumulated different levels of sucrose (Hubbard et al., 1989). The same relationship was seen in watermelon, in which sucrose-phosphate synthase activity was higher in cultivars that

accumulated more sucrose (Yativ et al., 2010) and was also differentially expressed across fruit development (Guo et al., 2011). Sucrose-phosphate synthase and sucrose-phosphate phosphatase were both significantly differentially expressed at four of the five developmental time points sampled, with higher expression in 'Sweet REBA'.

Taken together, the identification of starch and sucrose metabolic genes and analysis of their differential gene expression demonstrates that 'Sweet REBA' accumulates more starch because of higher expression of sugar metabolism and starch synthesis genes. This higher level of stored starch, in combination with higher expression of starch breakdown and sucrose synthesis genes, then permits 'Sweet REBA' to accumulate more sucrose while maintaining a high level of starch.

Conclusion

In this study, we sequenced the fruit and seed transcriptome of an acorn squash, 'Sweet REBA', and an oilseed pumpkin, 'Lady Godiva', across fruit development. These two cultivars were developed for different purposes, fruit versus seed consumption, and had a resulting contrast in the important fruit quality traits carotenoid content, percent dry matter, and percent soluble solids. Putative metabolic pathways for the synthesis of carotenoids, starch, and sugar were assembled from the literature and expression of the structural genes of these pathways varied both across fruit development and between the two cultivars. The overall trend was increased expression of metabolic genes during times of metabolite accumulation and in the cultivar that accumulated higher metabolite levels. This demonstrates that

transcriptional regulation of metabolic gene expression is an important determinant of winter squash fruit quality.

Acknowledgments

We thank the Center for Genomics and Bioinformatics at Indiana University for their sequencing and informatics effort, Li Li and Xiangjun Zhou for their assistance with the HPLC carotenoid analysis, Suzy Strickler and Lukas Mueller for their bioinformatics expertise, and Kent Loeffler for taking the fruit pictures included in the chapter. This research was supported through funds from Cornell University. Support for Lindsay Wyatt was provided by a Cornell University Presidential Life Sciences Fellowship, USDA National Needs Graduate Fellowship Competitive Grant No. 2008-38420-04755 from the National Institute of Food and Agriculture, and the Agriculture and Food Research Initiative Competitive Grant No. 2013-67011-21122 from the USDA National Institute of Food and Agriculture.

References

Altschul, S.F., T.L. Madden, A.A. Schaffer, J.H. Zhang, Z. Zhang, W. Miller, and D.J. Lipman, 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25:3389-3402.

Apweiler, R., M.J. Martin, C. O'Donovan, M. Magrane, Y. Alam-Faruque, E. Alpi, R. Antunes, J. Arganiska, E.B. Casanova, B. Bely, M. Bingley, C. Bonilla, R. Britto, B. Bursteinas, W.M. Chan, G. Chavali, E. Cibrian-Uhalte, A. Da Silva, M. De Giorgi, E. Dimmer, F. Fazzini, P. Gane, A. Fedotov, L.G. Castro, P. Garmiri, E. Hatton-Ellis, R. Hieta, R. Huntley, J. Jacobsen, R. Jones, D. Legge, W.D. Liu, J. Luo, A. MacDougall, P. Mutowo, A. Nightingale, S. Orchard, S. Patient, K. Pichler, D. Poggioli, S. Pundir, L. Pureza, G.Y. Qi, S. Rosanoff, T. Sawford, H. Sehra, E. Turner, V. Volynkin, T. Wardell, X. Watkins, H. Zellner, M. Corbett, M. Donnelly, P. van Rensburg, M. Goujon, H. McWilliam, R. Lopez, I. Xenarios, L. Bougueleret, A. Bridge, S. Poux, N. Redaschi, A. Auchincloss, K. Axelsen, P. Bansal, D. Baratin, P.A. Binz, M.C. Blatter, B. Boeckmann, J. Bolleman, E. Boutet, L. Breuza, E. de Castro, L. Cerutti, E. Coudert, B. Cuche, M. Doche, D. Dornevil, S. Duvaud, A. Estreicher, L. Famiglietti, M. Feuermann, E. Gasteiger, S. Gehant, V. Gerritsen, A. Gos, N. Gruaz-Gumowski, U. Hinz, C. Hulo, J. James, F. Jungo, G. Keller, V. Lara, P. Lemercier, J. Lew, D. Lieberherr, X. Martin, P. Masson, A. Morgat, T. Neto, et al., 2013. Update on activities at the Universal Protein Resource (UniProt) in 2013. *Nucleic Acids Research* 41:D43-D47.

Arumuganathan, K. and E.D. Earle, 1991. Nuclear DNA Content of Some Important Plant Species. *Plant Molecular Biology Reporter* 9:208-218.

Asif, M.H., D. Lakhwani, S. Pathak, P. Gupta, S.K. Bag, P. Nath, and P.K. Trivedi, 2014. Transcriptome analysis of ripe and unripe fruit tissue of banana identifies major metabolic networks involved in fruit ripening process. *BMC Plant Biology* 14.

Azevedo-Meleiro, C.H. and D.B. Rodriguez-Amaya, 2007. Qualitative and quantitative differences in carotenoid composition among *Cucurbita moschata*, *Cucurbita maxima*, and *Cucurbita pepo*. *Journal of Agricultural and Food Chemistry* 55:4027-4033.

Blanca, J., J. Canizares, C. Roig, P. Ziarsolo, F. Nuez, and B. Pico, 2011. Transcriptome characterization and high throughput SSRs and SNPs discovery in *Cucurbita pepo* (Cucurbitaceae). *BMC Genomics* 12.

Bolger, A.M., M. Lohse, and B. Usadel, 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114-2120.

Cazzonelli, C.I. and B.J. Pogson, 2010. Source to sink: regulation of carotenoid biosynthesis in plants. *Trends in Plant Science* 15:266-274.

Conesa, A., S. Gotz, J.M. Garcia-Gomez, J. Terol, M. Talon, and M. Robles, 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21:3674-3676.

Corrigan, V.K., P.L. Hurst, and J.F. Potter, 2001. Winter squash (*Cucurbita maxima*) texture: sensory, chemical, and physical measures. *New Zealand Journal of Crop and Horticultural Science* 29:111-124.

Corrigan, V.K., D.E. Irving, and J.F. Potter, 2000. Sugars and sweetness in buttercup squash. *Food Quality and Preference* 11:313-322.

Culpepper, C.W. and H.H. Moon, 1945. Differences in the composition of the fruits of *Cucurbita* varieties at different ages in relation to culinary use. *Journal of Agricultural Research* 71:111-136.

Cumarasamy, R., V. Corrigan, P. Hurst, and M. Bendall, 2002. Cultivar differences in New Zealand "Kabocha" (buttercup squash, *Cucurbita maxima*). *New Zealand Journal of Crop and Horticultural Science* 30:197-208.

Dai, N., S. Cohen, V. Portnoy, G. Tzuri, R. Harel-Beja, M. Pompan-Lotan, N. Carmi, G. Zhang, A. Diber, S. Pollock, H. Karchi, Y. Yeselson, M. Petreikov, S. Shen, U. Sahar, R. Hovav, E. Lewinsohn, Y. Tadmor, D. Granot, R. Ophir, A. Sherman, Z. Fei, J. Giovannoni, Y. Burger, N. Katzir, and A.A. Schaffer, 2011. Metabolism of soluble sugars in developing melon fruit: a global transcriptional view of the metabolic transition to sucrose accumulation. *Plant Molecular Biology* 76:1-18.

Degu, A., U. Hochberg, N. Sikron, L. Venturini, G. Buson, R. Ghan, I. Plaschkes, A. Batushansky, V. Chalifa-Caspi, F. Mattivi, M. Delledonne, M. Pezzotti, S. Rachmilevitch, G.R. Cramer, and A. Fait, 2014. Metabolite and transcript profiling of berry skin during fruit development elucidates differential regulation between Cabernet Sauvignon and Shiraz cultivars at branching points in the polyphenol pathway. *BMC Plant Biology* 14.

do Nascimento, J.R.O., A.V. Junior, P.Z. Bassinello, B.R. Cordenunsi, J.A. Mainardi, E. Purgatto, and F.M. Lajolo, 2006. Beta-amylase expression and starch degradation during banana ripening. *Postharvest Biology and Technology* 40:41-47.

Ferreira, S.J., M. Senning, S. Sonnewald, P.M. Kessling, R. Goldstein, and U. Sonnewald, 2010. Comparative transcriptome analysis coupled to X-ray CT

reveals sucrose supply and growth velocity as major determinants of potato tuber starch biosynthesis. *BMC Genomics* 11.

Fu, L., B. Niu, Z. Zhu, S. Wu, and W. Li, 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28:3150-3152.

Garcia, E. and F.M. Lajolo, 1988. Starch transformation during banana ripening - The amylase and glucosidase behavior. *Journal of Food Science* 53:1181-1186.

Geigenberger, P., M. Stitt, and A.R. Fernie, 2004. Metabolic control analysis and regulation of the conversion of sucrose to starch in growing potato tubers. *Plant Cell and Environment* 27:655-673.

Grabherr, M.G., B.J. Haas, M. Yassour, J.Z. Levin, D.A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q.D. Zeng, Z.H. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B.W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman, and A. Regev, 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29:644-652.

Grassi, S., G. Piro, J.M. Lee, Y. Zheng, Z.J. Fei, G. Dalessandro, J.J. Giovannoni, and M.S. Lenucci, 2013. Comparative genomics reveals candidate carotenoid pathway regulators of ripening watermelon fruit. *BMC Genomics* 14.

Guo, S., J. Liu, Y. Zheng, M. Huang, H. Zhang, G. Gong, H. He, Y. Ren, S. Zhong, Z. Fei, and Y. Xu, 2011. Characterization of transcriptome dynamics during watermelon fruit development: sequencing, assembly, annotation and gene expression profiles. *BMC Genomics* 12.

Harjes, C.E., T.R. Rocheford, L. Bai, T.P. Brutnell, C.B. Kandianis, S.G. Sowinski, A.E. Stapleton, R. Vallabhaneni, M. Williams, E.T. Wurtzel, J.B. Yan, and E.S. Buckler, 2008. Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. *Science* 319:330-333.

Harvey, W.J., D.G. Grant, and J.P. Lammerink, 1997. Physical and sensory changes during the development and storage of buttercup squash. *New Zealand Journal of Crop and Horticultural Science* 25:341-351.

Hemmerich, C., A. Buechlein, R. Podicheti, K.V. Revanna, and Q.F. Dong, 2010. An Ergatis-based prokaryotic genome annotation web server. *Bioinformatics* 26:1122-1124.

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

Huang, X.Q. and A. Madan, 1999. CAP3: A DNA sequence assembly program. *Genome Research* 9:868-877.

Hubbard, N.L., S.C. Huber, and D.M. Pharr, 1989. Sucrose phosphate synthase and acid invertase as determinants of sucrose concentration in developing muskmelon (*Cucumis melo* L.) fruits. *Plant Physiology* 91:1527-1534.

Huber, S.C. and J.L. Huber, 1996. Role and regulation of sucrose-phosphate synthase in higher plants. *Annual Review of Plant Physiology and Plant Molecular Biology* 47:431-444.

Irving, D.E., P.L. Hurst, and J.S. Ragg, 1997. Changes in carbohydrates and carbohydrate metabolizing enzymes during the development, maturation, and ripening of buttercup squash (*Cucurbita maxima* D 'Delica'). *Journal of the American Society for Horticultural Science* 122:310-314.

Irving, D.E., G.J. Shingleton, and P.L. Hurst, 1999. Starch degradation in buttercup squash (*Cucurbita maxima*). *Journal of the American Society for Horticultural Science* 124:587-590.

Itle, R.A. and E.A. Kabelka, 2009. Correlation Between L*a*b* Color Space Values and Carotenoid Content in Pumpkins and Squash (*Cucurbita* spp.). *Hortscience* 44:633-637.

Kang, B.S., W.E. Zhao, Y.B. Hou, and P. Tian, 2010. Expression of carotenogenic genes during the development and ripening of watermelon fruit. *Scientia Horticulturae* 124:368-375.

Keeling, P.L. and A.M. Myers, 2010. Biochemistry and Genetics of Starch Synthesis. *Annual Review of Food Science and Technology, Vol 1* 1:271-303.

Kim, D., G. Pertea, C. Trapnell, H. Pimentel, R. Kelley, and S.L. Salzberg, 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology* 14.

Liu, S.Q., W.S. Li, Y.M. Wu, C.M. Chen, and J.J. Lei, 2013. *De Novo* Transcriptome Assembly in Chili Pepper (*Capsicum frutescens*) to Identify Genes Involved in the Biosynthesis of Capsaicinoids. *Plos One* 8.

Lois, L.M., M. Rodriguez-Concepcion, F. Gallego, N. Campos, and A. Boronat, 2000. Carotenoid biosynthesis during tomato fruit development: regulatory role of 1-deoxy-D-xylulose 5-phosphate synthase. *Plant Journal* 22:503-513.

Loy, J.B., 2004. Morpho-physiological aspects of productivity and quality in squash and pumpkins (*Cucurbita* spp.). *Critical Reviews in Plant Sciences* 23:337-363.

Lu, S. and L. Li, 2008. Carotenoid metabolism: Biosynthesis, regulation, and beyond. *Journal of Integrative Plant Biology* 50:778-785.

Lv, P., N. Li, H. Liu, H.H. Gu, and W.E. Zhao, 2015. Changes in carotenoid profiles and in the expression pattern of the genes in carotenoid metabolisms during fruit development and ripening in four watermelon cultivars. *Food Chemistry* 174:52-59.

Malone, J.G., V. Mittova, R.G. Ratcliffe, and N.J. Kruger, 2006. The response of carbohydrate metabolism in potato tubers to low temperature. *Plant and Cell Physiology* 47:1309-1322.

Martinez-Lopez, L.A., N. Ochoa-Alejo, and O. Martinez, 2014. Dynamics of the chili pepper transcriptome during fruit development. *BMC Genomics* 15.

Murkovic, M., U. Mulleder, and H. Neunteufl, 2002. Carotenoid content in different varieties of pumpkins. *Journal of Food Composition and Analysis* 15:633-638.

Nakkanong, K., J.H. Yang, and M.F. Zhang, 2012a. Carotenoid Accumulation and Carotenogenic Gene Expression during Fruit Development in Novel Interspecific Inbred Squash Lines and Their Parents. *Journal of Agricultural and Food Chemistry* 60:5936-5944.

Nakkanong, K., J.H. Yang, and M.F. Zhang, 2012b. Starch accumulation and starch related genes expression in novel inter-specific inbred squash line and their parents during fruit development. *Scientia Horticulturae* 136:1-8.

Obrero, A., C.I. Gonzalez-Verdejo, J.V. Die, P. Gomez, M. Del Rio-Celestino, and B. Roman, 2013. Carotenogenic Gene Expression and Carotenoid Accumulation in Three Varieties of *Cucurbita pepo* during Fruit Development. *Journal of Agricultural and Food Chemistry* 61:6393-6403.

Richardson, P.T., D.A. Baker, and L.C. Ho, 1982. The chemical composition of cucurbit vascular exudates. *Journal of Experimental Botany* 33:1239-1247.

Schulz, M.H., D.R. Zerbino, M. Vingron, and E. Birney, 2012. Oases: robust *de novo* RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics* 28:1086-1092.

Sharkey, T.D., M. Laporte, Y. Lu, S. Weise, and A.P.M. Weber, 2004. Engineering plants for elevated CO2: A relationship between starch degradation and sugar sensing. *Plant Biology* 6:280-288.

Smith, A.M., 2012. Starch in the Arabidopsis plant. *Starch-Starke* 64:421-434.

Tjaden, J., T. Mohlmann, K. Kampfenkel, G. Henrichs, and H.E. Neuhaus, 1998. Altered plastidic ATP/ADP-transporter activity influences potato (*Solanum tuberosum* L.) tuber morphology, yield and composition of tuber starch. *Plant Journal* 16:531-540.

Trapnell, C., A. Roberts, L. Goff, G. Pertea, D. Kim, D.R. Kelley, H. Pimentel, S.L. Salzberg, J.L. Rinn, and L. Pachter, 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols* 7:562-578.

Van Eck, J., X. Zhou, S. Lu, and L. Li, 2010. Modulation of Carotenoid Accumulation in Transgenic Potato by Inducing Chromoplast Formation with Enhanced Sink Strength. In: A. Fett-Neto (ed.), Plant Secondary Metabolism Engineering.

Warnes, G.R., B. Bolker, L. Bonebakker, R. Gentleman, W.H.A. Liaw, T. Lumley, M. Maechler, A. Magnusson, S. Moeller, M. Schwartz, and B. Venables, 2015. gplots: Various R Programming Tools for Plotting Data. R package version 2.16.0.

Weber, A.P.M., 2004. Solute transporters as connecting elements between cytosol and plastid stroma. *Current Opinion in Plant Biology* 7:247-253.

Wegrzyn, T. and E. Macrae, 1995. Alpha-amylase and starch degradation in kiwifruit. *Journal of Plant Physiology* 147:19-28.

Wu, T., S. Luo, R. Wang, Y. Zhong, X. Xu, Y.e. Lin, X. He, B. Sun, and H. Huang, 2014. The first Illumina-based de novo transcriptome sequencing and analysis of pumpkin (*Cucurbita moschata* Duch.) and SSR marker development. *Molecular Breeding* 34:1437-1447.

Wyatt, L.E., S.R. Strickler, L.A. Mueller, and M. Mazourek, 2015. An acorn squash (*Cucurbita pepo* ssp. *ovifera*) fruit and seed transcriptome as a resource for the study of fruit traits in *Cucurbita*. *Horticulture Research* 2:1-7.

Yativ, M., I. Harary, and S. Wolf, 2010. Sucrose accumulation in watermelon fruits: Genetic variation and biochemical analysis. *Journal of Plant Physiology* 167:589-596.

Zdobnov, E.M. and R. Apweiler, 2001. InterProScan - an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* 17:847-848.

Zerbino, D.R. and E. Birney, 2008. Velvet: Algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Research* 18:821-829.

Zhang, M.K., M.P. Zhang, M. Mazourek, Y. Tadmor, and L. Li, 2014. Regulatory control of carotenoid accumulation in winter squash during storage. *Planta* 240:1063-1074.

CONCLUSION

This dissertation describes genomics research and applied breeding activities focused on improved fruit quality and disease resistance in pepper and squash. The molecular markers described in Chapter 2 are already being employed by pepper breeders to select for the *Pun1* gene that controls pungency, thus increasing breeding efficiency. The *Phytophthora capsici*-resistant bell pepper inbred lines reported in Chapter 3 have excellent disease resistance and also perform well as hybrid parents, meeting a current need for resistant peppers in the Northeast. Chapters 4 and 5 examine squash fruit quality from the genomics level, looking at the genes expressed in developing fruit and seed and investigating the gene expression differences underlying important culinary properties. The breadth of the research undertaken encompasses the spectrum of approaches that are required to develop successful cultivars. The applied pepper work will immediately benefit vegetable growers and breeders, and the more basic squash transcriptome research will provide the foundation for additional work to improve fruit quality.

There are many ways to build off of the research in this dissertation to continue to make gains in vegetable disease resistance and fruit quality. The study of pungency in pepper will benefit from additional analysis of the genetic basis of quantitative variation in capsaicinoid levels. Most QTL studies of pungency have relied on interspecific populations, but with the advent of genotyping-by-sequencing technology and the completion of the pepper genome it will be possible to map QTLs in a *C. annuum* x *C. annuum* population. Using the pepper genome, candidate genes can potentially be identified for major QTLs, resulting in a deeper understanding of the

regulation of capsaicinoid biosynthesis and enabling the development of molecular markers to use in selecting for high or low pungency levels when breeding *C. annuum*.

We are currently continuing our *P. capsici* breeding work by making crosses to improve the earliness and fruit size of our *P. capsici*-resistant bell peppers, as well as transfer the resistance to a variety of other fruit types. In addition to selecting highly resistant peppers with improved horticultural characteristics, this work provides the opportunity to examine the genetics of *P. capsici* resistance in our inbred lines. By analyzing the distribution of resistance phenotypes in F_2 populations, it will be possible to determine whether a major R gene or many quantitative resistance genes control resistance in the inbred lines. These F_2 populations can also be genotyped using genotyping-by-sequencing to map QTLs for resistance. Knowledge of the genetic basis of resistance will inform continued breeding efforts to combine high levels of resistance with larger fruit and earlier ripening and will help to predict the durability of the resistance.

The winter squash fruit and seed transcriptome provides an excellent resource for further study of fruit quality and fruit and seed development. Beyond identifying the structural genes involved in fruit metabolite biosynthesis, the data set as a whole can be analyzed to determine which genes are differentially expressed between the two cultivars and across development. Genes can also be clustered based on their expression patterns and clusters of genes expressed in the same patterns as the metabolite structural genes can be further examined. Transcription factors can then be identified and any transcription factors with sequence or expression differences between the two cultivars can be flagged for a more in-depth analysis. It will be

especially interesting to combine this gene expression analysis with association mapping to locate the genetic loci responsible for the observed variation in gene expression.

The seed transcriptome data will enable the study of the hull-less gene and aid in its identification. Preliminary work towards this objective is described in Appendix A. Briefly, homologs of secondary cell wall biosynthesis genes and secondary cell wall-related MYB transcription factors were identified and their observed gene expression supports the hypothesis that the hull-less gene may be a NAC transcription factor acting as a tissue-specific master level regulator of secondary cell wall biosynthesis. This gene expression data can now be combined with association mapping data from our hull-less breeding populations to identify the regions consistently introgressed along with the hull-less phenotype and then further select candidate genes within those regions.

Overall, it will be important to continue the study of vegetable fruit quality and disease resistance, as these two aspects of vegetable breeding are both essential to support the sustainable production of vegetables and promote their increased consumption.

APPENDIX A

TRANSCRIPTOMIC ANALYSIS OF THE HULL-LESS GENE IN SQUASH

Hull-less Pumpkin Seeds

Winter squash is a widely consumed vegetable crop in the United States. Squash was domesticated for consumption of its nutritious seeds (Robinson and Decker-Walters, 1997), but seeds are currently a seldom-used part of the squash fruit. A primary reason for the under-utilization of squash seeds is their tough, unpalatable seed coat. There is a major, recessive, qualitative mutation, with quantitative modifiers, that causes the seed coat to be greatly reduced (Stuart and Loy, 1983) and has great potential for increasing consumption of squash seeds.

This mutation, known as the hull-less mutation, causes a dramatic reduction in the development of the seed coat. All of the layers of the seed coat are present at the beginning of seed coat development, but the outer layers of the seed coat do not undergo normal secondary cell wall development and thus have greatly reduced levels of lignin, cellulose, and hemicellulose (Stuart, 1981; Stuart and Loy, 1983).

Regulation of Secondary Cell Wall Development

The regulation of tissue-specific secondary cell wall development has been described in many species such as Arabidopsis, poplar, rice, and Medicago (Mitsuda et al., 2005; Wang et al., 2011; Zhong et al., 2010b; Zhong et al., 2011). A group of NAC transcription factors act as tissue-specific master regulators, activating the same set of secondary-level MYB transcription factors in all tissues (Zhong et al., 2008). These MYBs then activate a set of secondary cell wall biosynthesis genes, which synthesize the cellulose, hemicellulose, and lignin of the secondary cell wall (Zhong et

al., 2010a). We hypothesize that the hull-less mutation affects a NAC transcription factor controlling secondary cell wall development in the seed coat. In support of this hypothesis, four genes involved in secondary cell wall biosynthesis were found to be down-regulated in seed coats of hull-less mutants (Bezold et al., 2005), indicating that the entire set of biosynthesis genes may be down-regulated as would be the case if the NAC master switch were mutated.

Complementation Test to Compare Multiple Sources of the Hull-less Trait

To determine if there were one or multiple genes causing the hull-less phenotype in *C. pepo*, we conducted a complementation test combining six different sources of the trait (PMR Naked Seeded Pumpkin, Lady Godiva, Austria's Finest, Kakai, PI 615133, and PI 531323). These six lines were crossed to each other in all combinations in summer 2011 in Freeville, NY and the F_1 plants were grown in summer 2012 in Freeville, NY. All F_1 plants displayed the hull-less phenotype, indicating that they all had mutations in the same major gene causing their hull-less phenotype.

Squash Comparative Transcriptome Analysis

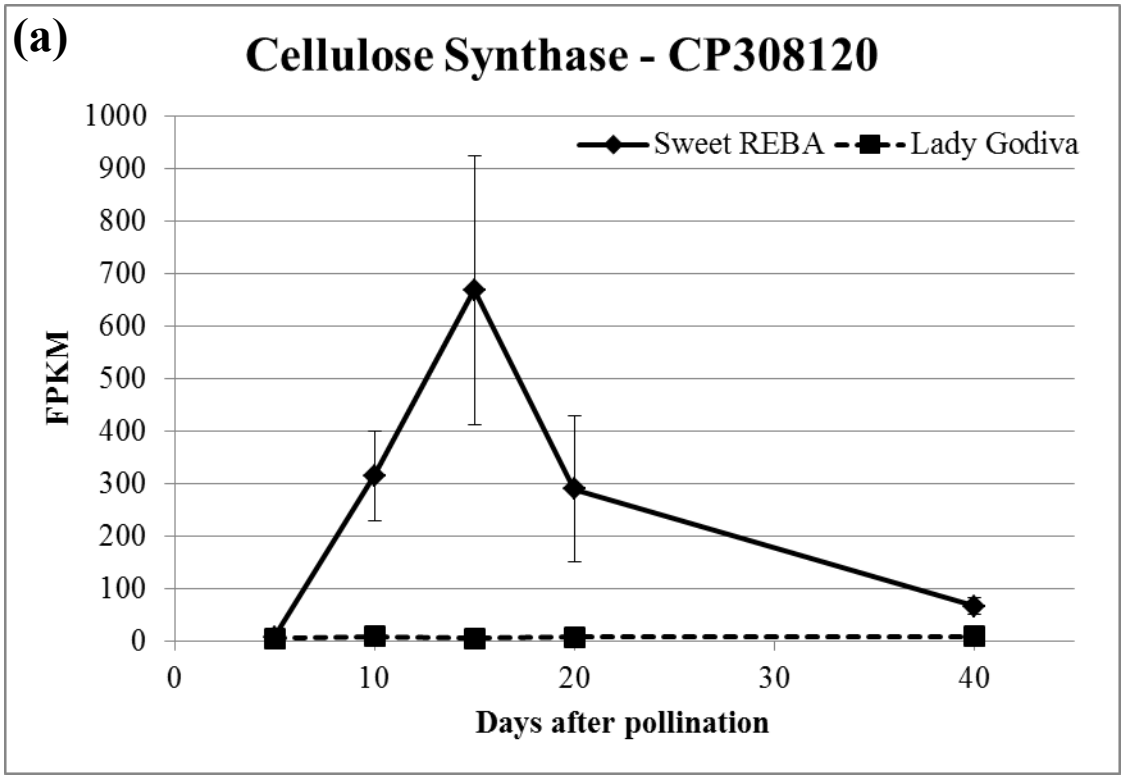
We performed transcriptome sequencing to examine differential gene expression in the regulatory and biosynthetic networks underlying secondary cell wall biosynthesis throughout seed development and between cultivars with and without the hull-less mutation. Two contrasting *Cucurbita pepo* cultivars were used for RNA-seq analysis: 'Lady Godiva,' an oilseed pumpkin with poor eating quality, and 'Sweet Reba,' an acorn squash with high eating quality.

Self-pollinated fruits were collected at five time points and fruit mesocarp and seeds were flash-frozen in liquid nitrogen for RNA extraction. Mesocarp and seed mRNA from each time point were extracted using a Qiagen RNeasy kit and sequenced in three biological replicates, each consisting of two pooled samples from different fruit. Library preparation, sequencing, and bioinformatics were performed by the Center for Genomics and Bioinformatics at Indiana University. An Illumina TruSeq Sample Prep Kit was used for library preparation and barcoded samples were sequenced on two lanes of an Illumina HiSeq 2000. Read quality and adapter trimming were performed using the software Trimmomatic (Bolger et al., 2014). The trimmed paired reads and singleton reads were normalized using Trinity's *in silico* normalization utility, normalize_by_kmer_coverage.pl (Grabherr et al., 2011). Following normalization, Velvet (Zerbino and Birney, 2008) and Oases (Schulz et al., 2012) were used to assemble the transcriptome and the longest unigene was chosen to represent each isogroup. This transcriptome was then combined with the published acorn squash fruit and seed transcriptome (Wyatt et al., 2015) and the resulting set of unigenes was clustered using CD-HIT (Fu et al., 2012), further assembled using CAP3 (Huang and Madan, 1999) and clustered using self-blast to group alternatively-spliced transcripts, with the longest unigene chosen to represent each cluster in the final transcriptome. After assembly, reads were mapped to the transcriptome using TopHat2 (Kim et al., 2013) and FPKM (fragments per kilobase of exon per million reads mapped) values were subsequently calculated using Cuffdiff (Trapnell et al., 2012). Secondary cell wall biosynthesis gene homologs and candidate secondary cell

wall-related MYB transcription factors were identified using BLAST (Altschul et al., 1997).

We identified putative homologs for the four indicator genes commonly used to detect regulatory changes in the secondary cell wall biosynthesis pathway (Figure A.1): cellulose synthase, FRA8 (involved in hemicellulose synthesis), 4-coumarate-CoA ligase, and caffeoyl CoA *O*-methyltransferase (both involved in lignin synthesis). These all showed a dramatic decrease in expression in the hull-less cultivar Lady Godiva, supporting the hypothesis that the hull-less gene is a regulatory gene controlling secondary cell wall biosynthesis in the seed coat. In addition, another hemicellulose synthesis gene (IRX9) and multiple lignin synthesis genes (phenylalanine ammonia-lyase, cinnamate 4-hydroxylase, caffeic acid *O*-methyltransferase, and cinnamyl alcohol dehydrogenase) were also identified and displayed the same decrease in expression in the hull-less cultivar (Figure A.2), further supporting our hypothesis. Four putative secondary cell wall-related MYB transcription factors were also identified (Figure A.3), which were also down-regulated in the hull-less cultivar, which provides additional support that the hull-less gene is a NAC transcription factor controlling secondary cell wall development in the seed coat via secondary-level MYB transcription factors. We were able to identify 80 candidate NAC transcription factors, but there was not sufficient homology to identify any specific candidates. The identification of the actual hull-less gene, aided by knowledge of the NACs expressed in the seed coat during development, will be an important topic for future study.

Figure A.1 Expression of secondary cell wall biosynthesis indicator genes. Multiple homologs were identified for secondary cell wall indicator genes. Graphs display the gene expression of the unigene(s) with both high expression levels and a higher expression in 'Sweet REBA'. Indicator genes homologs were identified for cellulose synthesis **(a-b)**, hemicellulose synthesis **(c)**, and lignin synthesis **(d-f)**. FPKM gene expression values of both genotypes are displayed over the five developmental time points, averaged from three biological replicates. Error bars indicate the standard deviation.


(a)
Cellulose Synthase - CP308120
Sweet REBA
Lady Godiva
FPKM
1000
900
800
700
600
500
400
300
200
100
0
0
10
20
30
40
Days after pollination

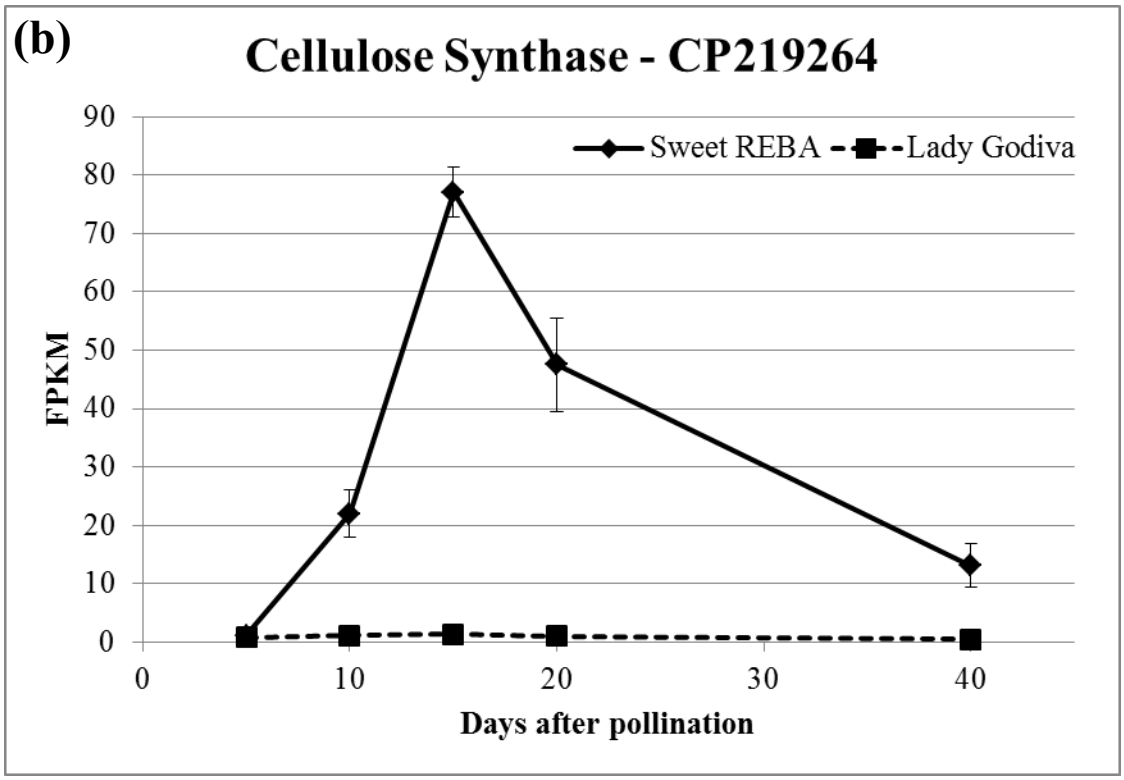
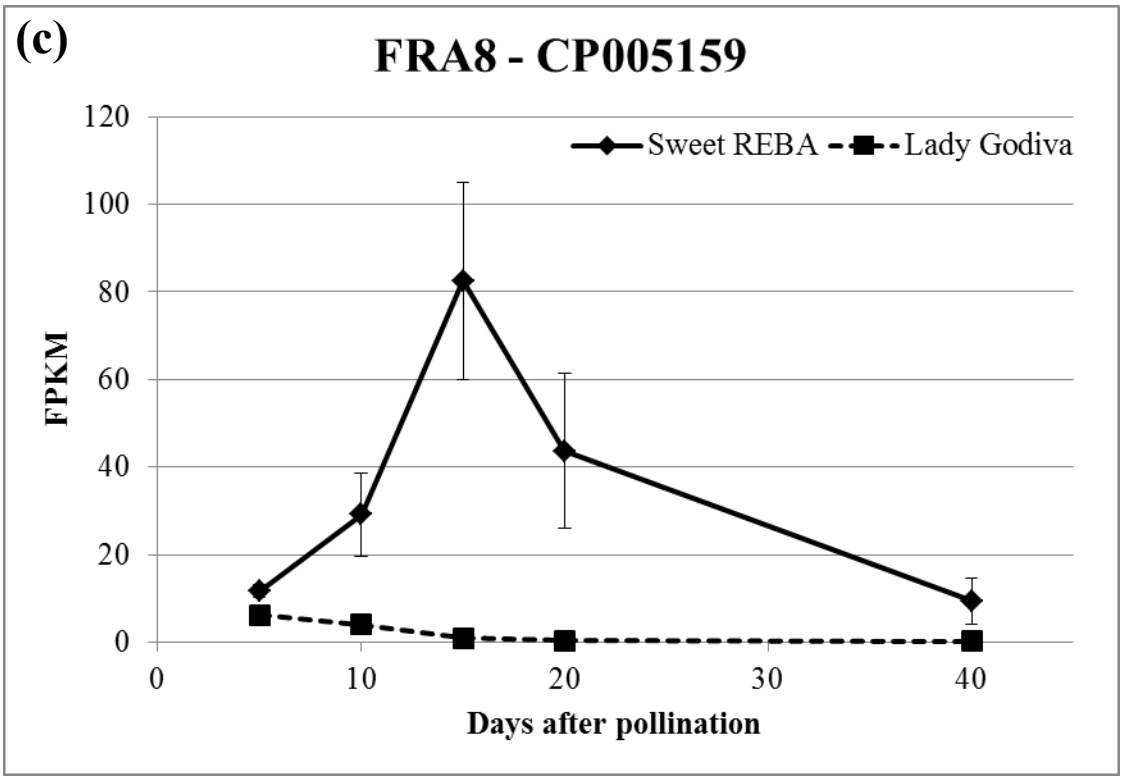

(b)
Cellulose Synthase - CP219264
Sweet REBA
Lady Godiva
FPKM
90
80
70
60
50
40
30
20
10
0
0
10
20
30
40
Days after pollination

Figure A.1 (Continued)



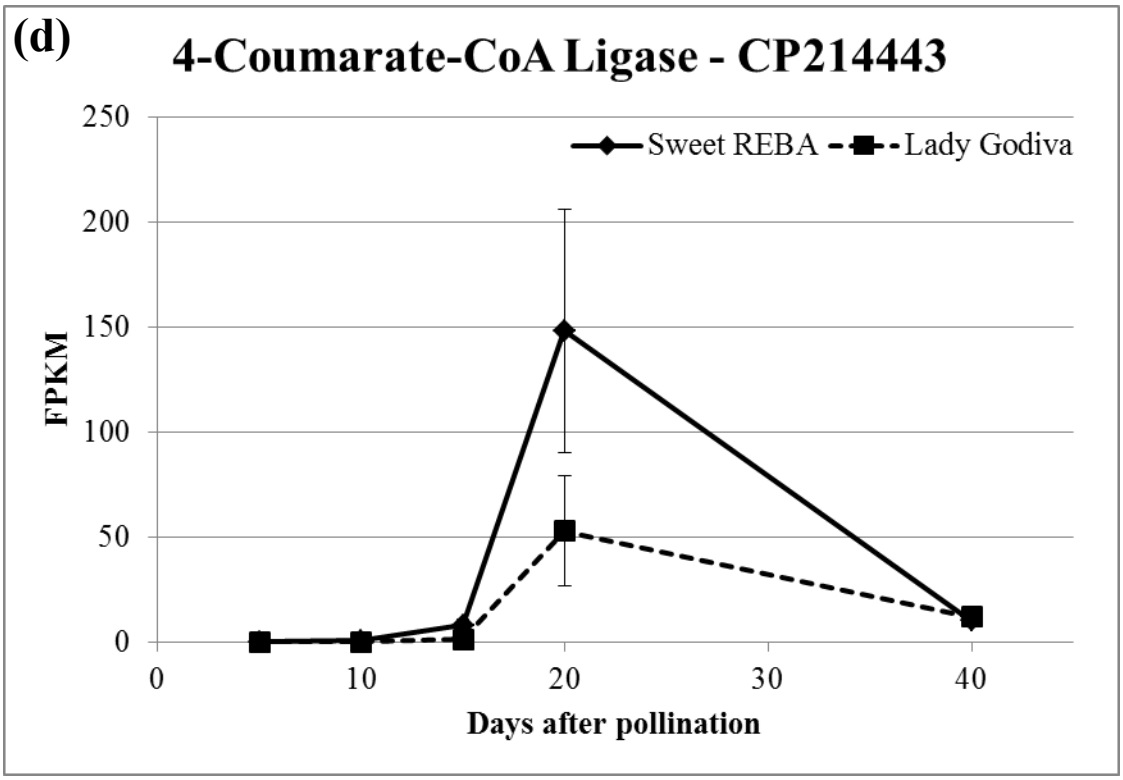
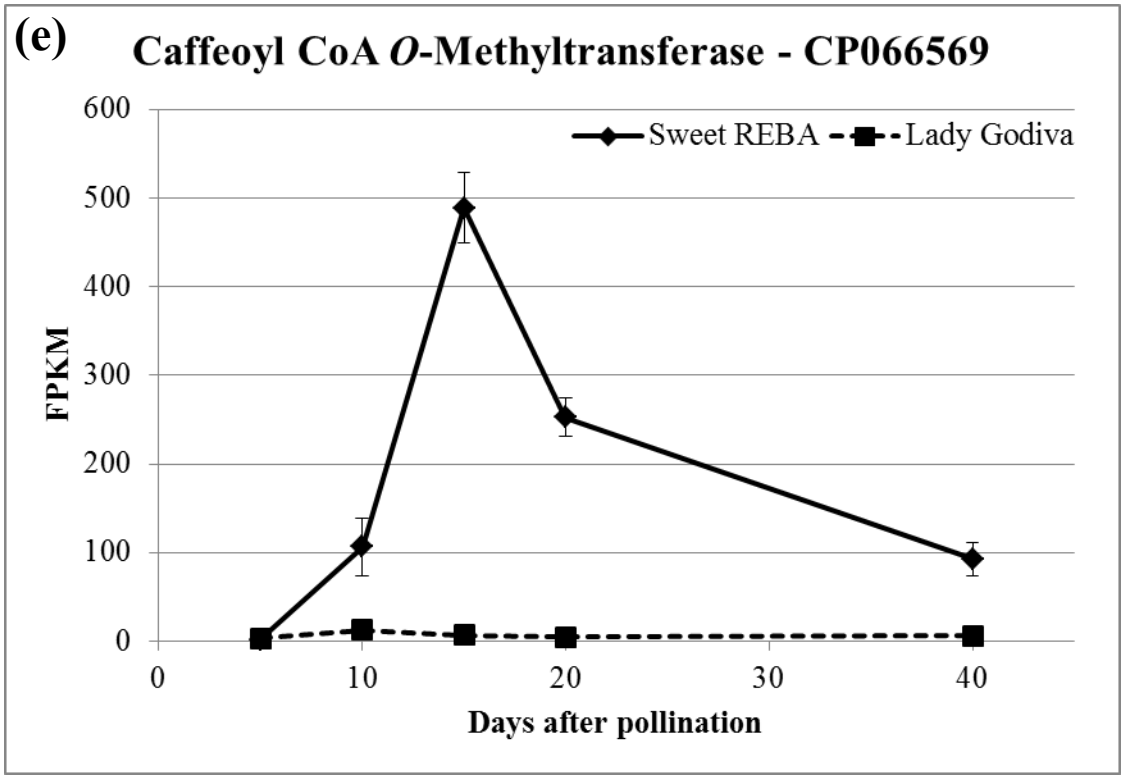


Figure A.1 (Continued)



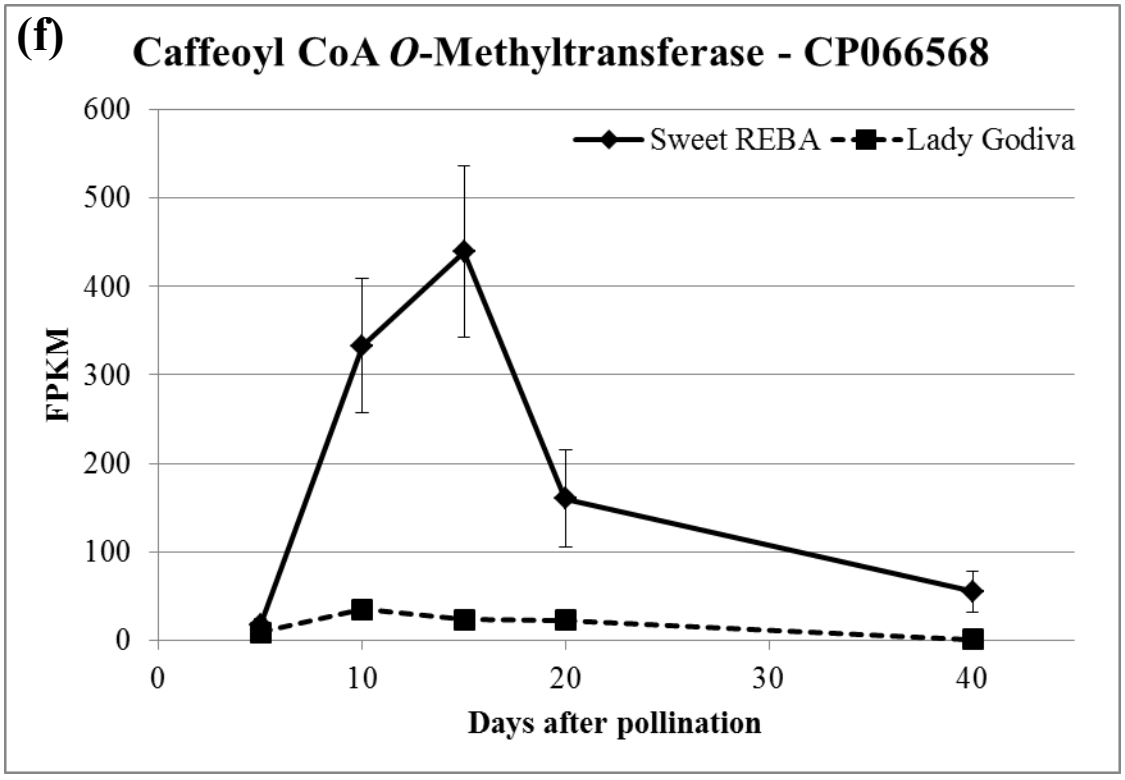
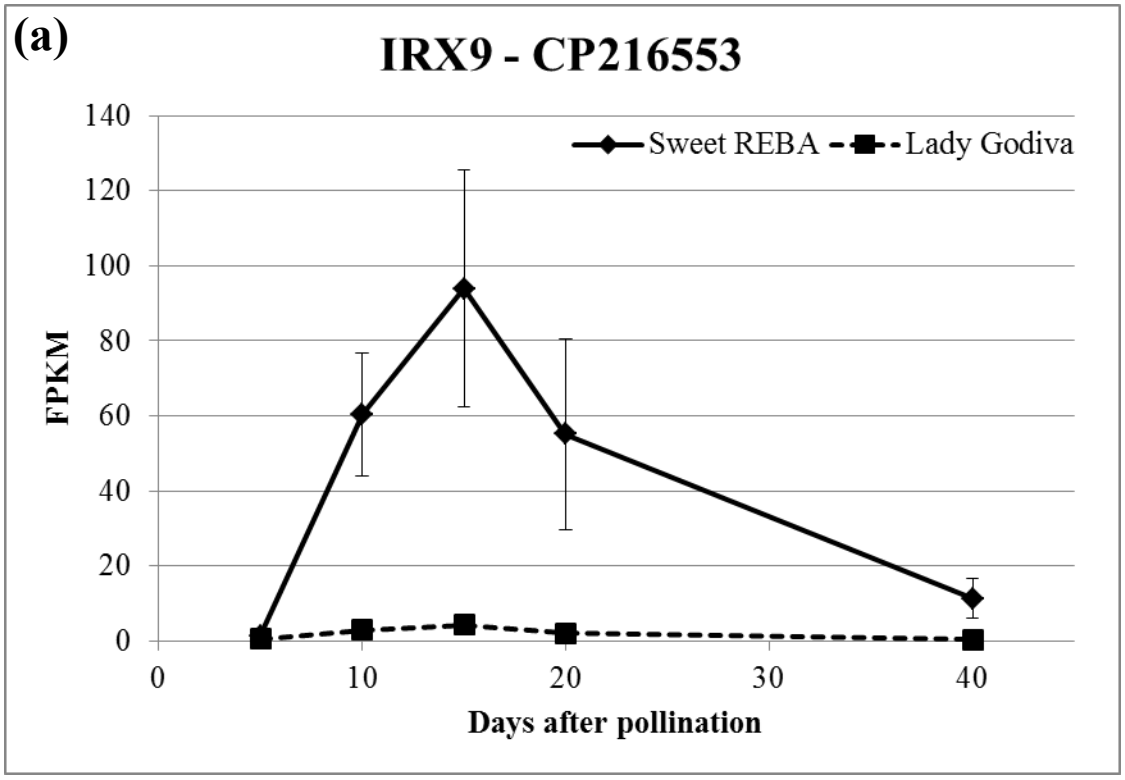


Figure A.2 Expression of additional secondary cell wall biosynthesis genes. Multiple homologs were identified for additional secondary cell wall biosynthesis genes. Graphs display the gene expression of the unigene(s) with both high expression levels and a higher expression in 'Sweet REBA'. Unigenes were identified that are involved in both hemicellulose synthesis **(a)** and lignin synthesis **(b-g)**. FPKM gene expression values of both genotypes are displayed over the five developmental time points, averaged from three biological replicates. Error bars indicate the standard deviation.


(a)
IRX9 - CP216553
Sweet REBA
Lady Godiva
FPKM
140
120
100
80
60
40
20
0
0
10
20
30
40
Days after pollination

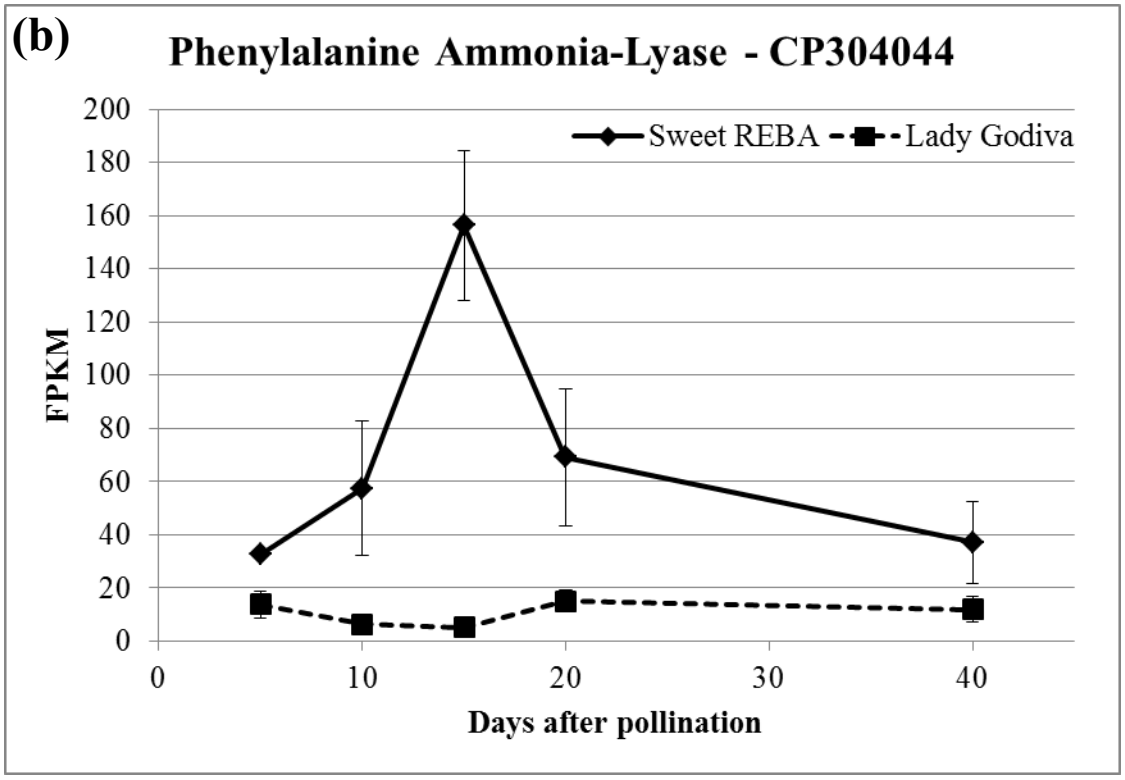
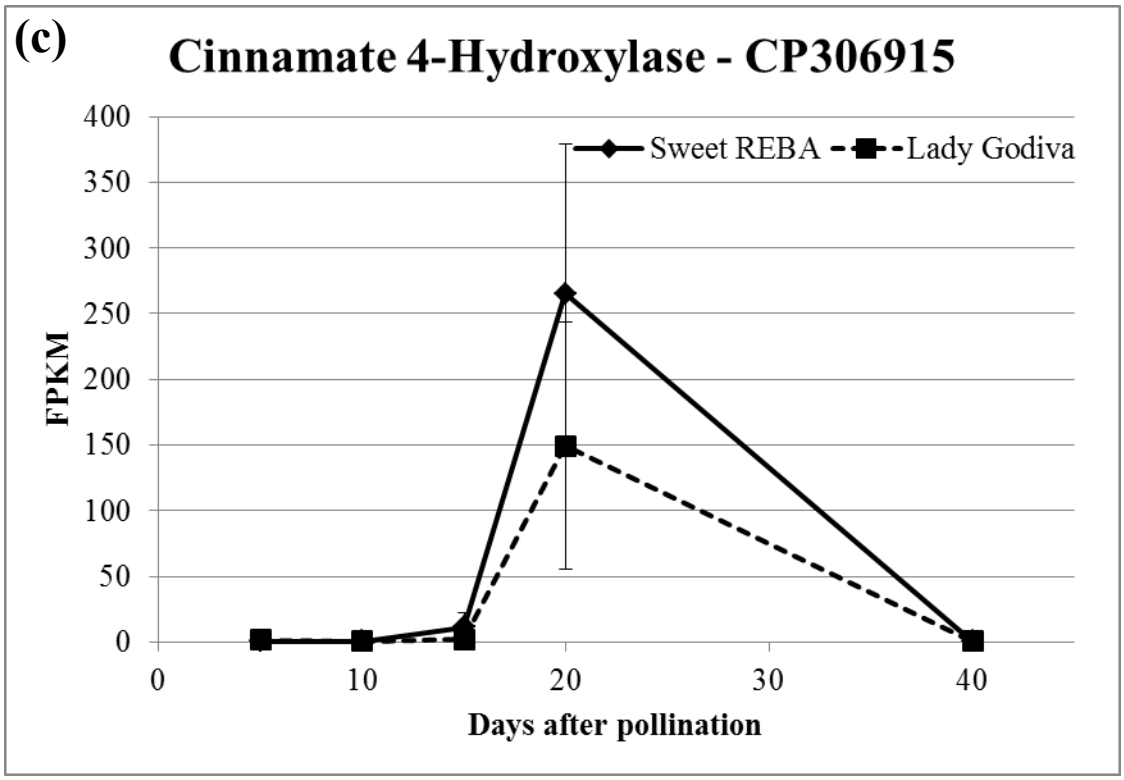

(b)
Phenylalanine Ammonia-Lyase - CP304044
Sweet REBA
Lady Godiva
FPKM
200
180
160
140
120
100
80
60
40
20
0
0
10
20
30
40
Days after pollination

Figure A.2 (Continued)



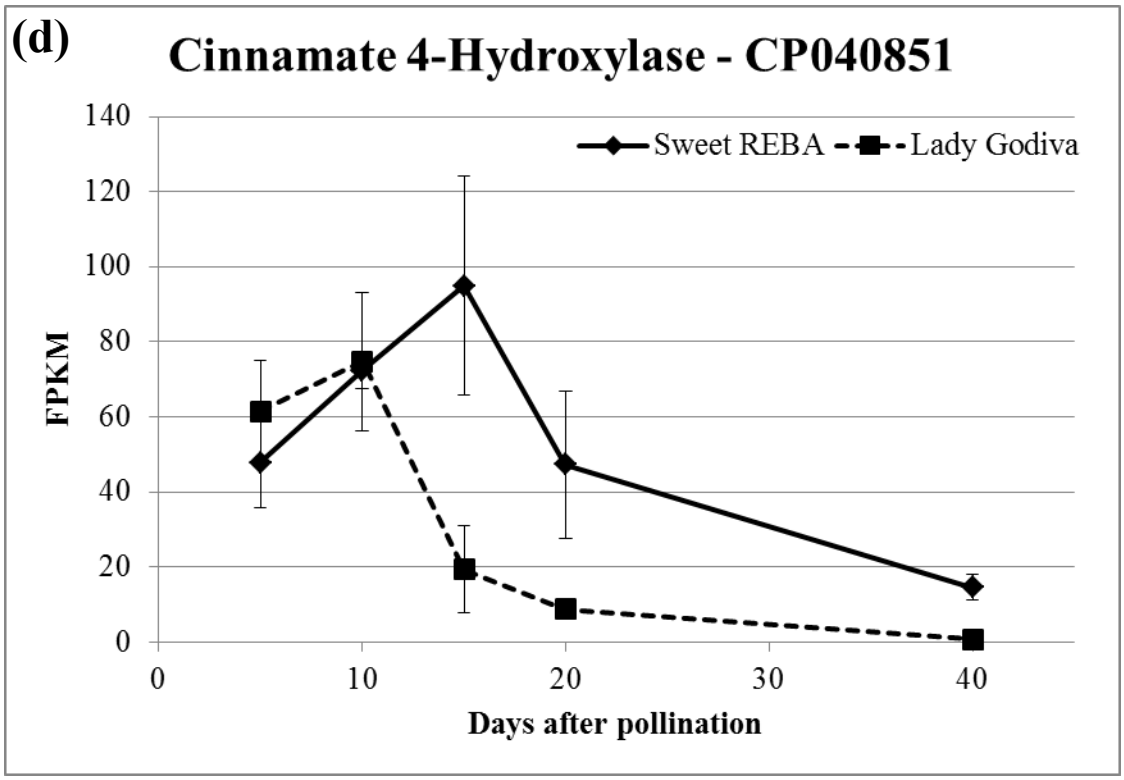
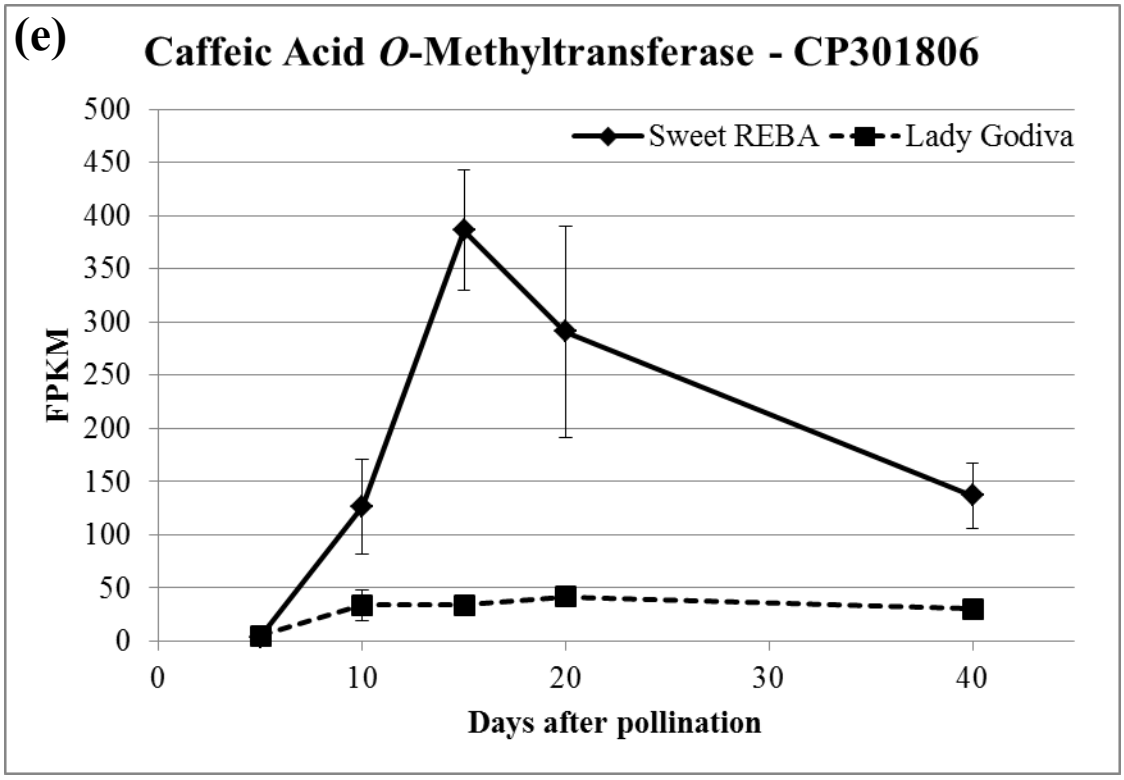


Figure A.2 (Continued)



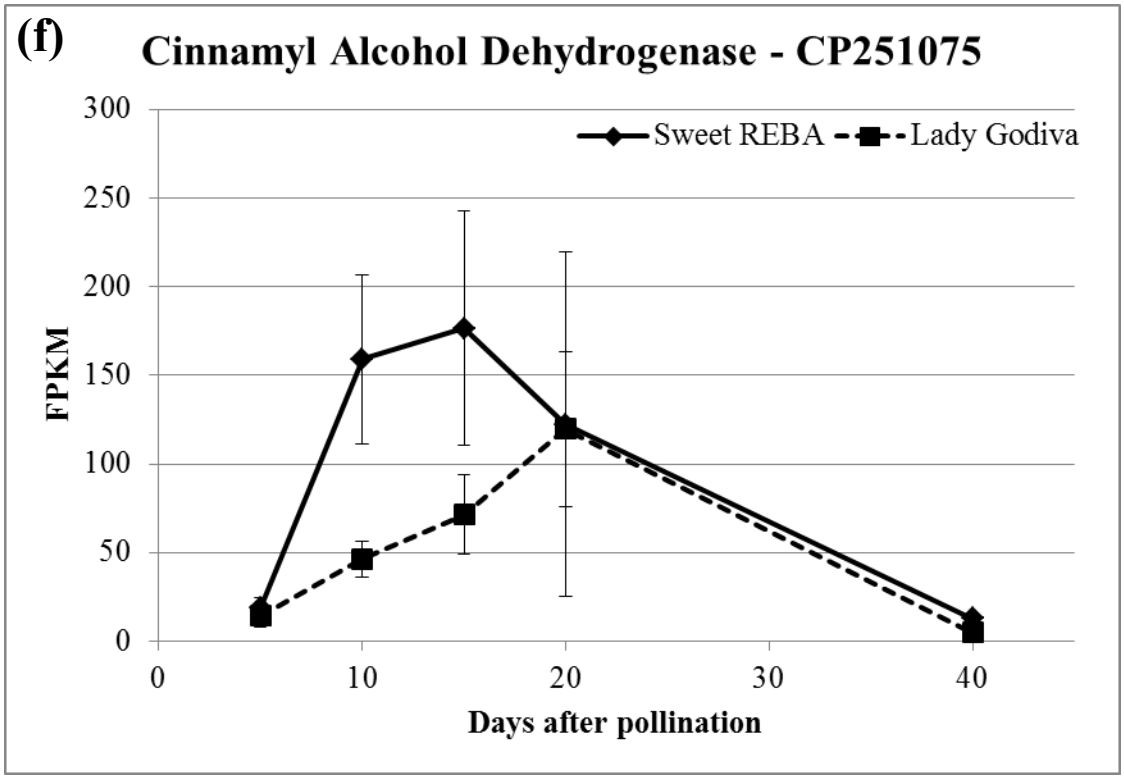


Figure A.2 (Continued)

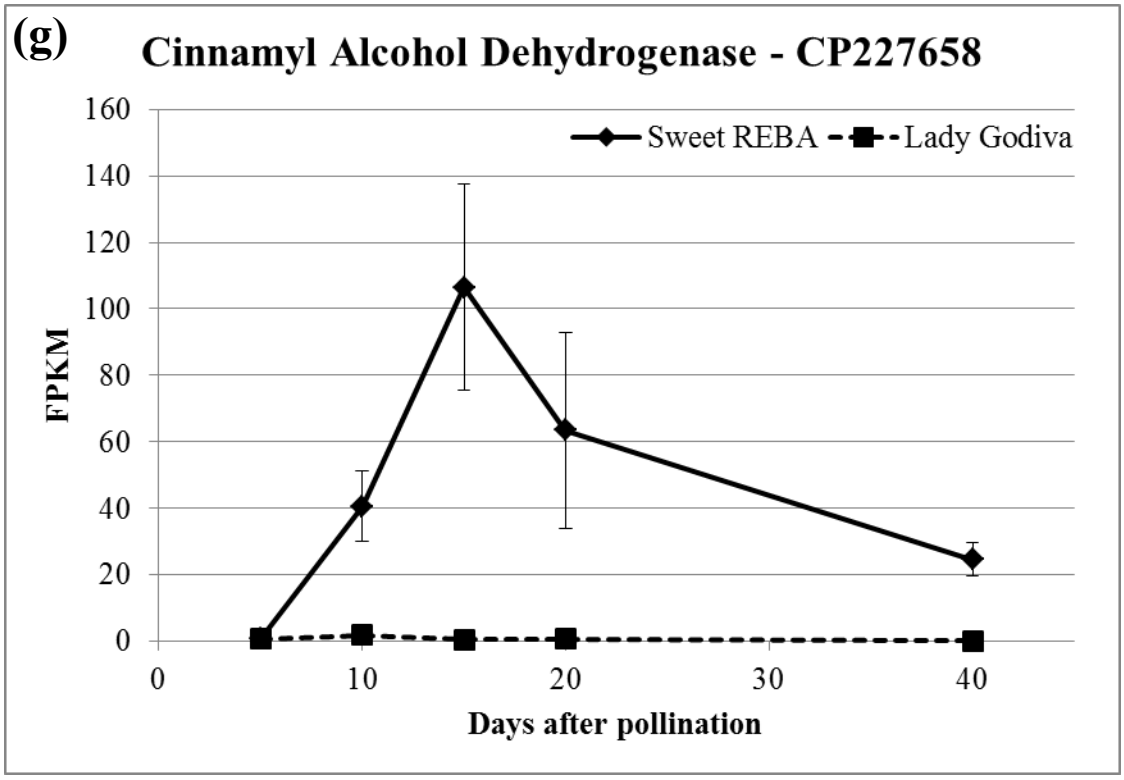
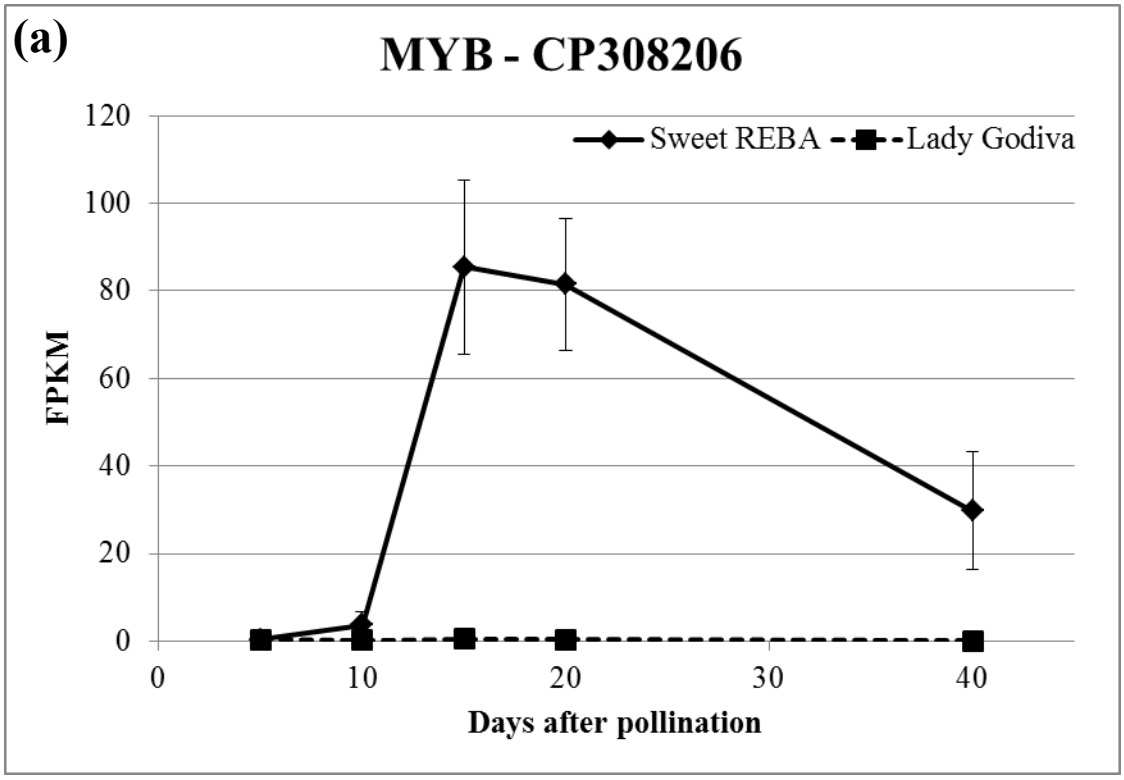


Figure A.3 Expression of secondary cell wall-related MYB transcription factors. Four candidate secondary cell wall-related MYB transcription factors were identified **(a-d)**. FPKM gene expression values of both genotypes are displayed over the five developmental time points, averaged from three biological replicates. Error bars indicate the standard deviation.


(a)
MYB - CP308206
Sweet REBA
Lady Godiva
FPKM
120
100
80
60
40
20
0
0
10
20
30
40
Days after pollination

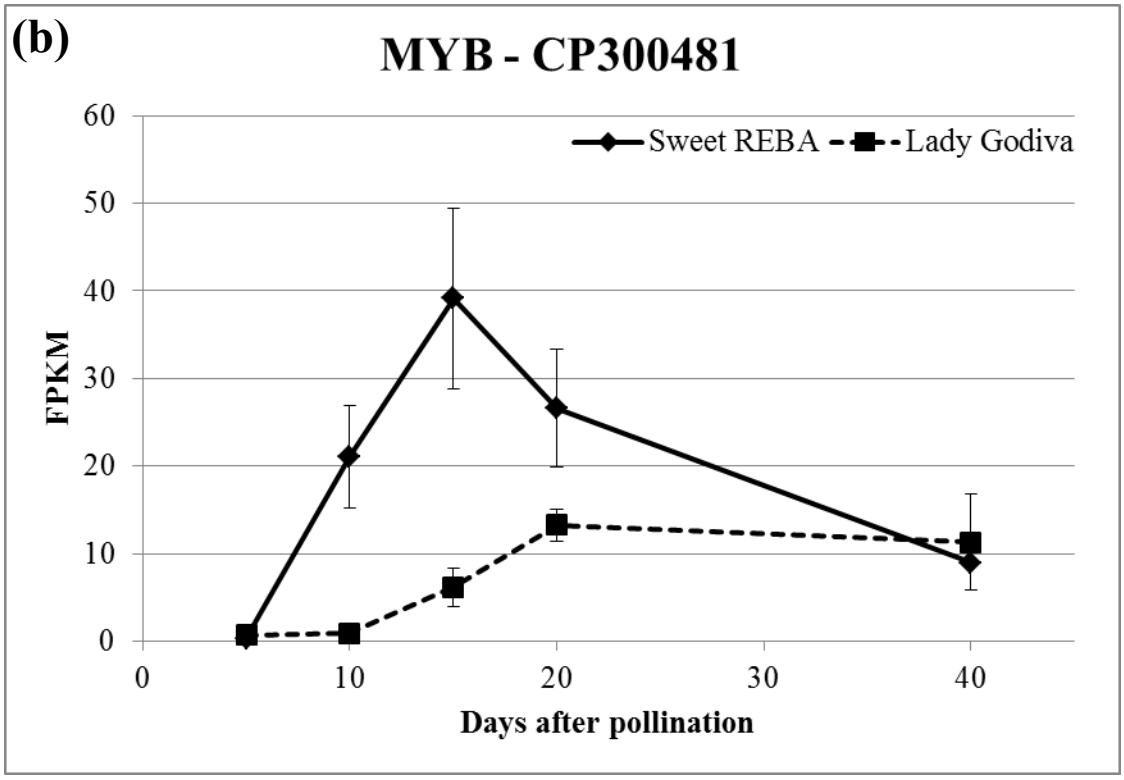
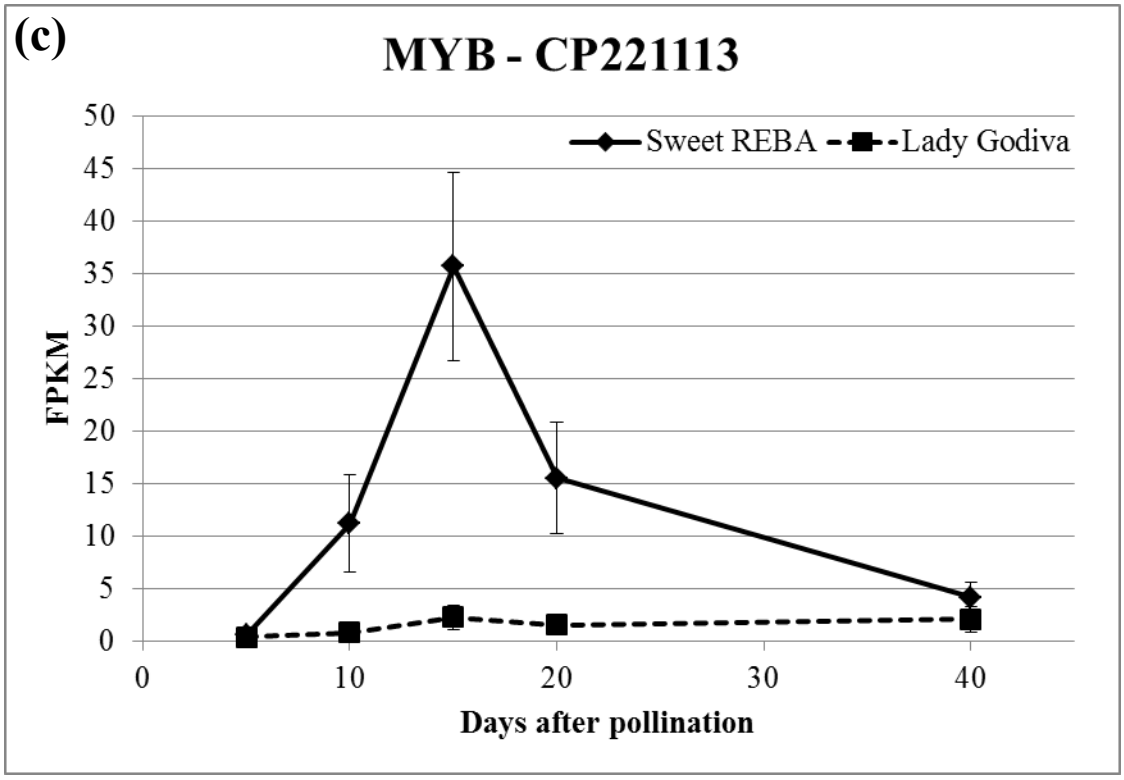
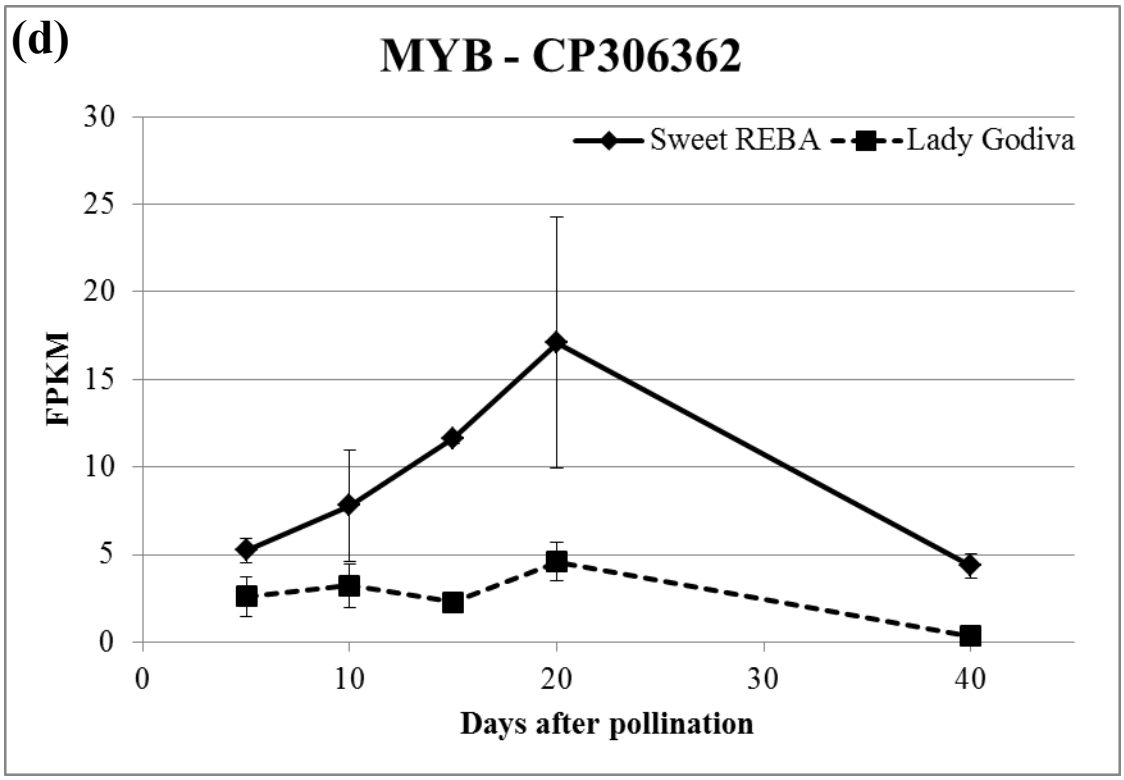

(b)
MYB - CP300481
Sweet REBA
Lady Godiva
FPKM
60
50
40
30
20
10
0
0
10
20
30
40
Days after pollination

Figure A.3 (Continued)





References

Altschul, S.F., T.L. Madden, A.A. Schaffer, J.H. Zhang, Z. Zhang, W. Miller, and D.J. Lipman, 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25:3389-3402.

Bezold, T.N., D. Mathews, J.B. Loy, and S.C. Minocha, 2005. Molecular analysis of the hull-less seed trait in pumpkin: expression profiles of genes related to seed coat development. *Seed Science Research* 15:205-217.

Bolger, A.M., M. Lohse, and B. Usadel, 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114-2120.

Fu, L., B. Niu, Z. Zhu, S. Wu, and W. Li, 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28:3150-3152.

Grabherr, M.G., B.J. Haas, M. Yassour, J.Z. Levin, D.A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q.D. Zeng, Z.H. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B.W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman, and A. Regev, 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29:644-652.

Huang, X.Q. and A. Madan, 1999. CAP3: A DNA sequence assembly program. *Genome Research* 9:868-877.

Kim, D., G. Pertea, C. Trapnell, H. Pimentel, R. Kelley, and S.L. Salzberg, 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology* 14:1-13.

Mitsuda, N., M. Seki, K. Shinozaki, and M. Ohme-Takagi, 2005. The NAC transcription factors NST1 and NST2 of Arabidopsis regulate secondary wall thickenings and are required for anther dehiscence. *Plant Cell* 17:2993-3006.

Robinson, R.W. and D.S. Decker-Walters, 1997. Cucurbits. CAB International, New York.

Schulz, M.H., D.R. Zerbino, M. Vingron, and E. Birney, 2012. Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics* 28:1086-1092.

Stuart, S.G., 1981. Comparative studies of testa development in normal and hull-less seed strains of *Cucurbita pepo* L. University of New Hampshire, MS thesis.

Stuart, S.G. and J.B. Loy, 1983. Comparison of testa development in normal and hull-less seeded strains of *Cucurbita pepo* L. *Botanical Gazette* 144:491-500.

Trapnell, C., A. Roberts, L. Goff, G. Pertea, D. Kim, D.R. Kelley, H. Pimentel, S.L. Salzberg, J.L. Rinn, and L. Pachter, 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols* 7:562-578.

Wang, H.Z., Q. Zhao, F. Chen, M.Y. Wang, and R.A. Dixon, 2011. NAC domain function and transcriptional control of a secondary cell wall master switch. *Plant Journal* 68:1104-1114.

Wyatt, L.E., S.R. Strickler, L.A. Mueller, and M. Mazourek, 2015. An acorn squash (*Cucurbita pepo* ssp. *ovifera*) fruit and seed transcriptome as a resource for the study of fruit traits in *Cucurbita*. *Horticulture Research* 2:1-7.

Zerbino, D.R. and E. Birney, 2008. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research* 18:821-829.

Zhong, R.Q., C. Lee, R.L. McCarthy, C.K. Reeves, E.G. Jones, and Z.H. Ye, 2011. Transcriptional Activation of Secondary Wall Biosynthesis by Rice and Maize NAC and MYB Transcription Factors. *Plant and Cell Physiology* 52:1856-1871.

Zhong, R.Q., C.H. Lee, and Z.H. Ye, 2010a. Evolutionary conservation of the transcriptional network regulating secondary cell wall biosynthesis. *Trends in Plant Science* 15:625-632.

Zhong, R.Q., C.H. Lee, and Z.H. Ye, 2010b. Functional Characterization of Poplar Wood-Associated NAC Domain Transcription Factors. *Plant Physiology* 152:1044-1055.

Zhong, R.Q., C.H. Lee, J.L. Zhou, R.L. McCarthy, and Z.H. Ye, 2008. A Battery of Transcription Factors Involved in the Regulation of Secondary Cell Wall Biosynthesis in Arabidopsis. *Plant Cell* 20:2763-2782.

APPENDIX B

ANALYSIS OF SQUASH SEED FATTY ACID COMPOSITION

Squash seeds are a good source of dietary fatty acids, both when eaten as a snack food and when pressed to create squash seed oil. The major fatty acids in squash oil are oleic and linoleic acid, both of which are unsaturated. There is a smaller amount of two saturated fatty acids, palmitic and stearic, as well as a very small amount of linolenic acid. To maximize the health benefits of squash seeds, fatty acid composition is an important target for improvement through plant breeding. To enable this breeding goal, it is necessary to investigate the existing variation in fatty acid composition of squash seeds.

In 2010, we surveyed the fatty acid profiles of the *C. pepo* and *C. moschata* Plant Introduction (PI) collections from the US Department of Agriculture's National Plant Germplasm System. We phenotyped seed sent directly from the PI collection, sending them to the Iowa State DNA facility where they determined the percent composition of five major fatty acids (palmitic, stearic, oleic, linoleic, and linolenic) in the seed oil.

We found that the percentages of all the main fatty acids varied substantially (Figures B.1 and B.2, Table B.1), but that the unsaturated fatty acids, especially linoleic acid, made up the majority of the seed oil. In general, the two species had similar oil compositions, but *C. moschata* seeds tended to have a slightly higher percentage of saturated fatty acids than *C. pepo* and a corresponding lower percentage of oleic acid.

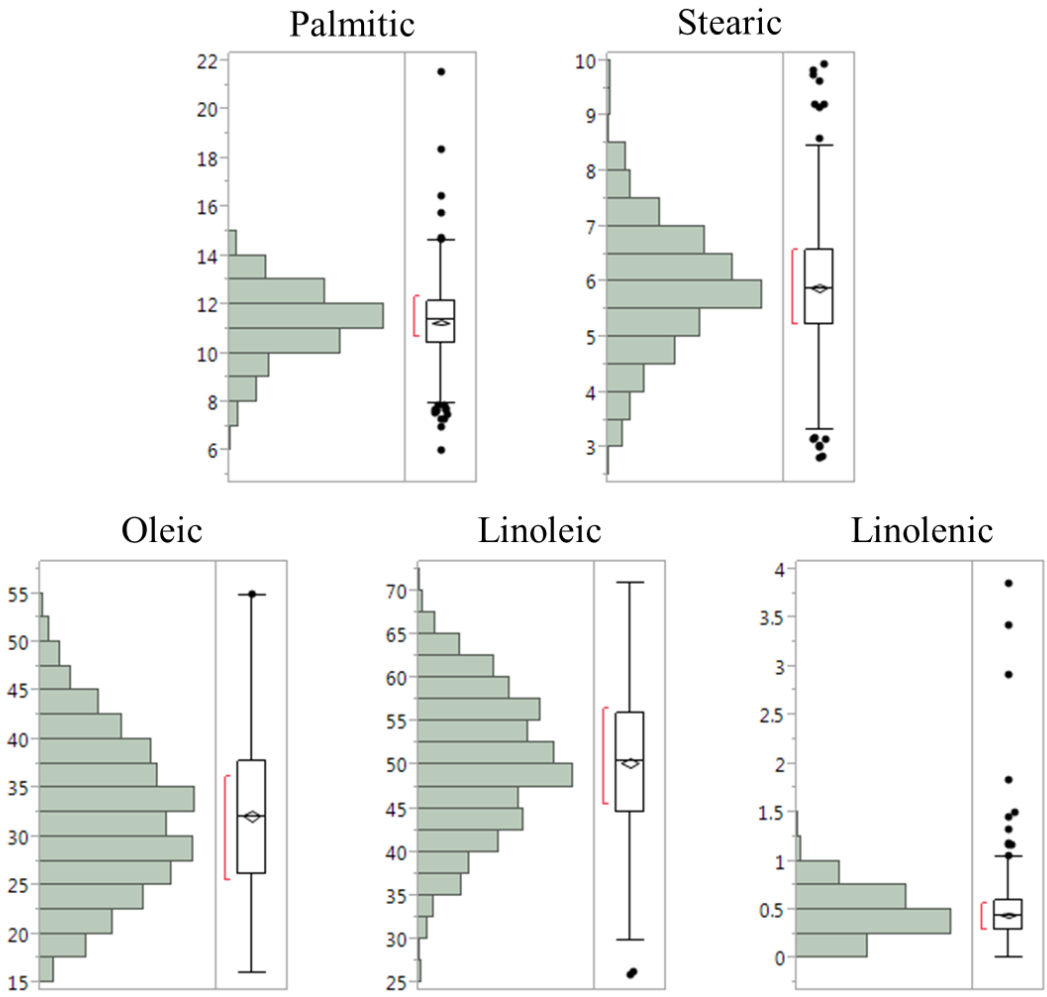


Figure B.1 Seed fatty acid composition of the *C. pepo* Plant Introduction collection. 799 *C. pepo* accessions were analyzed to determine the percentage of the total seed fatty acids comprised by each of the five major fatty acids. The distribution of these percentages for the entire collection is displayed for each fatty acid.

Because the PI collection seed was grown under different environmental conditions and stored for varying lengths of time, there was potential for these variables to affect the fatty acid profiles of squash seeds. To confirm and investigate the phenotypic extremes of the PI collections, we selected 10 *C. pepo* and 10 *C. moschata* accessions that represented these phenotypic extremes, focusing mainly on accessions with the highest or lowest ratios of oleic to linoleic acid. We grew these 20

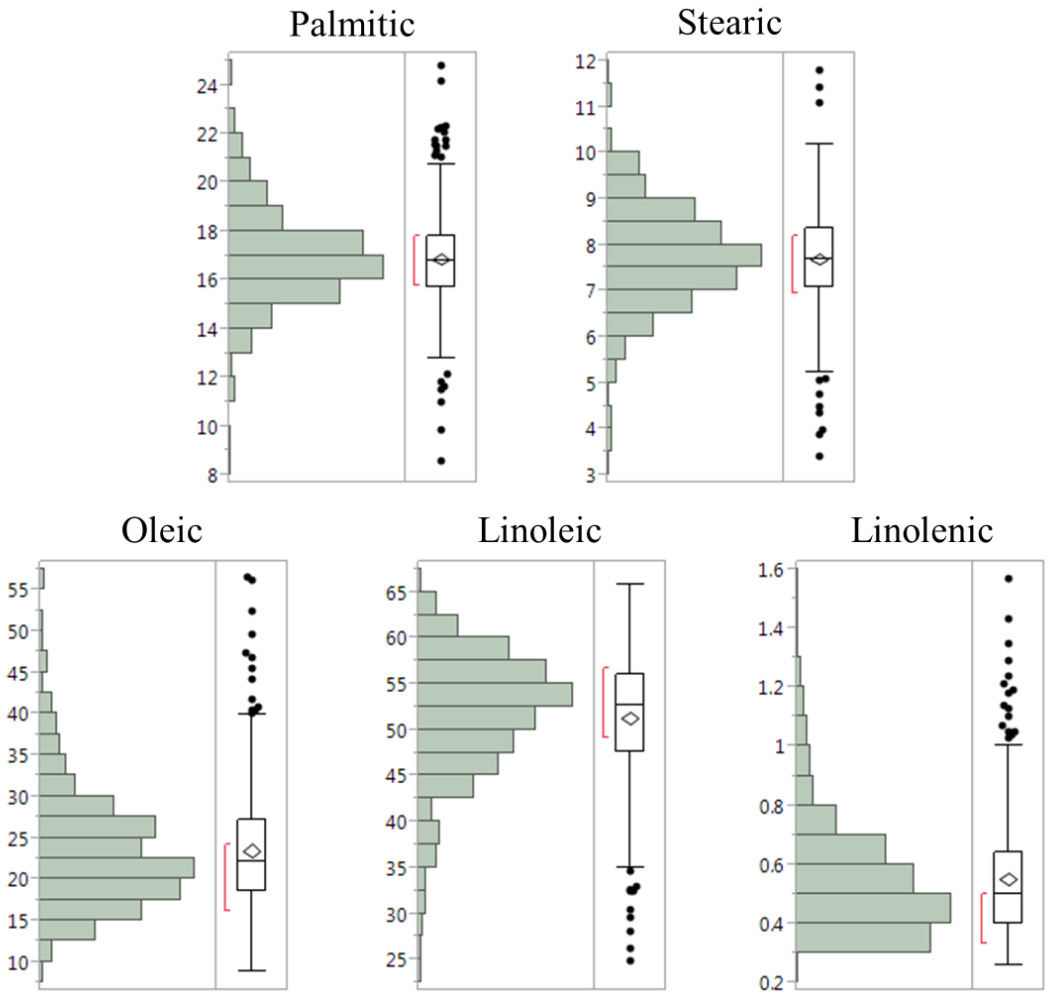


Figure B.2 Seed fatty acid composition of the *C. moschata* Plant Introduction collection. 368 *C. pepo* accessions were analyzed to determine the percentage of the total seed fatty acids comprised by each of the five major fatty acids. The distribution of these percentages for the entire collection is displayed for each fatty acid.

accessions in a common environment in Freeville, NY in the summer of 2010 and collected self-pollinated, open-pollinated, and reciprocally cross-pollinated seed. The seed collected was then sent to Iowa State for fatty acid composition analysis.

Our first experimental question was the degree of genetic versus environmental control of the seed oil fatty acid profiles. We found that there is indeed a large amount of environmental control, as the extreme profiles we saw in the parents were

Table B.1 Seed fatty acid profiles of the *C. pepo* and *C. moschata* Plant Introduction collections. 799 *C. pepo* accessions and 368 *C. moschata* accessions were analyzed to determine the percentage of the total seed fatty acids comprised by each of the five major fatty acids. The mean, maximum, and minimum percentages are reported for each of the fatty acids.

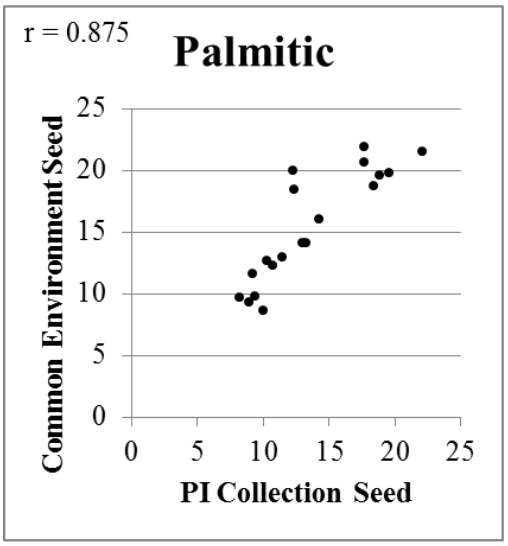
	C. pepo			***C. moschata***		
Fatty acid	**Mean**	**Maximum**	**Minimum**	**Mean**	**Maximum**	**Minimum**
Palmitic	11.27%	21.56%	6.01%	16.88%	24.77%	8.59%
Stearic	5.88%	9.93%	2.80%	7.69%	11.80%	3.42%
Oleic	32.16%	54.96%	15.30%	23.51%	56.53%	8.92%
Linoleic	50.25%	70.87%	25.92%	51.36%	65.84%	24.91%
Linolenic	0.44%	3.85%	0.00%	0.55%	1.57%	0.26%

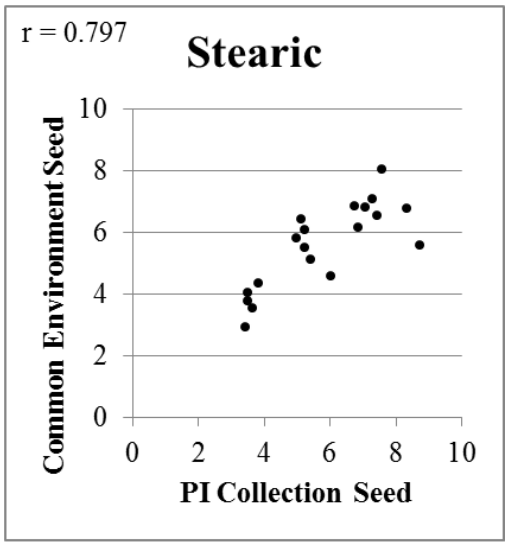
sometimes replaced by moderate profiles or even the opposite extreme when grown in the new location. Six of the twenty did still display the same extreme profile, suggesting that the PI collection seed we initially phenotyped for those accessions may have been grown in a similar environment. Overall, the correlation between the PI collection seed and the common environment seed for each fatty acid ranged from -0.084 (linolenic acid) to 0.875 (palmitic acid) (Figure B.3) and the two saturated fatty acids had a higher correlation than the unsaturated fatty acids. The lower correlation of the unsaturated fatty acids could potentially be due to their oxidation in the PI collection seed, which had been stored for varying lengths of time. The ratio between oleic and linoleic acids, the two primary unsaturated fatty acids, is also known to be affected by temperature during the growing season in multiple species including sunflower (Flagella et al., 2002; Harris et al., 1978) and soybean (Gibson and Mullen, 1996), which could explain why they are more environmentally variable. Because there was such a poor correlation overall between the PI collection seed and the

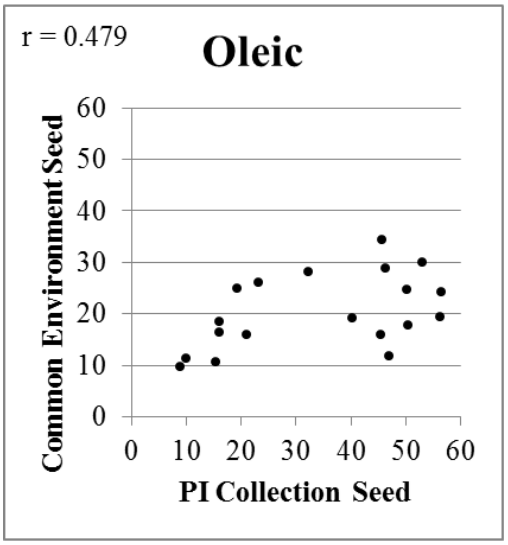
common environment seed, we decided that the environmental variation in fatty acid profiles was too great to warrant further research.

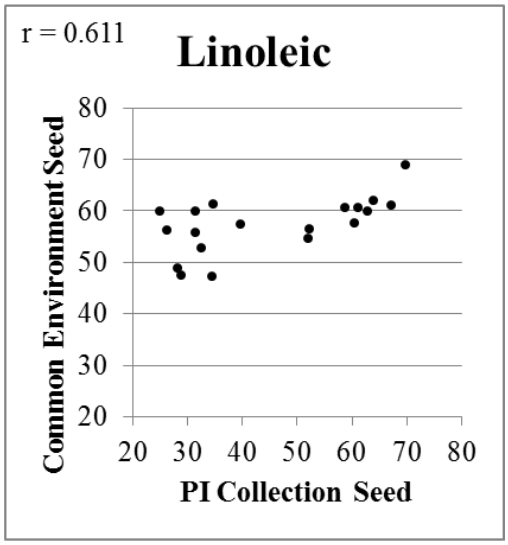
We also investigated the degree of maternal control of seed oil fatty acid composition by comparing the profiles of self-pollinated seed (paternal parent same as maternal parent), open-pollinated seed (random paternal parent), and reciprocally cross-pollinated seed (paternal parent of opposite extreme profile). In all cross-pollinated seed, the fatty acid profile was different than that of self-pollinated seed from the maternal parent. In more than half of the cases, the change was in the direction of the paternal parent, indicating that both parents may contribute to the fatty acid profile of the seed. This is logical, as the oil in squash seeds is contained in the embryo, which contains genetic material from both parents. While the low number of replicates and genotypes investigated in this study prevents firm conclusions, it highlights an area for future study and suggests that self-pollinated seed should be used for more replicable fatty acid analysis.

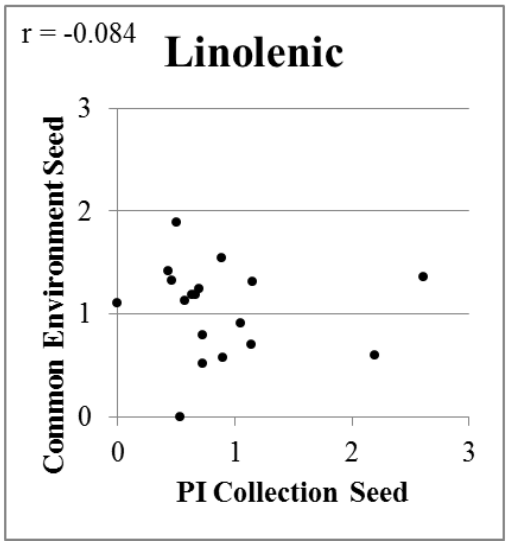
Figure B.3 Correlation between Plant Introduction seed and common environment seed fatty acid profiles. Both PI collection seed and self-pollinated seed produced in Freeville, NY in 2010 were analyzed for nineteen selected accessions to determine the percentage of the total seed fatty acids comprised by each of the five major fatty acids. These percentages are plotted against each other, depicting the degree of correlation between the two environments for each of the five fatty acids.


r = 0.875
Palmitic
Common Environment Seed
PI Collection Seed
0
5
10
15
20
25


r = 0.797
Stearic
Common Environment Seed
PI Collection Seed
0
2
4
6
8
10


r = 0.479
Oleic
Common Environment Seed
PI Collection Seed
0
10
20
30
40
50
60


r = 0.611
Linoleic
Common Environment Seed
PI Collection Seed
20
30
40
50
60
70
80


r = -0.084
Linolenic
Common Environment Seed
PI Collection Seed
0
1
2
3

References

Flagella, Z., T. Rotunno, E. Tarantino, R. Di Caterina, and A. De Caro, 2002. Changes in seed yield and oil fatty acid composition of high oleic sunflower (*Helianthus annuus* L.) hybrids in relation to the sowing date and the water regime. *European Journal of Agronomy* 17:221-230.

Gibson, L.R. and R.E. Mullen, 1996. Soybean seed composition under high day and night growth temperatures. *Journal of the American Oil Chemists Society* 73:733-737.

Harris, H.C., J.R. McWilliam, and W.K. Mason, 1978. Influence of temperature on oil content and composition of sunflower seed. *Australian Journal of Agricultural Research* 29:1203-1212.

APPENDIX C

SUMMARY OF THE 2014 STUDENT ORGANIC SEED SYMPOSIUM

In August 2014, Cornell University's Department of Plant Breeding and Genetics hosted the third annual Student Organic Seed Symposium with the support of faculty advisor Michael Mazourek. The Student Organic Seed Symposium is an event planned by and for graduate students with the goal of bringing together graduate students passionate about organic and sustainable seed systems. The 2014 symposium brought together 24 graduate students from around the United States, 8 distinguished speakers, and 12 additional participants from the seed industry for 4 days. The theme of the symposium was "Regional Adaptation for Sustainable Food and Seed Systems", which was highlighted by a day of tours which included visits to a local organic farm, a regional seed company, a local miller, and more. Other events included presentations of student breeding material grown in a demonstration garden at the Freeville Organic Farm, oral presentations about student research projects, and meals featuring locally-grown ingredients (see below). This event was made possible by the generous support of 14 donors (see below), who enabled our mission of bring students together to network, learn, and be inspired by the work of their peers. Capitalizing on the congregation of so many industry experts, the 2014 Organic Seed School was held the day before the symposium, bringing together seven seed companies, other seed industry experts, students, and local growers to discuss issues surrounding the production of organic seed. An optional trip to the Stone Barns Center for Food and Agriculture completed the event, in which students were able to tour the center and learn about the innovative growing practices and outreach that take place there.

2014 Planning Committee Members

Lindsay Wyatt, Cornell University
Dylan Wann, Texas Tech University
Hannah Swegarden, University of Minnesota
Jay Bost, University of Hawaii
John Hart, USDA-ARA Tropical Agriculture Research Station
Randi Jiménez, University of California, Davis

Schedule of Events

Sunday, August 17th – Day One

- 2014 Organic Seed School
- Welcome Speech by SOSS Organizing Committee
- Welcome Dinner in Tent on Resort Grounds
- Outdoor Social

Monday, August 18th – Day Two

- Session I – "*Foundations of Organic Seed*"
    - Mark Sorrells
    - E.T. Lammerts van Buren
- Session II – "*Integrated Systems*"
    - R. Ford Denison
    - Albert Culbreath
- Visit to Freeville Organic Farm
    - Lunch and Introduction
    - Farm Tour and Faculty Talks
    - Student Introductions and Trial Plot Demonstrations
- Dinner with Cornell Plant Breeding Department

Tuesday, August 19th – Day Three

- Depart for Field Tour Day
- Tour Kingbird Farm
- Tour Farmer Ground Flour
- Visit to Bejo Seeds
    - Lunch hosted by Bejo Seeds
    - Tour of Bejo Seeds Demonstration Gardens
- Tour USDA Vegetable Germplasm Repository
- Tour USDA Apple and Grape Germplasm Collection
- Tour Remembrance Farm with Fruition Seeds
- Dinner at Northstar Restaurant

Wednesday, August 20th – Day Four

- Session III – "*Utilizing Exotic Germplasm in Breeding*"
    - Major Goodman
    - Jane Denver
- Seed Swap
- Student Research Presentations
- Lunch at Park Next to Resort
- Session IV – "*Seed Solidarity*"
    - Jared Zystro
- Planning Meeting: SOSS 2015
- Final Remarks
    - Humberto Rios Labrada
    - Matthew Dillon
- Closing Dinner at Park Next to Resort

Thursday, August 21st – Day Five

- Visit to Stone Barns Center

<u>2014 Sponsors</u>

Department of Plant Breeding and Genetics, Cornell University
Cornell University Agricultural Experiment Station
Cornell University New York State Agricultural Experiment Station
Cornell University Cooperative Extension
Dr. Michael Mazourek – Cornell University
Seed Matters
The Ceres Trust
The Bost Family
Peter Meinig
Albert Lea Seed House
Bejo Seeds
High Mowing Organic Seeds
Johnny's Selected Seeds
Vitalis Organic Seeds