DIHYDROCHALCONES IN *MALUS* MILL. GERMPLASM AND HYBRID POPULATIONS

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

Benjamin Leo Gutierrez

December 2017

© 2017 Benjamin Leo Gutierrez

DIHYDROCHALCONES IN *MALUS* MILL. GERMPLASM AND HYBRID POPULATIONS

Benjamin Leo Gutierrez, Ph.D.

Cornell University 2017

Dihydrochalcones are abundant in *Malus* Mill. species, including the cultivated apple (*M.* ×*domestica* Borkh.). Phloridzin, the primary dihydrochalcone in *Malus* species, has beneficial nutritional qualities, including antioxidant, anti-cancer, and anti-diabetic properties. As such, phloridzin could be a target for improvement of nutritional quality in new apple cultivars. In addition to phloridzin, a few rare *Malus* species produce trilobatin or sieboldin in place of phloridzin and hybridization can lead to combinations of phloridzin, trilobatin, or sieboldin in interspecific apple progenies. Trilobatin and sieboldin also have unique chemical properties that make them desirable targets for apple breeding, including high antioxidant activity, anti-inflammatory, anti-diabetic properties, and a high sweetness intensity. We studied the variation of phloridzin, sieboldin, and trilobatin content in leaves of 377 accessions from the USDA National Plant Germplasm System (NPGS) *Malus* collection in Geneva, NY over three seasons and identified valuable genetic resources for breeding and researching dihydrochalcones. From these resources, five apple hybrid populations were developed to determine the genetic basis of dihydrochalcone variation. Phloridzin, sieboldin, and trilobatin appear to follow segregation patterns for three independent genes and significant trait-marker associations were identified using genetic data from genotyping-by-sequencing.

Dihydrochalcones are at much lower quantities in mature apple fruit compared with vegetative tissues. Within the fruit, phloridzin is more concentrated in the peel than in the flesh. We observed higher phloridzin content in the peel of russet apples compared to the peel of non-

russet apples. Russet is a disorder affecting the development of the fruit cuticle, causing a smooth waxy surface to be replaced by a rough, russet colored layer. We compared leaf, and fruit peel and flesh samples from 108 accessions in the USDA-NPGS *Malus* collection and identified a strong correlation between percent russeting and phloridzin content the fruit peel, especially in sport families with variation for russeting. Though russeting can severely impact commercial value of apple cultivars, russeted fruit have a unique nutritional profile compared to non-russet cultivars.

BIOGRAPHICAL SKETCH

Benjamin Leo Gutierrez was born July 31, 1984 in Provo, UT, the youngest child of Ronald Phillip Gutierrez and Jean Crandall Gutierrez. His love of science began with teeth. Throughout his youth, he learned his father's trade as dental technician, making dentures and other prosthetics in their basement laboratory, Timpanogos Dental Lab. From this, he developed a love for the diversity of form and function of teeth. To further his dental career, Benjamin enrolled at Utah Valley University, with a goal to pursue dentistry. The first in his family to attend college, he fell in love with science and learning. In 2005, Benjamin postponed his education to serve a two-year mission in Belgium and France for the Church of Jesus Christ of Latter-day Saints. Returning to Utah, Benjamin re-enrolled in college and took courses in genetics and plant physiology. Finding his love of plant science outweighed his love of teeth, he sought a new career as a scientist. He engaged in undergraduate research with Dr. Olga Ruiz Kopp, studying tissue culture of endangered Utah species, *Astragalus holmgreniorum* and pathogens of *Pinus longaeva.*

Upon graduating from Utah Valley University in 2012 with BS in Biology and Botany, Benjamin applied for graduate studies at Cornell University to study plant breeding. There, he had the opportunity to work in partnership with Cornell's apple breeder, Dr. Susan Brown, and Dr. Gan-Yuan Zhong of the USDA-ARS Plant Genetic Resources Unit at the New York State Agricultural Experiment Station (NYSAES). Throughout his graduate program, he had the unique opportunity to pursue advanced education and work for the USDA. In 2017, he was recognized for five years of service in federal government. The highlight of his experience was spending time in the apple repository, studying (and tasting) historic apple cultivars and wild species from around the world, and working with Dr. Brown, tasting the apples of the future.

For my wife Emily, and our children Della, Leo, and Sullivan

ACKNOWLEDGMENTS

I am grateful to **Dr. Susan Brown**, for taking me on as a student and for all the time she spent searching the literature, reading drafts, making observations in the field and greenhouses, and offering much needed encouragement. I have learned from her passion for plant breeding, genetics, and the art of communication. With great confidence, she gave me freedom to expand my research project to fit my interests, and helped me to develop our apple populations. From her abounding humor, I learned to take my research seriously but not myself, and how to appreciate family and those who support me. As a good plant breeder, Dr. Brown sees value in things for what they could be, and not for what they are (including graduate students).

I am grateful to **Dr. Gan-Yuan Zhong**, Research Leader, for the USDA-ARS Plant Genetic Resources Unit (PGRU). He provided important leadership both academically and professionally as my USDA supervisor and committee member. He was my voice to Area Office leaders in funding for research and ultimately a position within the unit through the USDA Pathways program. He advocated on my behalf of my value and contributions to the location. He was willing to talk about life, careers, and science, and was always very optimistic about my potential as a scientist in training.

I would also like to thank my committee members **Dr. Lailiang Cheng** and **Dr. Michael Gore**. They provided key insights into my projects, technical experience, read early drafts, and allowed me to drop by their offices, often unannounced. I learned the principles of chromatography from Dr. Cheng, and to look beyond genetics, keeping physiological processes in mind. Dr. Gore taught me to think critically about statistics and quantitative genetics.

I am grateful to have worked with **Dr. Thomas Chao**, **Dr. Gennaro Fazio**, and **Dr. Heidi Schwaninger** of PGRU. I learned much from them about *Malus* diversity, the value of

genetic resources, working with clonal crops, careful lab techniques, and how to think critically about my research. A special thanks to **Dr. Jie Arro**, a PGRU postdoc, who was always willing to talk about an idea for my projects, answer my questions about statistics, and how to love computer programming. Dr. Arro taught me how to run the GBS pipeline, and to think critically about data analysis.

I also thank the 'A-Team', **Bill Srmack**, **John Keeton, Bob Martens, and Julian Koob** of PGRU. Working under Bill Srmack's leadership, they keep the collections alive and in excellent condition. Truly, it is their work that people come from all over the world to see. They were always quick with a joke, and took time to teach me about horticulture and management of the clonal repository, even when I got vehicles stuck in the mud. Though he's moved on from our unit, I am grateful to **Andy Humiston**, who taught me to run and maintain our HPLC. We spent a lot of hours together learning the art of chromatography. I am grateful to Julian Koob for his help in preparing and extracting samples for HPLC.

I am grateful to **Kevin Maloney**, whose good humor and vast knowledge of apples improved the quality of my research and my life. Without Kevin and **Matt Garlick's** technical assistance, I would not have been able to pollinate, germinate, and plant the 1,000+ seedlings that were irreplaceable to my research. I am also grateful to former lab member, **Dr. Benjamin Orcheski**, who made crosses for another project that turned out to be invaluable for mine, and suggested we look at russeted apples.

I also thank **Dr. Rose Nichols**, **Tiffany Fisk**, and **Martin Hooker** with the USDA who keep everything running smoothly and are the heart of the USDA-ARS Geneva location. This project was funded by the USDA-ARS Plant Genetics Resources Unit, Geneva NY. I am grateful to the **Plant Breeding and Genetics section**, the **New York State Agricultural Experiment**

Station (NYSAES), and the **United States Department of Agriculture** for their support. I also thank the **staff, graduate students, and faculty** in Geneva, NY who make NYSAES a wonderful place to work.

I am grateful to the **Arnold Arboretum of Harvard University** for providing budwood of *Malus trilobata* 127-2009.

Finally, I am grateful to my wife **Emily Ann Gutierrez,** and to our children **Della Jean, Leo Jay, and Sullivan Ronald.** They were always supportive, but excited when I'd come home to play. My wife spent many hours discussing my project with me, and learned to pronounce botanical names and chemical compounds. My children gave me a lot of exercise. Collecting leaves, and hunting for rocks and bugs with them retaught me how to see the natural world, and enjoy it with childlike wonder. My family kept life in perspective and helped me stay focused on what matters most.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

DIHYDROCHALCONES IN *MALUS* SPECIES

INTRODUCTION

Though present in other plant species, dihydrochalcones (DHCs) are a unique component of *Malus* species, including cultivated apple, *Malus* ×*domestica.* Apples are the most significant source of dietary DHCs, primarily in the form of phloridzin. Since its discovery, phloridzin has been studied for its effects on human and animal physiology, with little emphasis on biochemical synthesis or regulation, or physiological significance in apple and its wild relatives. The purpose of this chapter is to provide an overview of DHCs and their prevalence in *Malus* species.

Dihydrochalcones

DHCs belong to the flavonoid class of phenolic compounds. Structurally, they are similar to chalcones, with three carbon-bridge joining two phenyl rings, but differ in the reduction of the Cα-Cβ double bond. This double bond reduction separates DHCs early in the synthesis of other flavonoids. They are derived from both the phenylpropanoid and polyketide pathways with *p*-coumaroyl-CoA as the structural precursor to flavonoids. DHC synthesis in apple begins with reduction of *p*-coumaroyl-CoA (phenylpropanoid pathway) to *p*-dihydrocoumaroyl-CoA (Dare et al. 2013a; Ibdah et al. 2014). Three molecules of malonyl-CoA (polyketide pathway) are added to *p*-dihydrocoumaroy*l*-CoA forming phloretin via chalcone synthase, which is glycosylated to phloridzin in apple (Gosch et al. 2009). Recently, the DHC pathway has been more fully elucidated, including description of key enzymes and genes from *M.* ×*domestica* related to phloridzin synthesis. Dare et al. (2013a) and Ibdah et al. (2014) independently described two distinct gene families in apple, capable of reducing the 2-3 double bond in *p*-coumaroyl-CoA.

Though several enzymes can glycosylate phloretin *in vitro*, UGT88F1 and UGT88F4 have a prominent role in phloridzin synthesis *in planta* (Dare et al. 2017; Gosch et al. 2010a; Zhou et al. 2017). (Figure 1.1). Chalcone synthase (CHS) is the first committed step of flavonoid synthesis. There are several CHS-like genes in apple, however, *MdCHS1*, *MdCHS2*, and *MdCHS3* were shown to be associated with phloridzin synthesis (Dare et al. 2013b; Yahyaa et al. 2017).

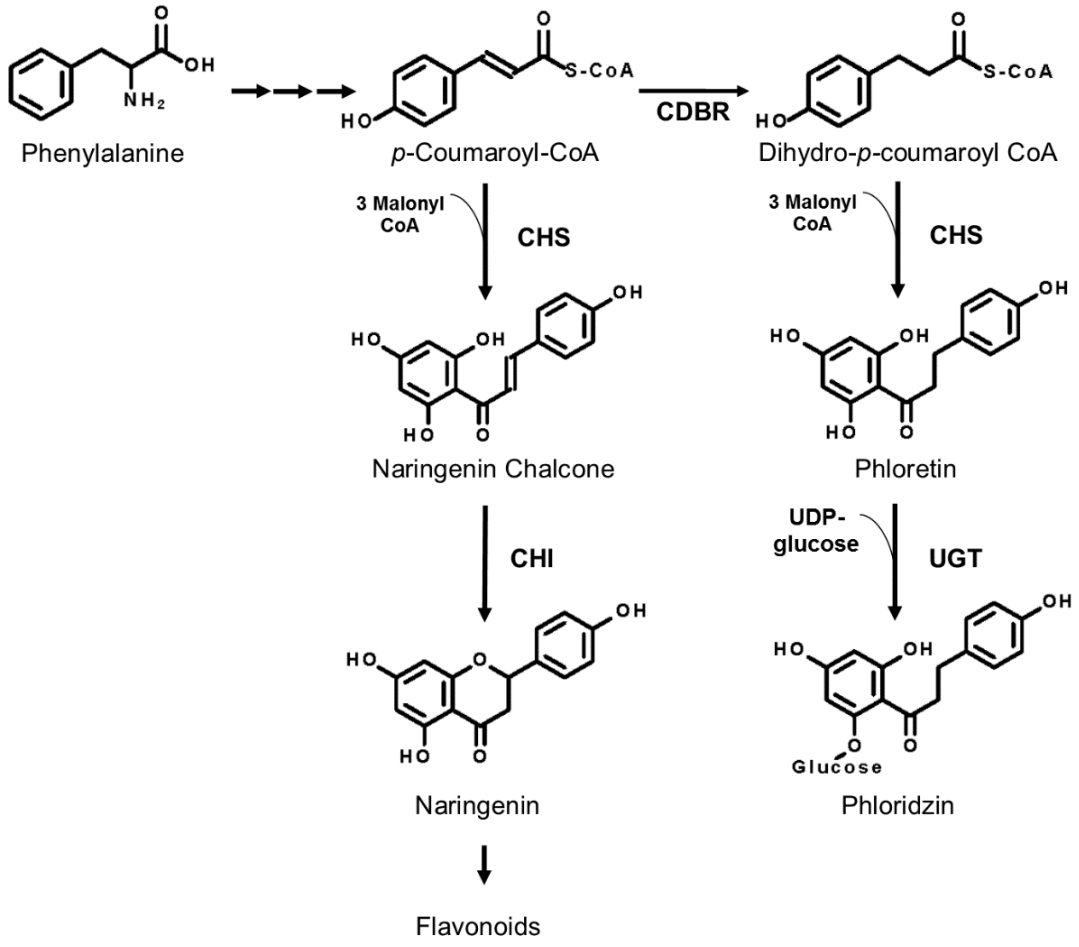


Figure 1.1 Proposed dihydrochalcone pathway in *Malus* (right) and beginning of flavonoid pathways (left). Abbreviations: carbon double bond reductase (CDBR), chalcone synthase (CHS), chalcone isomerase (CHI), uridine diphosphate glycosyltransferase (UGT)

The basic structure of DHCs allows for a variety of substitutions, including hydroxylation, methylation, and glycosylation leading to many potential compounds. However, only 256 naturally occurring DHCs have been described across a diverse range of plant families, including pteridophytes and angiosperms. Dietary sources of DHCs are rare, as many species that produce DHCs are used for medicinal or ornamental use. Significant sources of DHCs include aspalathin from rooibos tea (*Aspalathus linearis*); neohesperidin dihydrochalcone, a synthetic sweetener; and phloridzin from cultivated apples (Rivière 2016). In apple and its wild relatives, several DHCs have been identified, with phloridzin being the most prominent (Table 1.1). Phloridzin was first described by De Konink in 1835 as a bitter compound in the bark of apple trees (Gosch et al. 2010b). Initially, phloridzin was thought to be restricted to *Malus* species, but has since been identified in species both within the Rosaceae, including rose (Hvattum 2002) and strawberry (Hilt et al. 2003), and outside the Rosaceae, including, *Lithocarpus* species (Zhao et al. 2014), and cranberry (*Vaccinium macrocarpon*) (Turner et al. 2005).

Table 1.1 List of dihydrochalcones isolated from *Malus* species

Dihydrochalcone	Reference
Phloretin	Gosch et al. 2010b
Phloridzin (phloretin-2′-*O*-glucoside)	Gosch et al. 2010b
Sieboldin (3-hydroxyphloretin 4′-*O*-glucoside)	Dugé de Bernonville et al. 2010
Trilobatin (phloretin-4′-*O*-glucoside)	Dugé de Bernonville et al. 2010
Phloretin-2′-O-xyloglucoside	De Paepe et al. 2015
3-hydroxyphloridzin	Hutabarat et al. 2016
Phloretin-pentosyl-hexoside	Ramírez-Ambrosi et al. 2015
Phloretin-hexosyl-hexoside	Ramírez-Ambrosi et al. 2015

Trilobatin and sieboldin are other DHCs found in select *Malus* species and hybrids. *Malus toringo* produces sieboldin in place of phloridzin, and *Malus trilobata* produces trilobatin in place of phloridzin (Williams 1961). Interspecific hybridization with these *Malus* results in progeny with multiple dominant DHCs. In experiments with hybrid apple populations, Williams and Jarret (1975) observed segregation of phloridzin, sieboldin, and trilobatin in F_1 progeny, and categorized the parents as either homozygous or heterozygous for each of the DHCs based on progeny ratios. Though described decades ago, no candidate genes or loci for these DHCs have been proposed, nor have their pathways been described. Trilobatin and sieboldin differ from phloridzin in the placement of the glucose moiety on the C4′ position rather than C2′, and sieboldin from trilobatin by the C3 hydroxylation. It was suggested that a specific 4′-glycosyltransferase is responsible for the glycosylation of phloretin to trilobatin (Yahyaa et al. 2016) and 4′-glycosylation follows C3 hydroxylation of phloretin in sieboldin. While, phloridzin has been evaluated biochemically and medicinally, sieboldin and trilobatin have been poorly researched in apple. Most of what is understood about trilobatin, originates from studies of *Lithocarpus* species (Dong et al. 2012; Wang et al. 2016), and very few have reported on sieboldin (Dugé de Bernonville et al. 2010; Xiao et al. 2017).

Phloridzin has been researched extensively for nutraceuticals properties, including; diabetes, cancer, antioxidant, water soluble dyes, and cosmetics (Gosch et al. 2010b). The effects of phloridzin as an anti-diabetic has been documented. Phloridzin aids in blood sugar regulation through inhibition of sodium linked glucose transporters in the intestines, and induction of glycosuria (excretion of sugar in the urine) (Ehrenkranz et al. 2005). Trilobatin also impacts diabetes, but through α-glucosidase inhibition (Dong et al. 2012), and sieboldin is a significant antioxidant (Dugé de Bernonville et al. 2010).

Apple is one of the most important food crops in the US and in the world. Consumed as fresh fruit, juice, and processed products, they are among the most popular and valuable commodities, and are positively associated with human health (Hyson 2011). Nutritional improvement in apple has emphasized selection for higher phenolic content, including red-fleshed cultivars (Bars-Cortina et al. 2017; Farneti et al. 2015). However, there are several diverse sources of dietary phenolics (Francini and Sebastiani 2013; McDougall 2017), whereas, DHCs are obtained primarily from apple. Selection for increased DHC content would further diversify the nutritional profile of apple from other foods. This process could include selection for increased phloridzin content, and introgression of trilobatin and sieboldin into advanced selections. Though sieboldin and trilobatin are found in some ornamental apples, there are no commercial cultivars with significant sieboldin or trilobatin content.

DHC content disproportionately accumulates in vegetative tissues over fruit. Previous reports in *Malus* suggest phloridzin can represent 66 to 90% of total phenolic content in vegetative tissues and more than 15% of leaf dry weight, but only 2 to 6% of phenolics in mature fruit (Gosch et al. 2010b). The difference in phloridzin concentrations between vegetative tissues and fruit is around three orders of magnitude. Phloridzin content decreases throughout fruit development. Genetic studies have focused on either leaf phloridzin synthesis (Yahyaa et al. 2017) or fruit (Baldi et al. 2017; Khan et al. 2012), but none have considered the two tissues together. A clear understanding of phloridzin regulation between fruit and vegetative tissues may aid in breeding for higher DHC content.

The cultivated apple is a complex interspecific hybrid. The domestication process of apple from its primary progenitor, *Malus sieversii* (from Kazakhstan), included not only selected for key traits such as, fruit size, flavor, sweetness, and acidity, but introgression from wild

species, including *Malus sylvestris* and *Malus orientalis*. (Cornille et al. 2012; Duan et al. 2017; Khan et al. 2014). There is a distinct pattern of introgression of wild species from European and North American cultivars from Asia and Middle Eastern cultivars (Gharghani et al. 2009; Ma et al. 2017). Modern apple breeding has involved introgression with additional wild species, such as *Malus floribunda*, particularly for apple rootstock breeding (Brown 2012). Leaf phloridzin content variation within and among cultivated apple and its progenitors is comparable, but lower than other wild *Malus* (Gutierrez et al. 2017a), and fruit phloridzin variation varies significantly in *M.* ×*domestica* (Anastasiadi et al. 2017; Gutierrez et al. 2017b).

Malus taxonomy is not entirely resolved. Within the genus *Malus* there are approximately 59 species and interspecific hybrids, with geographic distributions in Central Asia, East Asia, Southeast Asia, Europe, and North America (Nikiforova et al. 2013; Volk et al. 2015a). Genetic relationships between wild *Malus* species is poorly understood, as research into the domestication of apple has been prioritized. Though China is the center of diversity for *Malus*, three Eastern North American species, *Malus angustifolia*, *Malus coronaria*, and *Malus ioensis* appear to be more distantly related to other *Malus* species (Hokanson et al. 2001; Volk et al. 2015b). Due to obligate outcrossing, interspecific hybridization, and overlapping generations through clonal propagation, *Malus* species maintain a high level of genetic diversity. Genome sequencing of 'Golden Delicious' revealed an genome duplication of nine ancestral chromosomes, resulting in 17 chromosome pairs, following the loss of one chromosome (Velasco et al. 2010). Genome size and ploidy varies significantly between *Malus* species (Chagné et al. 2015; Höfer and Meister 2010). Throughout its domestication history, cultivated apple has maintained its genetic diversity. Comparative analysis between ancient apple cultivars

and modern cultivars, has shown little reduction in genetic diversity across centuries, relative to its wild progenitors (Gross et al. 2014).

There are many questions concerning the evolutionary significance of DHCs in *Malus* species. All *Malus* species accumulate high levels of DHCs, whereas closely related genera, such as *Pyrus*, do not (Gosch et al. 2010b). Presumably, phloridzin, the prominent DHC in North American, European, and Asian species, represents the ancestral DHC in *Malus*. Divergent evolution is possible in the DHCs of *M. trilobata*, native to Turkey, Lebanon, and Israel, and *M. toringo,* native to Southeast Asia. In addition to its unique chemistry, *M. trilobata* is the sole member of the section Eriolobus, and genetic and morphological studies place it close to the three Eastern North American *Malus* species (Höfer et al. 2014; Hokanson et al. 2001; Robinson et al. 2001). The DHC pathways in these species presents an intriguing question: given the apparent dependency of sieboldin synthesis on the trilobatin pathway, did trilobatin evolve twice in these unrelated and geographically isolated species? With a better understanding of the genetic controllers of DHC, we may understand how these pathways may have evolved.

Genetic studies of *Malus* typically involve pedigreed based linkage mapping of one or multiple F_1 populations. Association mapping is a valuable method for genetic analysis, using large collections of unrelated germplasm with more diversity for an array of traits (Zhu et al. 2008). It is particularly advantageous for perennial crops due to long generation times and high cost of population development and maintenance. Association studies in *Malus* are challenging due to rapid decay of linkage disequilibrium, however, through advancements in genotyping platforms for apple, they are becoming more broadly used (Di Guardo et al. 2017; Kumar et al. 2013; Lozano et al. 2014). Utilizing genomic resources and modern tools of molecular biology,

including genome editing, dissecting metabolic traits in apple is becoming feasible (Nishitani et al. 2016; Zhou et al. 2017).

There is some evidence for physiological roles of DHCs in *Malus*, particularly phloridzin. Early research suggested phloridzin was a plant growth regulator due to its effects on physiological processes in a variety of plant species, including photosynthesis, respiration, and root development (Stenlid 1968). In the 1960s, phloridzin was implicated in auxin regulation through increased IAA oxidase activity (Stenlid 1968). However, many experiments neglected to understand DHCs in plant species, in which they are natively produced. Recently, Dare et al. (2013; 2017) demonstrated that reducing phloridzin content through silencing of both chalcone synthase and specific phloretin glycosyltransferases in 'Royal Gala', results in reduced auxin transport and aberrant tree morphology. Whether or not trilobatin or sieboldin have similar roles needs to be investigated. Stenlid (1968) observed that phloridzin promotes IAA oxidase activity, whereas sieboldin does not.

The purpose of the research presented in the following chapters was to better understand the genetic diversity of DHC content in a *Malus*, identify genetic controllers for further analysis, and to characterize the variation of phloridzin content in fruit of cultivated apple. We identified genetic resources, diverse in DHC content and composition, which are valuable for further genetic studies and breeding. In apple germplasm we determined that DHC content has high heritability, and we identified a previously undescribed locus on apple linkage group 7 associated with accessions with trilobatin and sieboldin. Using linkage mapping in apple F_1 hybrids, we observed that trilobatin and sieboldin segregate as independent loci and identified a locus on linkage group 8 associated with sieboldin content, as well as the same locus on linkage group 7 described in the germplasm, associated with trilobatin content. Though DHC content is generally

low in fruit comparted to vegetative tissues, we identified unique cultivars with high phloridzin content in the fruit peel. In particular, we identified an association with russeting, a cuticle disorder, and higher peel phloridzin content.

REFERENCES

Anastasiadi M, Mohareb F, Redfern SP, Berry M, Simmonds MSJ, Terry LA (2017) Biochemical profile of heritage and modern apple cultivars and application of machine learning methods to predict usage, age, and harvest season. J Agric Food Chem. doi:10.1021/acs.jafc.7b00500

Baldi P, Moser M, Brilli M, Vrhovsek U, Pindo M, Si-Ammour A (2017) Fine-tuning of the flavonoid and monolignol pathways during apple early fruit development. Planta 245:1021–1035. doi:10.1007/s00425-017-2660-5

Bars-Cortina D, Macià A, Iglesias I, Romero MP, Motilva MJ (2017) Phytochemical profiles of new red-fleshed apple varieties compared with traditional and new white-fleshed varieties. J Agric Food Chem 65:1684–1696. doi:10.1021/acs.jafc.6b02931

Brown S (2012) Apple. In: Badenes ML, Byrne DH (eds) Fruit Breeding. Springer, New York, pp 329–367

Chagné D, Kirk C, Whitworth C, Erasmuson S, Bicknell R, Sargent DJ, Kumar S, Troggio M (2015) Polyploid and aneuploid detection in apple using a single nucleotide polymorphism array. Tree Genet Genomes 11:94. doi:10.1007/s11295-015-0920-8

Nishitani C, Hirai N, Komori S, et al (2016) Efficient genome editing in apple using a CRISPR/Cas9 system. Scientific Reports 6:31481

Cornille A, Gladieux P, Smulders MJ, Roldán-Ruiz I, Laurens F, Cam BL, et al (2012) New insight into the history of domesticated apple: secondary contribution of the European wild apple to the genome of cultivated varieties. PLoS Genet 8:e1002703. doi:10.1371/journal.pgen.1002703

Dare A, Hellens R (2013) RNA interference silencing of *CHS* greatly alters the growth pattern of apple (*Malus* x *domestica*). Plant Signal Behav 8:e25033. doi:10.4161/psb.25033

Dare AP, Tomes S, Cooney JM, Greenwood DR, Hellens RP (2013a) The role of enoyl reductase genes in phloridzin biosynthesis in apple. Plant Physiol Biochem 72:54–61. doi:10.1016/j.plaphy.2013.02.017

Dare AP, Tomes S, Jones M, McGhie TK, Stevenson DE, Johnson RA, Greenwood DR, Hellens RP (2013b) Phenotypic changes associated with RNA interference silencing of chalcone synthase in apple (*Malus* × *domestica*). Plant J 74:398–410. doi:10.1111/tpj.12140

Dare AP, Yauk Y-K, Tomes S, McGhie TK, Rebstock RS, Cooney JM, Atkinson RG (2017) Silencing a phloretin-specific glycosyltransferase perturbs both general phenylpropanoid biosynthesis and plant development. Plant J Cell Mol Biol 91:237–250. doi:10.1111/tpj.13559

De Paepe D, Valkenborg D, Noten B, Servaes K, Diels L, Loose MD, Van Droogenbroeck B, Voorspoels S (2015) Variability of the phenolic profiles in the fruits from old, recent and

new apple cultivars cultivated in Belgium. Metabolomics 11:739–752. doi:10.1007/s11306-014-0730-2

Di Guardo M, Bink MCAM, Guerra W, Letschka T, Lozano L, Busatto N, et al (2017) Deciphering the genetic control of fruit texture in apple by multiple family-based analysis and genome-wide association. J Exp Bot 68:1451–1466. doi:10.1093/jxb/erx017

Dong H-Q, Li M, Zhu F, Liu F-L, Huang J-B (2012) Inhibitory potential of trilobatin from *Lithocarpus polystachyus* Rehd against α-glucosidase and α-amylase linked to type 2 diabetes. Food Chem 130:261–266. doi:10.1016/j.foodchem.2011.07.030

Duan N, Bai Y, Sun H, Wang N, Ma Y, Li M, et al (2017) Genome re-sequencing reveals the history of apple and supports a two-stage model for fruit enlargement. Nat Commun. doi:10.1038/s41467-017-00336-7

Dugé de Bernonville T, Guyot S, Paulin J-P, Gaucher M, Loufrani L, Henrion D, et al (2010) Dihydrochalcones: Implication in resistance to oxidative stress and bioactivities against advanced glycation end-products and vasoconstriction. Phytochem 71:443–452. doi:10.1016/j.phytochem.2009.11.004

Ehrenkranz JR, Lewis NG, Kahn CR, Roth J (2005) Phlorizin: a review. Diabetes Metab Res Rev 21:31–38

Farneti B, Masuero D, Costa F, Magnago P, Malnoy M, Costa G, Vrhovsek U, Mattivi F (2015) Is there room for improving the nutraceutical composition of apple? J Agric Food Chem 63:2750–2759. doi:10.1021/acs.jafc.5b00291

Francini A, Sebastiani L (2013) Phenolic compounds in apple (*Malus* x domestica Borkh.): compounds characterization and stability during postharvest and after processing. Antioxidants 2:181–193

Gharghani A, Zamani Z, Talaie A, Oraguzie NC, Fatahi R, Hajnajari H, Wiedow C, Gardiner SE (2009) Genetic identity and relationships of Iranian apple (*Malus* × *domestica* Borkh.) cultivars and landraces, wild *Malus* species and representative old apple cultivars based on simple sequence repeat (SSR) marker analysis. Genet Resour Crop Evol 56:829–842. doi:10.1007/s10722-008-9404-0

Gosch C, Halbwirth H, Kuhn J, Miosic S, Stich K (2009) Biosynthesis of phloridzin in apple (*Malus domestica* Borkh.). Plant Sci 176:223–231. doi:10.1016/j.plantsci.2008.10.011

Gosch C, Halbwirth H, Schneider B, Hölscher D, Stich K (2010a) Cloning and heterologous expression of glycosyltransferases from *Malus* x *domestica* and *Pyrus communis*, which convert phloretin to phloretin 2′-O-glucoside (phloridzin). Plant Sci 178:299–306. doi:10.1016/j.plantsci.2009.12.009

Gosch C, Halbwirth H, Stich K (2010b) Phloridzin: biosynthesis, distribution and physiological relevance in plants. Phytochem 71:838–843. doi:10.1016/j.phytochem.2010.03.003

Gross BL, Henk AD, Richards CM, Fazio G, Volk GM (2014) Genetic diversity in *Malus* x*domestica* (Rosaceae) through time in response to domestication. Am J Bot 101:1770–1779

Gutierrez BL, Zhong G-Y, Brown SK (2017a) Genetic diversity of dihydrochalcone content in *Malus* germplasm. Manuscript submitted for publication to Genetic Resources and Crop Evolution on 8/15/17.

Gutierrez BL, Zhong G-Y, Brown SK (2017b) Increased phloridzin content associated with russeting in apple fruit. Manuscript in preparation for submission to Genetic Resources and Crop Evolution.

Hilt P, Schieber A, Yildirim C, Arnold G, Klaiber I, Conrad J, Beifuss U, Carle R (2003) Detection of phloridzin in strawberries (*Fragaria* x *ananassa* Duch.) by HPLC-PDA-MS/MS and NMR spectroscopy. J Agric Food Chem 51:2896–2899. doi:10.1021/jf021115k

Höfer M, Meister A (2010) Genome size variation in *Malus* species. J Bot 2010:e480873. doi:10.1155/2010/480873

Höfer M, Ali M, Sellmann J, Peil A (2014) Phenotypic evaluation and characterization of a collection of *Malus* species. Genet Resour Crop Evol 61:943–964. doi: 10.1007/s10722-014-0088-3

Hokanson SC, Lamboy WF, Szewc-McFadden AK, McFerson JR (2001) Microsatellite (SSR) variation in a collection of *Malus* (apple) species and hybrids. Euphytica 118:281–294. doi:10.1023/A:1017591202215

Hutabarat OS, Flachowsky H, Regos I, Miosic S, Kaufmann C, Faramarzi S, et al (2016) Transgenic apple plants overexpressing the *chalcone 3-hydroxylase* gene of *Cosmos sulphureus* show increased levels of 3-hydroxyphloridzin and reduced susceptibility to apple scab and fire blight. Planta 243:1213–1224. doi:10.1007/s00425-016-2475-9

Hvattum E (2002) Determination of phenolic compounds in rose hip (*Rosa canina*) using liquid chromatography coupled to electrospray ionisation tandem mass spectrometry and diode-array detection. Rapid Commun Mass Spectrom RCM 16:655–662. doi:10.1002/rcm.622

Hyson DA (2011) A comprehensive review of apples and apple components and their relationship to human health. Adv Nutr Int Rev J 2:408–420. doi:10.3945/an.111.000513

Ibdah M, Berim A, Martens S, Valderrama ALH, Palmieri L, Lewinsohn E, Gang DR (2014) Identification and cloning of an NADPH-dependent hydroxycinnamoyl-CoA double bond reductase involved in dihydrochalcone formation in *Malus* × *domestica* Borkh. Phytochem 107:24–31

Khan MA, Olsen KM, Sovero V, Kushad MM, Korban SS (2014) Fruit quality traits have played critical roles in domestication of the apple. Plant Genome 7:1–18. doi:10.3835/plantgenome2014.04.0018

Khan SA, Chibon P-Y, de Vos RCH, Schipper BA, Walraven E, Beekwilder J, et al (2012) Genetic analysis of metabolites in apple fruits indicates an mQTL hotspot for phenolic compounds on linkage group 16. J Exp Bot 63:2895–2908. doi:10.1093/jxb/err464

Kumar S, Garrick DJ, Bink MC, Whitworth C, Chagné D, Volz RK (2013) Novel genomic approaches unravel genetic architecture of complex traits in apple. BMC Genomics 14:393-2164-14–393

Lozano L, Iglesias I, Micheletti D, Troggio M, Kumar S, Volz RK, Allan AC, Chagné D, Gardiner SE (2014) Feasibility of genome-wide association analysis using a small single nucleotide polymorphism panel in an apple breeding population segregating for fruit skin color. J Am Soc Hortic Sci 139:619–626

Ma B, Liao L, Peng Q, Fang T, Zhou H, Korban SS, Han Y (2017) Reduced representation genome sequencing reveals patterns of genetic diversity and selection in apple. J Integr Plant Biol 59:190–204. doi:10.1111/jipb.12522

McDougall GJ (2017) Phenolic-enriched foods: sources and processing for enhanced health benefits. Proc Nutr Soc 76:163–171. doi:10.1017/S0029665116000835

Nikiforova SV, Cavalieri D, Velasco R, Goremykin V (2013) Phylogenetic analysis of 47 chloroplast genomes clarifies the contribution of wild species to the domesticated apple maternal line. Mol Biol Evol 30:1751–1760. doi:10.1093/molbev/mst092

Ramírez-Ambrosi M, López-Márquez DM, Abad-García B, Dapena E, Berrueta LÁ, Gallo B (2015) Comparative study of phenolic profile of fruit and juice samples of a progeny of 'Meana' × 'Florina' from an Asturian cider apple breeding program. Eur Food Res Technol 241:769–784. doi:10.1007/s00217-015-2502-2

Rivière C (2016) Dihydrochalcones: Occurrence in the plant kingdom, chemistry and biological activities. In: Atta-ur-Rahman (ed) Studies in Natural Products Chemistry, 1st edn. Elsevier, pp 253–381

Robinson J, Harris S, Juniper B (2001) Taxonomy of the genus *Malus* Mill. (Rosaceae) with emphasis on the cultivated apple, *Malus domestica* Borkh. Plant Systematics and Evolution 226:35–58

Stenlid G (1968) On the physiological effects of phloridzin, phloretin and some related substances upon higher plants. Physiol Plant. doi:10.1111/j.1399-3054.1968.tb07314.x

Turner A, Chen S-N, Joike MK, Pendland SL, Pauli GF, Farnsworth NR (2005) Inhibition of uropathogenic *Escherichia coli* by cranberry juice: a new antiadherence assay. J Agric Food Chem 53:8940–8947. doi:10.1021/jf052035u

Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, et al (2010) The genome of the domesticated apple (*Malus* × *domestica* Borkh.). Nat Genet 42:833–839. doi:10.1038/ng.654

Volk GM, Chao CT, Norelli J, Brown SK, Fazio G, Peace C, McFerson J, Zhong G-Y, Bretting P (2015a) The vulnerability of US apple (*Malus*) genetic resources. Genet Resour Crop Evol 62:765–794. doi:10.1007/s10722-014-0194-2

Volk GM, Henk AD, Baldo A, Fazio G, Chao CT, Richards CM (2015b) Chloroplast heterogeneity and historical admixture within the genus *Malus*. Am J Bot 102:1198–1208. doi:10.3732/ajb.1500095

Wang J, Huang Y, Li K, Chen Y, Vanegas D, McLamore ES, Shen Y (2016) Leaf Extract from *Lithocarpus polystachyus* Rehd. Promote Glycogen Synthesis in T2DM Mice. PLoS ONE 11:e0166557. doi:10.1371/journal.pone.0166557

Williams A, Jarrett J (1975) Hybridization of *Malus*. In: Report - Long Ashton Research Station 1974. University of Bristol, Bristol, Eng, p 44

Williams AH (1961) Dihydrochalcones of *Malus* species. J Chem Soc 155:4133–4136. doi:10.1039/JR9610004133

Xiao Z, Zhang Y, Chen X, Wang Y, Chen W, Xu Q, Li P, Ma F (2017) Extraction, identification, and antioxidant and anticancer tests of seven dihydrochalcones from *Malus* 'Red Splendor' fruit. Food Chem 231:324–331. doi:10.1016/j.foodchem.2017.03.111

Yahyaa M, Ali S, Davidovich-Rikanati R, Ibdah M, Shachtier A, Eyal Y, Lewinsohn E, Ibdah M (2017) Characterization of three chalcone synthase-like genes from apple (*Malus* x *domestica* Borkh.). Phytochem 140:125–133. doi:10.1016/j.phytochem.2017.04.022

Yahyaa M, Davidovich-Rikanati R, Eyal Y, Sheachter A, Marzouk S, Lewinsohn E, Ibdah M (2016) Identification and characterization of UDP-glucose:Phloretin 4′-O-glycosyltransferase from *Malus* x *domestica* Borkh. Phytochem 130:47–55. doi:10.1016/j.phytochem.2016.06.004

Zhao Y, Li X, Zeng X, Huang S, Hou S, Lai X (2014) Characterization of phenolic constituents in *Lithocarpus polystachyus*. Anal Methods 6:1359–1363. doi:10.1039/C3AY41288A

Zhou K, Hu L, Li P, et al (2017) Genome-wide identification of glycosyltransferases converting phloretin to phloridzin in *Malus* species. Plant Science 265:131–145. doi: 10.1016/j.plantsci.2017.10.003

Zhu C, Gore M, Buckler ES, Yu J (2008) Status and prospects of association mapping in plants. Plant Genome 1:5–20. doi:10.3835/plantgenome2008.02.0089

CHAPTER 2

GENETIC DIVERSITY OF DIHYDROCHALCONE CONTENT IN *MALUS* GERMPLASM

ABSTRACT

Dihydrochalcones, beneficial phenolic compounds, are abundant in *Malus* Mill. species, particularly in vegetative tissues and seeds. Phloridzin (phloretin 2′-*O*-glucoside) is the primary dihydrochalcone in most *Malus* species including cultivated apple, *Malus* ×*domestica* Borkh. A few species contain sieboldin (3-hydroxyphloretin 4′-*O*-glucoside) or trilobatin (phloretin 4′-*O*-glucoside) in place of phloridzin, and interspecific hybrids may contain combinations of phloridzin, sieboldin, and trilobatin. Proposed health benefits of phloridzin include anti-cancer, antioxidant, and anti-diabetic properties, suggesting the potential to breed apples for nutritional improvement. Sieboldin and trilobatin are being investigated for nutritional value and unique chemical properties. Although some of the biosynthetic steps of dihydrochalcones are known, little is known about the extent of variation within *Malus* germplasm. This research explores the genetic diversity of leaf dihydrochalcone content and composition in *Malus* germplasm. Dihydrochalcone content was measured using high performance liquid chromatography (HPLC) from leaf samples of 377 accessions, representing 50 species and interspecific hybrids from the USDA-Agricultural Research Service (ARS) National Plant Germplasm System *Malus* collection. Within the accessions sampled, 284 accessions contained phloridzin as the primary dihydrochalcone, one had only trilobatin, two had phloridzin and trilobatin, 36 had sieboldin and trilobatin, and 54 had all three. Leaf phloridzin content ranged from 17.3 to 113.7 mg/g with a heritability of 0.76 across all accessions. Beyond the potential of dihydrochalcones for breeding

purposes, dihydrochalcone composition may be indicative of hybridization or species misclassification.

INTRODUCTION

Dihydrochalcones are a key characteristic of *Malus* Mill. species, distinguishing them from closely related genera; dihydrochalcones are absent in *Pyrus* L. and *Prunus* L. Dihydrochalcones also differentiate *Malus* species; phloridzin is the primary dihydrochalcone in most species, but a few species replace phloridzin with either sieboldin (3-hydroxyphloretin-4′-*O*-glucoside) or trilobatin (phloretin-4′-*O*-glucoside) (Challice 1974; Williams 1961, 1982). Williams (1961) suggested that trilobatin originates from Eastern Mediterranean species *Malus trilobata* (Poir) C.K. Schneid. and sieboldin from Southeast Asian species of the series Sieboldiane: *Malus sargentii* Rehder, *Malus toringo* (Siebold) de Vriese, and *Malus zumi* (Matsum.) Rehder. Interspecific hybridization with these species results in co-dominant inheritance of sieboldin and trilobatin with phloridzin (Hunter 1975; Williams 1966). Dihydrochalcone content also varies between species and could be indicative of a taxon (Tang et al. 2015) along with sugar and phenolic composition (Anastasiadi et al. 2017).

Understanding dihydrochalcones in *Malus* may offer unique insight into the evolution and physiology of apple. Phloridzin, trilobatin, and sieboldin are found in a few other plant genera, but only in *Malus* do these dihydrochalcones accumulate in high concentrations (66 to 90% of total phenolic content) in vegetative tissues and seeds (Gosch et al. 2010b). Presumably, phloridzin represents the ancestral dihydrochalcone type, whereas sieboldin and trilobatin represent divergent evolution in a few geographically isolated species. Suggested physiological roles of phloridzin include disease resistance and growth regulation, but results are inconclusive

(Dare and Hellens 2013; Dare et al. 2013b, 2017; Dugé de Bernonville et al. 2010; Gaucher et al. 2013a; Gosch et al. 2009).

The health benefits of phloridzin are well documented. Notable are the anti-diabetic properties of phloridzin, which were used in the development of sodium-dependent glucose transporter 2 (SGLT-2) inhibitors to treat diabetes (Grempler et al. 2012; Rosenwasser et al. 2013; Zinman et al. 2015). Sieboldin forms complexes with iron with strong affinity, reinforcing its antioxidant properties and potential for iron seclusion and/or storage (Dugé de Bernonville et al. 2010; Gaucher et al. 2013b). Trilobatin is of interest as a sweetener and for treatment of diabetes through α-glucosidase inhibition (Dong et al. 2012; Yahyaa et al. 2016). Other dihydrochalcones in *Malus* species include, phloretin-2′-O-xyloglucoside and 3-hydroxyphloridzin (3-hydroxyphloretin-2′-O-glucoside). These compounds are found in lower concentrations in vegetative tissues (De Paepe et al. 2015; Fromm et al. 2012; Jakobek and Barron 2016) and may also have physiological or nutritive roles (Eichenberger et al. 2017; Hutabarat et al. 2016; Xiao et al. 2017), but are not included in this study.

The USDA-ARS National Plant Germplasm System (NPGS) *Malus* collection in Geneva, NY is one of the largest and most diverse collections of apple genetic resources, maintaining over 4,000 unique field accessions of apple cultivars, hybrids, and wild species; *Malus* ×*domestica* Borkh. (n=1374), the cultivated apple, and *Malus sieversii* (Ledeb.) M. Roem. (n=789) the wild progenitor of apple, represent a large portion of the collection. Worldwide, there are approximately 38 wild and 21 cultivated *Malus* species and hybrids across six sections (Volk et al. 2015a). Little is known about the relationships between *Malus* species; research has emphasized the relationships between *M.* ×*domestica* and its wild progenitors (Cornille et al. 2012; Nikiforova et al. 2013; Velasco et al. 2010). Interspecific hybridization and admixture is

common among *Malus* species, making taxonomy and phylogenetic relationships difficult to resolve (Cornille et al. 2013; Robinson et al. 2001; Volk et al. 2015b). Admixture from *M.* ×*domestica* has a particularly negative impact on wild *Malus* populations in terms of species conservation (Cornille et al. 2015; Wagner et al. 2014). Distinguishing a pure species from a hybrid is challenging in *ex situ* collections, but use of molecular markers with phenotypic and geographic information can facilitate distinction (Gross et al. 2012a; Höfer et al. 2014; Khan et al. 2014a). However, these methods should be applied cautiously to underrepresented species, where there is no baseline for phenotypic or genetic diversity and little information on their origin (Höfer et al. 2014). Most of the *Malus* species in the NPGS collection are represented by fewer than 20 accessions and when questions arise concerning their classification or hybrid-status it is difficult to resolve them. Currently, there are 372 NPGS accessions listed as *Malus* hybrid or *Malus* species, categories designated for accessions of uncertain pedigree or taxonomic status.

To better understand the chemical diversity within the *Malus* collection, we surveyed 377 accessions for leaf phenolic content, with an emphasis on phloridzin, sieboldin, and trilobatin. We identified five distinct dihydrochalcone profiles across fifty species and interspecific hybrids. Dihydrochalcones were the most abundant compound isolated and the strongest factor in differentiating accessions on the basis of phenolic content. Some accessions did not fit species-specific chemical description, but fit the chemical profile of an interspecific hybrid (phloridzin detected with sieboldin or trilobatin) or that of another species. Other accessions were outliers in dihydrochalcone content relative to other accessions of the same taxon. Further examination of genetic and morphological variation helped to verify these discrepancies, and provided a stronger understanding of dihydrochalcone content variation in these taxa.

MATERIALS AND METHODS

Plant Material

The USDA-ARS Plant Genetic Resources Unit (PGRU) in Geneva, NY maintains 2,800 permanent *Malus* accessions on semi-dwarf 'EMLA 7' rootstock (E7), including a core collection of 258 accessions to represent the genetic diversity of the collection (Gross et al. 2013). The core collection was selected by the US Apple Crop Germplasm Committee based on fruit quality traits, disease and pest resistance, and morphological diversity (Potts et al. 2012). Duplicates of these accessions were grafted on 'Budagovsky 9' (B9) dwarfing rootstock. In mid-August (8/10 to 8/20) from 2013 to 2015, mature leaves were sampled from each accession in the core collection on B9 and E7 rootstocks. In 2015, 134 accessions were added to our study based on taxonomic or genotypic information; 123 accessions were from the permanent collection, ten from the US National Arboretum, and one from the Arnold Arboretum of Harvard University (*M. trilobata* 127-2009). In total, 377 accessions were sampled. Table 2.1 lists the species and the number of accessions included, with *Malus* taxonomy based on information from the Germplasm Resource Information Network (GRIN; USDA 2014). All tissues were frozen in liquid nitrogen, ground to a powder using a GenoGrinder 2000 (Spex SamplePrep, Metuchen, NJ, USA), and stored at –80℃.

Table 2.1 Number of *Malus* species[a] and hybrids from the USDA National Plant Germplasm System *Malus* collection sampled for dihydrochalcone content

Species	No.	Species	No.	Species	No.
M. angustifolia (Aiton) Michx.	2	*Malus* hybrid	47	*M. sieversii* (Ledeb.) M. Roem.	28
M. ×*arnoldiana* (Rehder) Sarg. ex Rehder	2	*M. ioensis* (Alph. Wood) Britton	5	*M. sikkimensis* (Wenz.) Koehne ex Schneid.	4
M. ×*asiatica* Nakai	3	*M. kansuensis* (Batalin) Schneid.	6	*M.* ×*soulardii* (LH Bailey) Britton	1
M. ×*astracanica* hort. ex Dum. Cours.	1	*M.* ×*magdeburgensis* Hartwig	1	*M. spectabilis* (Aiton) Borkh.	5
M. ×*atrosanguinea* (hort. ex Späth) Schneid.	2	*M. mandishurica* (Maxim.) Kom. ex Skvortsov	1	*Malus* spp.	12
M. baccata (L.) Borkh.	10	*M.* ×*micromalus* Makino	15	*M.* ×*sublobata* (Dippel) Rehder	4
M. brevipes (Rehder) Rehder	1	*M.* ×*moerlandsii* Door.	1	*M. sylvestris* (L.) Mill.	7
M. coronaria (L.) Mill.	3	*M. ombrophila* Hand.-Mazz.	2	*M. toringo* (Siebold) Siebold ex de Vriese	17
M. ×*dawsoniana* Rehder	1	*M. orientalis* Uglitzk.	12	*M. toringoides* (Rehder) Hughes	4
M. ×*domestica* Borkh.	66	*M. orthocarpa* Lavallée ex anon.	1	*M. transitoria* (Batalin) Schneid.	3
M. doumeri (Bois) A. Chev.	1	*M.* ×*platycarpa* Rehder	1	*M. trilobata* (Poir.) Schneid.	1
M. florentina (Zuccagni) Schneid.	2	*M. prattii* (Hemsl.) Schneid.	3	*M. tschonoskii* (Maxim.) Schneid.	3
M. floribunda Siebold ex Van Houtte	10	*M. prunifolia* (Willd.) Borkh.	35	*M. yunnanensis* (Franch.) Schneid.	3
M. fusca (Raf.) Schneid.	4	*M. pumila* Mill.	2	*M. zhaojiaoensis* N.G. Jiang	2
M. halliana Koehne	3	*M.* ×*purpurea* (A. Barbier) Rehder	3	*M. zumi* (Matsum.) Rehder	6
M. ×*hartwigii* Koehne	2	*M.* ×*robusta* (Carriere) Rehder	3		
M. honanensis Rehder	1	*M. sargentii* Rehder	18		
M. hupehensis (Pamp.) Rehder	7	*M.* ×*scheideckeri* (L.H. Bailey) Spath ex Zabel	1		

[a] Taxonomy based on GRIN (USDA 2014)

High Performance Liquid Chromatography

Phenolics were extracted from a 25 mg of a pooled sample of five leaves, in 1.5 ml of 70% methanol (99.9%) with 2% formic acid (98–100%). Two technical replicate samples were shaken in a GenoGrinder at 1,100 strokes/min for 10 minutes then centrifuged at 10,000*g* for 10

minutes. The supernatant was filtered with a 0.20 µm syringe filter. Compounds were separated on a Hitachi LaChrom Elite (San Jose, CA, USA) HPLC equipped with (L-2450) 2-channel diode array detector, and analyzed with EZChrom Elite 3.1.4 (Santa Clara, CA, USA) software. An Inertsil ODS-3 column (4.6×250 mm; 5 µm, GL Sciences Inc., Tokyo, Japan) at 30℃ and a Nova-Pack C_{18} guard column (3.9×20mm, 4 µm, Waters, Milford, MA, USA) were used in the separation. Leaf phenolics were separated in a 15% acetonitrile (99.9%), 10% formic acid (98–100%) isocratic mobile phase at a flow rate of 1.0 ml/min, with a sample injection volume of 10 µl and a run time of 34 minutes. Spectral scans were from 210 to 600 nm, with dihydrochalcones quantified at 280 nm. Acetonitrile, formic acid, methanol, and phloridzin standard were purchased from Sigma-Aldrich (St. Louis, MO, USA). Sieboldin and trilobatin standards were purchased from Extrasythèse (Genay Cedex, France). Only phloridzin, sieboldin, and trilobatin were analyzed among potential phloretin derivatives. Standard curves for phloridzin, sieboldin, and trilobatin were calculated using EZChrom Elite 3.1.4 from seven concentrations ranging from 7.8 µg to 1,000 µg with $R^2 > 0.99$. Sample values were expressed in mg/g fresh weight.

Genotyping-by-Sequencing

DNA from 50 mg leaf tissue was extracted using DNeasy 96 Plant Kits (Qiagen, Valencia CA, USA) from 1,947 *Malus* accessions including 280 accessions used in this study. DNA was isolated and sequenced from 1,890 of these accessions in collaboration between USDA-ARS PGRU and the Apple Diversity Group of Dalhousie University (Migicovsky et al. 2016). Single nucleotide polymorphisms (**SNP**) were identified through genotyping-by-sequencing (**GBS**) (Elshire et al. 2011). Library construction with *Ape*KI restriction enzyme and Illumina HiSeq 2000 (Illumina Inc., San Diego CA, USA) sequencing were completed at the Genomics Diversity Facility, Cornell University, Ithaca, NY, USA. Sequence data were

processed using the GBS analysis pipeline in TASSEL 5.0 (Glaubitz et al. 2014). Sequence tags were aligned to the *M.* ×*domestica* Whole Genome Reference Assembly v.2 (https://www.rosaceae.org/) using BWA software (Li and Durbin 2009) under default parameters. Master tag list was created by trimming reads to 64 bp (not including barcodes), and filtering reads with missing data and tags with less than five counts. VCFtools (Danecek et al. 2011) was used to filter data; 320 individuals were removed for mean depth below 18 (15^{th} percentile); sites were filtered for read depth of 10, minor allele frequency of 0.10, < 20% missing data, and QC less than 98. The majority of genotyped accessions were *M.* ×*domestica* (n=1,141); to balance data, this number was reduced to 342 by filtering accessions of *M.* ×*domestica* with high genetic similarity based on PCA.

Data Analysis

R 3.3.2 (R Core Team 2016) and JMP® (SAS Institute Inc., Cary, NC) version 12.0.1 were used for statistical analysis. Metabolite data were log transformed to meet normality assumption. Data were fit using the linear model:

$Y_{ijk} = \mu + G_i + Y_j + R_k + GY_{ij} + GR_{ik} + YR_{jk} + e_{ijk}$, where Y_{ijk} is the phenolic measurement on, genotype *i*, year *j* and rootstock *k*; μ is the overall mean; G_i is the random effect of genotype *i*; Y_j is the random effect of year *j* and R_k the random effect of rootstock *k*; GY_{ij}, GR_{ik}, and YR_{jk}, are the random interaction effects of genotype *i*, year *j* and rootstock *k*; and e_{ijk} is the residual error. The model is framed with the following assumptions:

$G_i \sim N(0, \sigma_G^2), Y_j \sim N(0, \sigma_Y^2), R_k \sim N(0, \sigma_R^2), GY_{ij} \sim N(0, \sigma_{GY}^2), GR_{ik} \sim N(0, \sigma_{GR}^2), YR_{jk} \sim N(0, \sigma_{YR}^2)$, and $e_{ijk} \sim N(0, \sigma_e^2)$.

Principal component analysis was completed using R package 'SNPrelate' (Zheng et al. 2012) on 817 *Malus* accessions with a set of 5,131 SNPs filtered for LD ($r^2 < 0.20$).

Phylogenetic trees were constructed using 'SNPhylo' (Lee et al. 2014) and R package 'phangorn' (Schliep 2011). Descriptor data for 336 accessions used in this research were downloaded from GRIN (USDA, 2014). Variables and individuals with more than 15% missing data were excluded; seven additional variables were dropped with a measure of sampling adequacy (Cerny and Kaiser 1977) lower than 0.65. Remaining sixteen descriptors included four continuous (fruit length and width; pedicel length; and fruit peel overcolor intensity), seven ordinal (calyx appearance and persistence; fruit flesh firmness; pedicel thickness; fruit texture and weight; and harvest season), and five nominal (fruit flesh color and texture; fruit peel ground and overcolor, and fruit size uniformity). Quantitative variables were centered and scaled. Gower distance matrix was calculated on sixteen variables across 327 accessions using R package 'cluster' (Maecheler et al. 2016) and hierarchical clustering was performed using the complete linkage method in R (R Core Team 2016).

RESULTS

Dihydrochalcone Profiles

Each accession was placed into one of five profiles (Table 2.2) depending on the major dihydrochalcones present in leaf samples. Of the 249 core collection accessions sampled, 227 had phloridzin only (**P**); 11 had sieboldin and trilobatin (**ST**); and 11 had sieboldin, phloridzin, and trilobatin (**SPT**). Profile ST included four taxa: *Malus* ×*micromalus* Makino, *M. sargentii*, *Malus* ×*sublobata* (Dippel) Rehder, and *M. toringo*. Profile SPT included seven taxa: *Malus* ×*arnoldiana* (Rehder) Sarg. ex Rehder, *Malus* ×*atrosanguinea* (hort. ex Späth) C.K. Schneid., *Malus floribunda* Siebold ex Van Houtte, *Malus kansuensis* (Batalin) C.K. Schneid., *M.* ×*micromalus*, *Malus prunifolia* (Willd.) Borkh., *M.* ×*sublobata*, and *Malus* hybrids 'White

Angel' and 'Prairie Fire'. To identify additional dihydrochalcone diversity, 106 accessions were sampled based on the nine taxa in profiles ST and SPT and taxa with little to no representation in the core collection. These taxonomical selections identified additional ST and SPT accessions, as well as a new group, trilobatin only (**T**) in *M. trilobata*. Neighbor-joining trees developed from GBS data available on 817 *Malus* accessions identified additional dihydrochalcone diversity; many of the ST and SPT accessions clustered together, with twenty-two accessions of unknown dihydrochalcone profile from which eight additional SPT accessions and two accessions with phloridzin and trilobatin (**PT**) were identified. Eighteen taxa with accessions in profiles T, PT, ST, or SPT are listed in Table 2.3. Not all combinations were observed in the germplasm; sieboldin and phloridzin (**SP**), and sieboldin only (**S**) profiles were absent. Consistent with other reports, sieboldin was not found independent of trilobatin (Zhou et al. 2017).

Table 2.2 Dihydrochalcone profiles identified with HPLC across *Malus* germplasm sets

DHC Group	Core Accessions (n= 249)	Taxonomical Selections (n=106)	Genotype Selections (n=22)	Total Accessions (N=377)
P	227	45	12	284
T	0	1	0	1
PT	0	0	2	2
ST	11	25	0	36
SPT	11	35	8	54

Table 2.3 Number of accessions per taxon[a] in each dihydrochalcone profile

Species	Section (Series)/Pedigree	P	T	PT	ST	SPT
M. ×arnoldiana	*M. baccata* × *M. floribunda*	-	-	-	-	2
M. ×atrosanguinea	*M. halliana* × *M. toringo*	-	-	-	-	2
M. baccata	Gymnomeles	8	-	1	-	1
M. floribunda	*M. toringo* × *M. baccata*	1	-	-	-	9
M. hupehensis	Gymnomeles	6	-	-	-	1
Malus hybrid		40	-	1	-	6
M. kansuensis	Sorbomalus (Kansuenses)	5	-	-		1
M. ×micromalus	*M. baccata* × *M. spectablilis*	6	-	-	6	3
M. prunifolia	Malus	24	-	-	-	11
M. ×purpurea	*M. ×atrosanguinea* × *M. pumila*	2	-	-	-	1
M. sargentii	Sorbomalus (Sieboldiane)	-	-	-	11	7
M. ×scheideckeri	*M. floribunda* × *M. prunifolia*	1	-	-	-	-
M. sikkimensis	Gymnomeles	3	-	-	-	1
M. ×scheideckeri	*M. floribunda* × *M. prunifolia*	1	-	-	-	-
Malus spp.		10	-	-	-	2
M. ×sublobata	*M. prunifolia* × *M. toringo*	-	-	-	1	3
M. toringo	Sorbomalus (Sieboldiane)	-	-	-	15	2
M. trilobata	Eriolobus	-	1	-	-	-
M. zumi	Sorbomalus (Sieboldiane)	-	-	-	3	3

[a] 34 taxa omitted were profile P exclusively

Quantitative Variation of Dihydrochalcones

The mixed linear model explained a high proportion of the variation of dihydrochalcone content. Across all accessions, broad-sense heritability estimates of phloridzin, sieboldin, and trilobatin were 0.76, 0.85, and 0.88, respectively. Year and genotype × year were significant but explained less than 2% of the variation, except for genotype × year for phloridzin, which explained 6.6%. There were no significant rootstock or rootstock × genotype effects between the two in this study, although significant rootstock effects on phenolic content in apple were reported in other studies (Kviklys et al. 2014; Polat and Yildirim 2016). Seasonal correlations from 2013 to 2015 varied from 0.74 to 0.96 and rootstock correlations varied from 0.86 to 0.96 for phloridzin, sieboldin, trilobatin and total dihydrochalcones. These high correlations suggest

that leaf dihydrochalcones are stable from year to year and between rootstocks. Guo et al. (2016) reported a seasonal correlation of $r = 0.70$ for phloridzin in fruit from *Malus* germplasm.

Phloridzin content ranged from 17.3 to 113.7 mg/g with a mean of 50.9±15.6 mg/g across all samples. Species with less than 3 accessions or less than two seasons of data were not included in assessing within species variation (Table 2.4). Accessions of *Malus baccata* (L.) Borkh., *Malus* ×*robusta* (Carriere) Rehder, *M.* ×*domestica*, and *M. sieversii* had the lowest average dihydrochalcone content. Accessions of *Malus halliana* Koehne, *Malus transitoria* (Batalin) C.K. Schneid, *Malus hupehensis* (Pamp.) Rehder, and *Malus fusca* (Raf.) C.K. Schneid. had the highest average dihydrochalcone content. The observed variation was consistent with a previous study of dihydrochalcones in *Malus* leaves (Tang et al. 2015). Heritability estimates for phloridzin content in subgroups *M. domestica*, *Malus* hybrid, *M. prunifolia*, and *M. sieversii* were 0.42, 0.57, 0.73, and 0.70, respectively.

Table 2.4 Mean of phloridzin (Phl), sieboldin (Sie), trilobatin (Tri) and range of total dihydrochalcone content in mg/g from leaf samples across *Malus* species.

Species	No.	Phl	Sie	Tri	Total DHC
M. angustifolia	2	69.4	*	*	52.4–86.4
M. ×*arnoldiana*	2	36.1	9.6	3.5	42.3–56.2
M. ×*asiatica*	3	65.2	*	*	54.9–79.4
M. ×*astracanica*	1	36.6	*	*	
M. ×*atrosanguinea*	2	31.5	23.7	8.9	60.3–68.1
M. baccata	6	42.8	*	*	28.2–55.1
M. brevipes	1	64.8	*	*	
M. coronaria	3	57.3	*	*	55.8–58.4
M. ×*dawsoniana*	1	70.4	*	*	
M. ×*domestica*	65	48.5	*	*	34.1–66.9
M. florentina	2	60.9	*	*	54.9–66.8
M. floribunda	10	36.8	21.2	8.4	44.2–78.3
M. fusca	4	90.8	*	*	80.2–97.7
M. halliana	3	75.1	*	*	54.2–95.2
M. ×*hartwigii*	1	41.9	*	*	
M. honanensis	1	101.2	*	*	
M. hupehensis	6	83.8	*	*	66.1–113.7

Table 2.4 (Continued)

Species	No.	Phl	Sie	Tri	Total DHC
Malus hybrid	39	46.9	28.1	1.5	25.6–65.5
M. ioensis	5	57.9	*	*	44.6–87.7
M. kansuensis	6	53.1	18.4	7.5	39.7–70.5
M. magdeburgensis	1	54.7	*	*	
M. mandishurica	1	50.2	*	*	
M. ×micromalus	15	44.7	42.0	7.9	32.0–97.5
M. ×moerlandsii	1	61.0	*	*	
M. ombrophila	2	59.9	*	*	54.3–65.6
M. orientalis	12	54.9	*	*	42.4–61.7
M. orthocarpa	1	39.8	*	*	
M. platycarpa	1	70.6	*	*	
M. prattii	3	72.1	*	*	48.9–97.5
M. prunifolia	33	49.2	14.5	7.9	35.2–90.3
M. pumila	2	43.6	*	*	42.4–44.9
M. ×purpurea	3	44.5	1.3	13.6	31.8–59.8
M. ×robusta	3	45.8	*	*	45.4–46.2
M. sargentii	17	18.8	42.6	3.6	44.7–70.3
M. ×scheideckeri	1	45.0	*	*	
M. sieversii	28	51.8	*	*	33.8–68.1
M. sikkimensis	1	63.4	*	*	
M. ×soulardii	1	49.3	*	*	
M. spectabilis	5	48.0	*	*	41.1–55.8
Malus spp.	11	57.0	15.6	7.2	32.7–89.8
M. ×sublobata	4	27.3	20.6	9.9	37.9–58.1
M. sylvestris	7	52.6	*	*	44.3–68.9
M. toringo	11	26.7	53.6	5.3	41.4–85.4
M. toringoides	4	69.2	*	*	60.2–76.1
M. transitoria	3	77.7	*	*	51.3–101.3
M. tschonoskii	3	47.5	*	*	33.4–54.8
M. yunnanensis	3	64.7	*	*	56.2–73.0
M. zhaojiaoensis	2	52.1	*	*	49.5–54.7
M. zumi	3	32.3	47.6	8.8	39.4–91.1

* Not detected

Genetic Diversity

Principal component analysis was used to determine relationships between 817 *Malus* accessions. The first PCs explain 20.5% and 10.2% of the total variation between accessions (Figure 2.1). PC1 primarily separates *M. ×domestica* and *M. sieversii* from other wild and hybrid

accessions and PC2 separates *M. sieversii* from *M.* ×*domestica*. Three main groupings are observed. Group one (bottom left of Figure 2.1) contains 286 individuals, consisting primarily of *M.* ×*domestica* (n=252) and some accessions of *Malus orientalis* Uglitzk. (n=1), *Malus pumila* Mill. (n=1), *M. sieversii* (n=16), *Malus sylvestris* (L.) Mill. (n=3), and *Malus* hybrids (n=13). Group two (top left of Figure 2.1) consists primarily of *M. sieversii* (n=164), with a few accessions of *M.* ×*domestica* (n=40), *M. orientalis* (n=12), *M. prunifolia* (n=1), and *M. pumila* (n=4). Several of the *M.* ×*domestica* accessions of this group originated from Eastern Europe, Middle East, and Russia, which differ from cultivars from western Europe or the United States in the genetic contribution of wild *Malus* species (Cornille et al. 2012; Gharghani et al. 2009). On the far right of Figure 2.1, group three contains 99 accessions, including wild Asian and North American species. Between these three groups are 201 accessions that do not group relative to their classification. A small grouping of *M. sylvestris* accessions are at the bottom of Figure 2.1. Accessions of dihydrochalcone profiles T, ST, PT, and SPT fall into group three and are scattered throughout the ungrouped accessions with no clear pattern of dihydrochalcone distribution. These results were similar to the structure of 117 diverse *Malus* accessions observed by Duan et al. (2017).

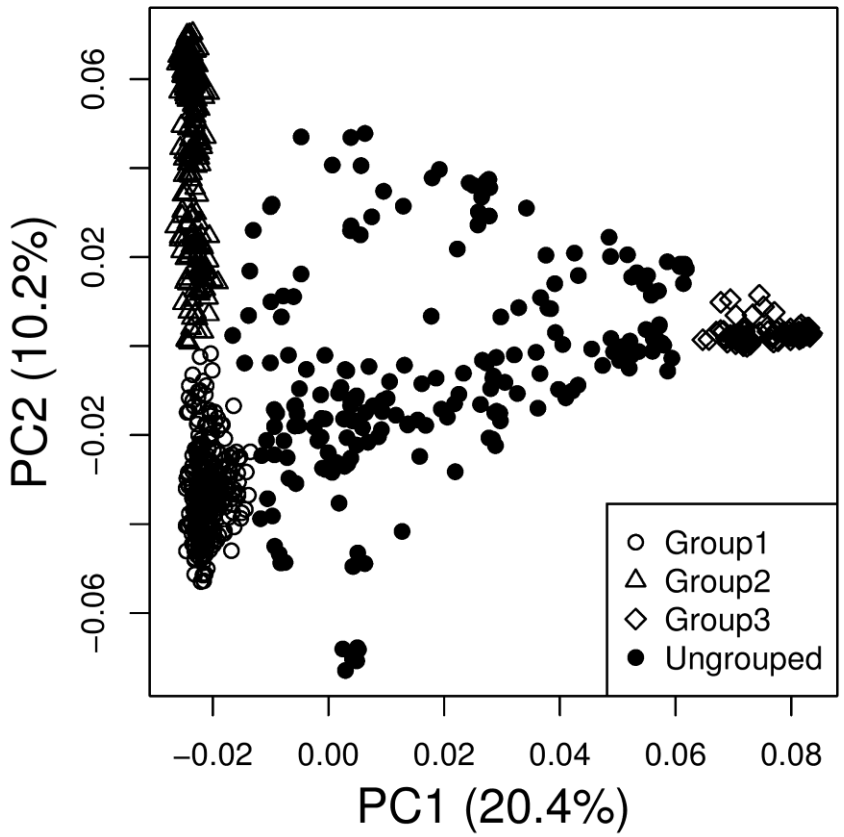


Figure 2.1 Principal component analysis of 817 *Malus* accessions using Genotyping-by-Sequencing data.

Hierarchical Clustering Analysis

Eight clusters were identified using hierarchical clustering based on morphological data for 327 accession (Figure 2.2). Cluster three was the largest and most chemically diverse, whereas clusters five, seven, and eight consisted only of phloridzin genotypes (Table 2.5). Clusters one, two, and three were primarily wild *Malus* species with smaller fruit, longer and thinner pedicels, astringent flavors, and variable calyx appearance and persistence. Clusters seven and eight were predominantly *M.* ×*domestica* and *Malus* hybrids, including a few *M. sieversii* accessions, with larger, sweeter fruit, and persistent calyces. Clusters four, five, and six included progenitor species, *M. orientalis* and *M. sieversii*, and a few *M.* ×*domestica* and *Malus*

hybrids. A single *M.* ×*domestica* accession, 'Russian sldg.' (PI 589312) grouped in cluster one. *M. prunifolia* accessions were identified in clusters one, three, four, five, and six; Figure 2.3 illustrates fruit morphological variation for these accessions.

Table 2.5 Dihydrochalcone profiles distributed across clusters based on *Malus* descriptor data from GRIN

Cluster	No. in DHC Profile		
	P	ST	SPT
1	31	0	5
2	8	10	7
3	64	18	26
4	18	0	2
5	10	0	0
6	32	0	1
7	80	0	0
8	15	0	0

Figure 2.2 Dendrogram of 327 accessions from the USDA *Malus* collection in Geneva, NY based on hierarchical clustering analysis of fruit descriptors using complete linkage method. Axis measures distance between accessions. PI number included after species. *M.* ×*domestica* (blue) and *M. prunifolia* (red) are highlighted. Species abbreviations: *M. angustifolia* (ang), *M.* ×*arnoldiana* (arn), *M.* ×*asiatica* (asi), *M.* ×*astracanica* (ast), *M.* ×*atrosanguinea* (atr), *M. baccata* (bac), *M. brevipes* (bre), *M. coronaria* (cor), *M. dawsoniana* (daw), *M.* ×*domestica* (dom), *M. florentina* (flore), *M. floribunda* (flori), *M. fusca* (fus), *M. halliana* (hal), *M.* ×*hartwigii* (har), *M. honanensis* (hon), *M. hupehensis* (hup). *Malus* hybrid (hyb), *M. ioensis* (ioe), *M. kansuensis* (kan), *M.* ×*magdeburgensis* (mag), *M. mandishurica* (man), *M.* ×*micromalus* (mic), *M. moerlandsii* (moe), *M. ombrophila* (omb), *M. orientalis* (ori), *M. orthocarpa* (ort), *M.* ×*platycarpa* (pla), *M. prattii* (pra), *M. prunifolia* (pru), *M. pumila* (pum), *M.* ×*purpurea* (pur), *M.* ×*robusta* (rob), *M. sargentii* (sar), *M.* ×*scheideckeri* (sch), *M. sikkimensis* (sik), *M.* ×*soulardii* (sou), *M. spectabilis* (spe), *Malus* spp. (spp), *M.* ×*sublobata* (sub), *M. sylvestris* (syl), *M. toringo* (tor), *M. toringoides* (bhu), *M. transitoria* (tra), *M. tschonoskii* (tsc), *M. yunnanensis* (yun), *M. zhaojiaoensis* (zha), and *M. zumi* (zum).

Figure 2.2 (Continued)

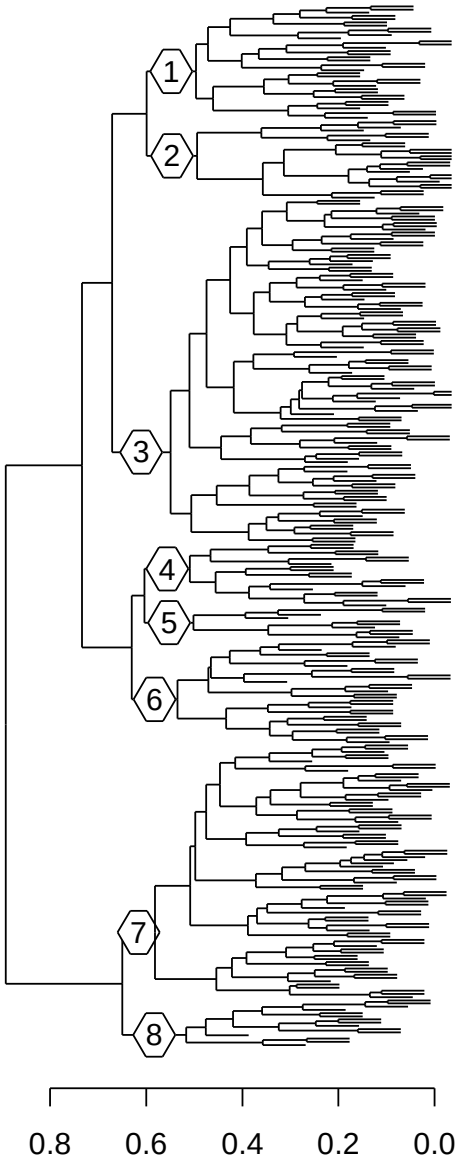


Figure 2.2 (Continued)

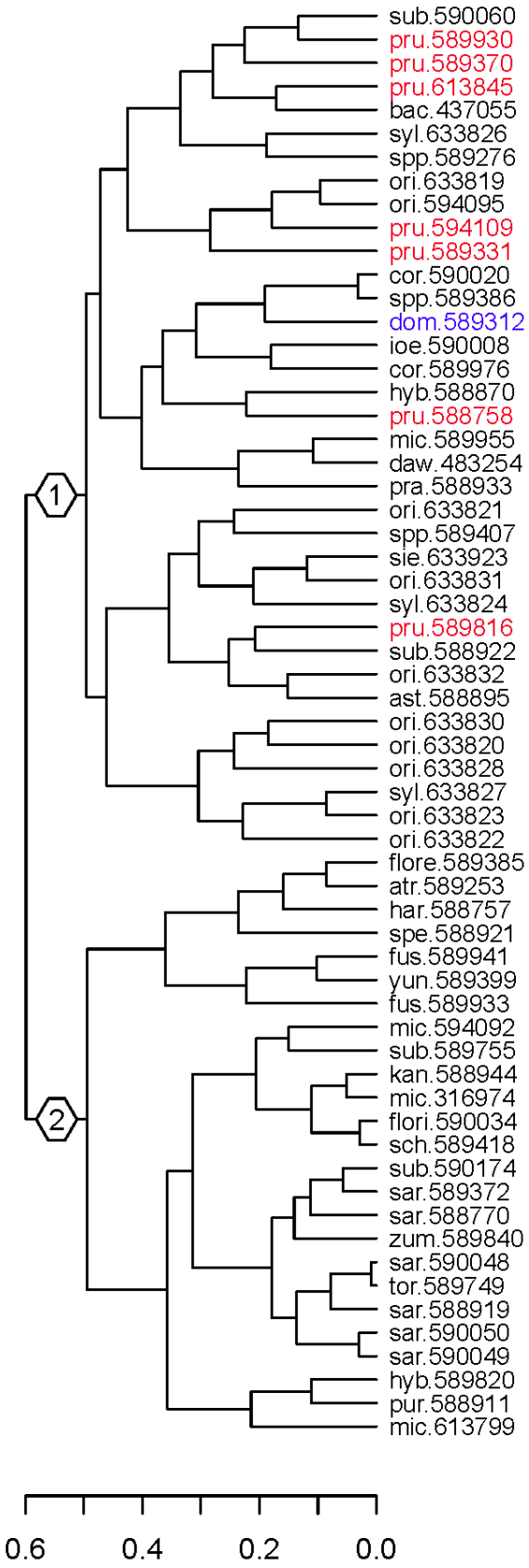


Figure 2.2 (Continued)

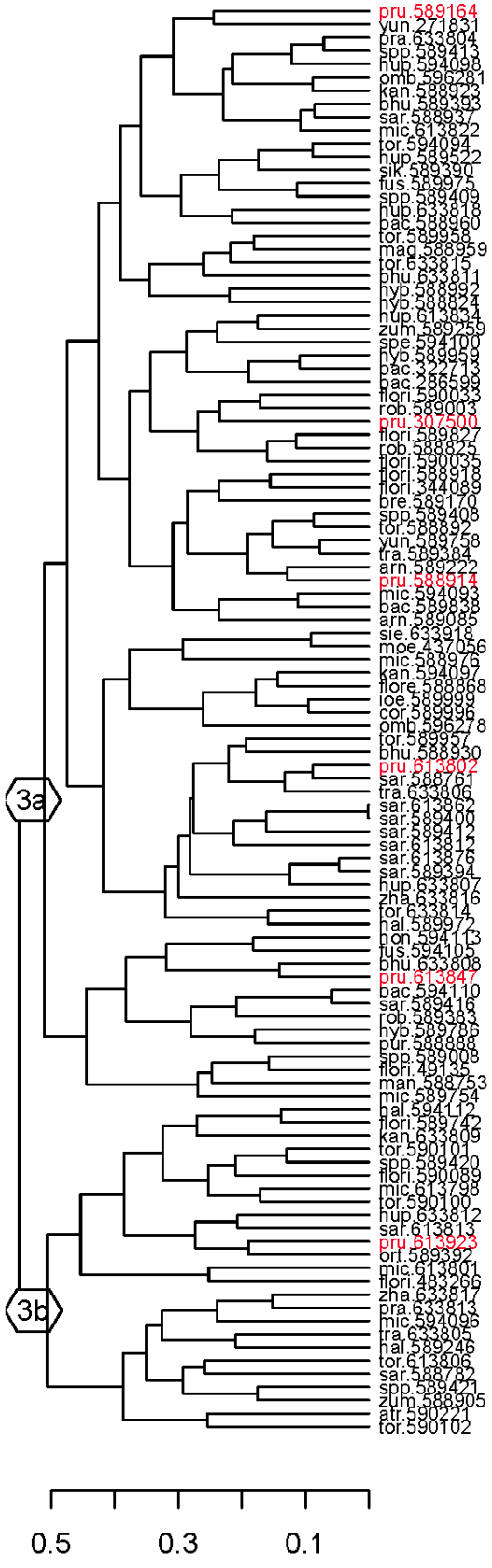


Figure 2.2 (Continued)

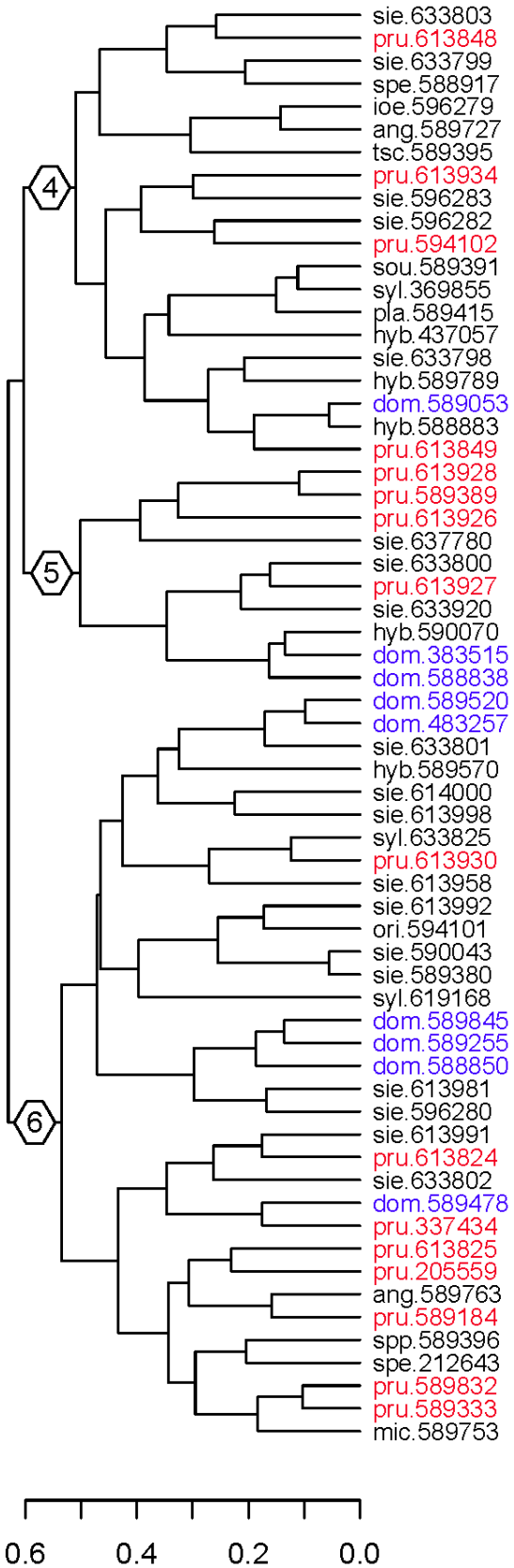


Figure 2.2 (Continued

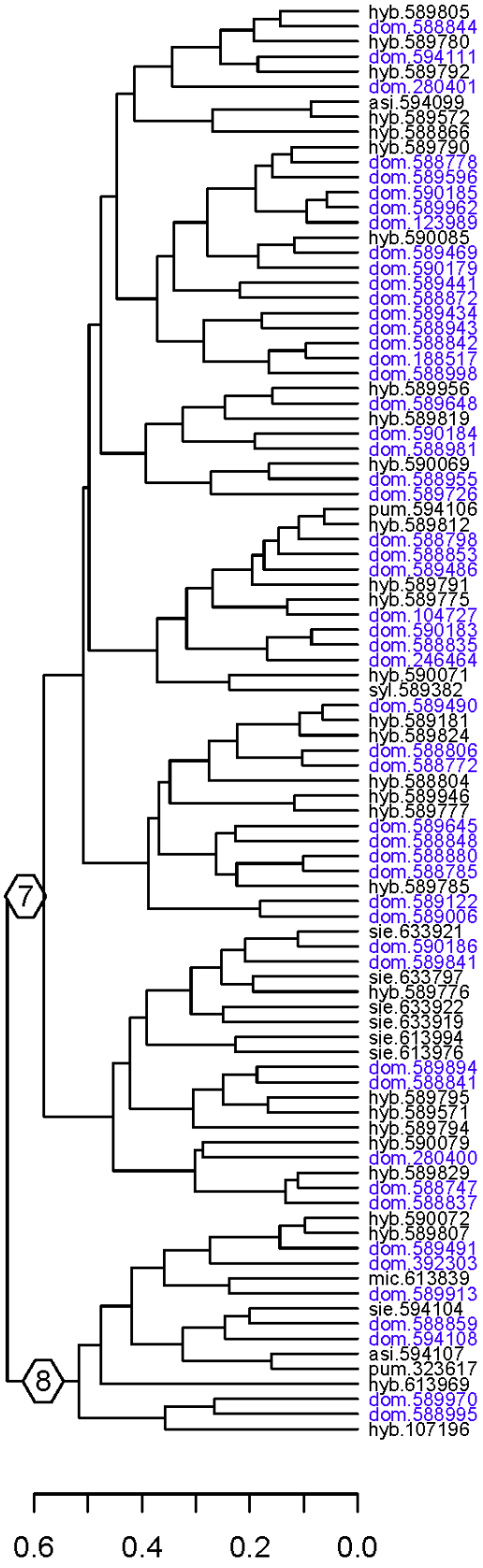


Figure 2.3 Fruit morphology variation of twenty-eight accessions classified as *M. prunifolia* in the USDA *Malus* collection in Geneva, NY on 2.54-cm^2 grids.

DISCUSSION

Dihydrochalcones are a target for commercialization and breeding, because of their unique properties. Eichenberger et al. (2017) and Zhang et al. (2017) discuss potential of metabolic engineering for industrial production of dihydrochalcones. However, given the high accumulation of dihydrochalcones in *Malus* species, plant based extractions could be more efficient (Xiao et al. 2017); this germplasm research identified several accessions that accumulate multiple dihydrochalcones in higher concentrations than most commercial apple cultivars. Additionally, new apple cultivars could be nutritionally enhanced through introgression of sieboldin and trilobatin and through selecting for higher concentrations of dihydrochalcones in fruit of dessert or cider apples. Although *M.* ×*domestica* had one of the lower average values of

dihydrochalcone content, there was a broad range (34.1–66.9 mg/g) and a moderate heritability (H^2 = 0.42) across sixty-three genotypes. This variation may be adequate for genetic improvement of dihydrochalcone content. Though there are species with higher dihydrochalcone content or greater variation for dihydrochalcone profiles, these species have significantly lower fruit quality. The range of dihydrochalcone content (33.8–68.1 mg/g) and higher heritability (H^2 = 0.70) in *M. sieversii* suggests it could be a genetic resource for dihydrochalcone content with somewhat better fruit quality traits than other wild species. Phenolic composition may be more varied and abundant in wild *Malus* species and heirloom cultivars of *M.* ×*domestica* relative to modern cultivars (Anastasiadi et al. 2017; De Paepe et al. 2015; Jakobek and Barron 2016; Khan et al. 2014b). However, high fruit quality as well as high phenolic content was identified in advanced breeding selections suggesting that breeding for increased phenolic content need not rely on introgression from wild species or older cultivars (Volz and McGhie 2011). Morphological analysis shows that sieboldin and trilobatin tend to be associated with small, low quality fruit, which poses some challenges for integrating these dihydrochalcones into new, high quality cultivars. Some of these genetic resources could also be utilized for other traits; *M. floribunda* 821 (SPT) and *M.* ×*atrosanguinea* 804 (SPT) are sources of apple scab (*Venturia inaequalis* (Cooke) G. Winter) resistance, and *Malus* hybrid 'White Angel' (SPT) and *M. zumi* (ST and SPT) are sources for powdery mildew (*Podosphaera leucotricha* (Ellis & Everh.) E.S. Salmon) resistance (Brown 2012). Though introgression from wild species into *M.* ×*domestica* is hampered by long juvenility and restricted backcrossing due to self-incompatibility (Le Roux et al. 2012), utilization of germplasm resources will enhance the genetic diversity of commercial apples (Noiton and Alspach 1996).

There is much to learn about the mechanism of hyper-accumulation of dihydrochalcones in *Malus*, and their potential role in the evolution and physiology of *Malus*. Dihydrochalcones, mainly phloridzin, have been explored for potential roles in apple. Dare et al. (2013b, 2017) observed aberrant morphology in *M.* ×*domestica* 'Royal Gala' and altered auxin transport when independently silencing chalcone synthase and phloretin glycosyltransferase *UGT88F1*. Several groups researched a connection between dihydrochalcone content and disease resistance but results are often inconclusive. Dugé de Bernonville et al. (2011) found no correlation between dihydrochalcone composition and resistance to fire blight in rootstock cultivar 'MM106' (profile P), ornamental cultivar *Malus* 'Evereste' (profile SPT), or their F1 progeny. However, Hutabarat et al. (2016) observed a significant correlation between reduced apple scab and fire blight susceptibility and increased hydroxy-phloridzin content in transgenic *M.* ×*domestica* 'Pinova' overexpressing *chalcone 3-hydroxylase*. It was posited that the utilization of phloretin oxidative products contributes to resistance rather than total dihydrochalcone content (Gosch et al. 2010a); Gaucher et al. (2013a) demonstrated that phloridzin is converted to phloretin differently in 'MM106' (fire blight susceptible) and 'Evereste' (fire blight resistant). Utilizing the variation in dihydrochalcone profiles may help to better understand the evolution and physiology of dihydrochalcones in *Malus*. For example, if phloridzin regulates auxin transport in *Malus*, one could assess whether sieboldin or trilobatin have similar roles in *Malus* species when they replace phloridzin.

Williams (1966) described dihydrochalcone variation in *Malus* species and hybrids and reported that trilobatin originated from *M. trilobata*, and sieboldin originated from *M. toringo* (formerly *M. sieboldii* (Regel) Rehder) and *M. sargentii*, which some consider a variety of *M. toringo*. Interspecific hybridization with these species produces offspring with different

combinations of phloridzin, trilobatin, or sieboldin (Hunter 1975; Williams 1966). Hybrids with *M. toringo*, such as *M.* ×*atrosanguinea* (*M. halliana* × *M. toringo*), *M. floribunda* (*M. toringo* × *M. baccata*), and *M.* ×*sublobata* (*M. prunifolia* × *M. toringo*) have accessions in profiles ST and SPT. Many secondary hybrids with *M. floribunda*, such as *M.* ×*arnoldiana* (*M. baccata* × *M. floribunda*) and *M.* ×*scheideckeri* (*M. floribunda* × *M. prunifolia*) also have accessions in profiles PT, ST, and SPT. The relationship of these hybrids is supported by genetic and morphological similarities. Inconsistencies to this pattern are highlighted in *M.* ×*micromalus* (*M. baccata* × *Malus spectabilis* (Aiton) Borkh.), with accessions across four profiles (P, PT, ST, and SPT), and *M. prunifolia*, with accessions in profiles P or SPT. Additionally, there were seven *M. sargentii*, two *M. toringo*, and three *M. zumi* accessions in profile SPT. Assuming that sieboldin and trilobatin originate from *M. trilobata* and from the series Sieboldiane, misclassification or admixture could account for these inconsistencies. Admixture was reported in several studies of the NPGS *Malus* collection. In a study of *M.* ×*domestica*, *M. sieversii*, *M. sylvestris*, and *M. orientalis*, evidence of admixture was found in 43 of 406 accessions, with a greater proportion of admixture in wild species than in *M.* ×*domestica* (Gross et al. 2012a). Khan et al. (2014a) identified potential admixture of *M. baccata*, *M.* ×*micromalus*, *M. orientalis*, and *M. sieversii* in five accessions of *M. prunifolia*; and one of four *M.* ×*micromalus* accessions showed signs of admixture from *M. orientalis* or *M. sieversii*.

Correct taxonomic classification impacts trait discovery and utilization of genetic resources. Although most *Malus* species in the NPGS collection lack adequate sample size ($n < 10$), there are sufficient accessions of *M. prunifolia* ($n = 35$, Table 2.6) to investigate genetic and phenotypic diversity. *M. prunifolia* is a valuable resource for ornamental and rootstock breeding for abiotic stress resistance, including drought and salt tolerance, and ease of propagation (Bao et

al. 2016; Fu et al. 2013; Moriya et al. 2015; Wan et al. 2011). *M. prunifolia* is often cited as one of the progenitors to *M.* ×*domestica* (Velasco et al. 2010; Volk et al. 2015b). Volk et al. (2015b) identified seven chloroplast haplotypes in twenty-seven accessions of *M. prunifolia* from the NPGS collection. *M. prunifolia* haplotype M21 (three accessions) was shared with *M. floribunda*; haplotype M15 (seven accessions) was shared with *M. baccata* and *M. mandishurica*; haplotypes M26 (five accessions) and M32 (one accession) were shared with *M.* ×*domestica* and *M. orientalis*; and haplotype M28 (ten accessions) was shared with *M. orientalis* and *M. sieversii*, and included seven *M. prunifolia* of profile SPT. In a separate study, two accessions of *M. prunifolia* from the Fondazione Edmund Mach apple collection clustered closely with *M.* floribunda, *M. sargentii*, and *M. toringo* based on chloroplast sequences (Nikiforova et al. 2013). We observed a high amount of genetic and morphological variation for accessions classified as *M. prunifolia*. Several accessions were more closely related to *M.* ×*domestica* based on genetic markers and fruit morphology; and others are associated with wild *Malus* species based on our analysis and that of Hokanson et al. (2001) using SSR markers. Some accession passport information provided in GRIN also suggests misclassification; *M. prunifolia* PI 589816 (profile SPT) was formerly listed as *M.* ×*arnoldiana*, a designation better fitting of its genetic, morphological, and chemical descriptions. Therefore, the inconsistencies between taxonomy and dihydrochalcone profile in *M. prunifolia* is likely due to the admixture prevalent in *Malus*, or incorrect classification rather than unique dihydrochalcone variation within a species, and classification of these accessions should be reassessed. NPGS accessions of *M. prunifolia* should be reassessed on the basis of available genetic, morphological, and chemical data. Accessions from China are primarily of the phloridzin only type, while in Japanese accessions, SPT is more common, suggesting greater introgression.

Table 2.6 Summary of 35 *M. prunifolia* accessions from the *Malus* collection in Geneva, NY

Plant Intro. Number	Plant Name	Geographical Origin	DHC Profile	Descriptor Cluster	Chloroplast Haplotype[a]
205559	Sikora (Type 1)		P	6	M15
307500	Golden Gem		SPT	3	M28
337434	B-87		P	6	M26
588758	Vinti Sdlg. 101D		P	1	
588914	Fastigiata		P	3	M15
588994	Rinkii	East Asia	P		M15
589164	DE 229	Northeast Asia	SPT	3	M28
589184	Dulcis		P	6	M26
589331	Xanthocarpa NA3604		P	1	M21
589333	DE 279	Northeast Asia	SPT	6	
589370	GMAL 1769		SPT	1	M28
589389	Macrocarpa		P	5	M26
589816	19651		SPT	1	
589832	Xanthocarpa		P	6	M21
589930	Nagano	Japan	SPT	1	M28
589932	M0-84	Japan	SPT		M28
594102	GMAL 1890		P	4	M26
594103	Inuringo	Japan	P		M28
594109	Microcarpa		P	1	M21
613802	Rinkii		P	3	M15
613824	GMAL 1522.a1	South Korea	P	6	
613825	GMAL 1522.d1	South Korea	P	6	
613845	GMAL 1890.b1		P	1	M26
613846	Ringo Asami	Japan	SPT		M28
613847	Ringo Asami	Japan	P	3	M28
613848	Inuringo	Japan	SPT	4	M28
613849	Inuringo	Japan	SPT	4	M28
613923	GMAL 3227.c	China	P	3	
613926	GMAL 3231.h1	China	P	5	613926
613927	GMAL 3232.g1	China	P	5	M15
613928	GMAL 3233.a	China	P	5	M32
613930	GMAL 3236.e1	China	P	6	M15
613934	GMAL 3241.e1	China	P	4	
666192	Seishi	Japan	SPT		
GMAL5082	287	China	P		

[a] Volk et al. 2015b

CONCLUSIONS

There is much to discover concerning dihydrochalcone accumulation in *Malus*, and this research suggests that genetic factors contribute significantly to leaf dihydrochalcone content. Though phloridzin is the dominant dihydrochalcone in *Malus* species, we identified unique

accessions with combinations of sieboldin, phloridzin, and trilobatin. Further research into the genetics and biochemistry of dihydrochalcones should include some of these accessions.

Along with genetic and morphological data, chemical data can assist in species classification. *Malus* species identification is difficult due to interspecific hybridization and a significant challenge for germplasm management and utilization. Inaccurate classification can result in false representation of within species variation, or neglect of valuable genetic resources for genetic and crop improvement programs. Though some NPGS *Malus* germplasm was collected from plant expeditions, much of the collection originated from disparate sources (arboretums, breeding programs, and conservatories). Accessions often have little passport information on origin, pedigree, or geographical location, which would help validate their classification. Many cultivated hybrids, such as *M.* ×*micromalus*, are challenging because they are not found in the wild. In *Malus* species, dihydrochalcone content and composition is particularly useful to distinguish species; accessions with dihydrochalcone content outside the expected range may indicate misclassification, and presence of phloridzin with sieboldin and trilobatin indicates interspecific hybridization.

Of particular interest for future germplasm preservation in the NPGS collection are *M. prunifolia*, *M. trilobata*, and *M.* ×*micromalus*. Our results suggest that accessions of *M.* ×*micromalus* and *M. prunifolia* in the Geneva collection do not adequately represent these taxa. *M. prunifolia* has ornamental value and potential for rootstock and scion breeding and was included on the threatened species list as data deficient. Capturing the diversity of this species should be a high priority. *M. trilobata* is also unique within the genus *Malus*, as the sole member of the section Eriolobus, with unique leaf morphology and chemistry for ornamental breeding.

Currently, the Geneva collection has a single accession and less than a hundred seeds to represent this species.

REFERENCES

Anastasiadi M, Mohareb F, Redfern SP, Berry M, Simmonds MSJ, Terry LA (2017) Biochemical profile of heritage and modern apple cultivars and application of machine learning methods to predict usage, age, and harvest season. J Agric Food Chem. doi:10.1021/acs.jafc.7b00500

Bao L, Li K, Liu Z, Han M, Zhang D (2016) Characterization of the complete chloroplast genome of the Chinese crabapple *Malus prunifolia* (Rosales: Rosaceae: Maloideae). Conserv Genet Resour 8:227–229. doi:10.1007/s12686-016-0540-0

Brown S (2012) Apple. In: Badenes ML, Byrne DH (eds) Fruit Breeding. Springer, New York, pp 329–367

Cerny BA, Kaiser HF (1977) A study of a measure of sampling adequacy for factor-analytic correlation matrices. Multivar Behav Res 12:43–47. doi:10.1207/s15327906mbr1201_3

Challice JS (1974) Rosaceae chemotaxonomy and the origins of the Pomoideae. Bot J Linn Soc 69:239–259. doi:10.1111/j.1095-8339.1974.tb01629.x

Cornille A, Feurtey A, Gelin U, Ropars J, Misvanderbrugge K, Gladieux P, Giraud T (2015) Anthropogenic and natural drivers of gene flow in a temperate wild fruit tree: a basis for conservation and breeding programs in apples. Evol Appl 8:373–384. doi:10.1111/eva.12250

Cornille A, Gladieux P, Giraud T (2013) Crop-to-wild gene flow and spatial genetic structure in the closest wild relatives of the cultivated apple. Evol Appl 6:737–748. doi:10.1111/eva.12059

Cornille A, Gladieux P, Smulders MJ, Roldán-Ruiz I, Laurens F, Cam BL, et al (2012) New insight into the history of domesticated apple: secondary contribution of the European wild apple to the genome of cultivated varieties. PLoS Genet 8:e1002703. doi:10.1371/journal.pgen.1002703

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al (2011) The variant call format and VCFtools. Bioinformatics 27:2156–2158. doi:10.1093/bioinformatics/btr330

Dare A, Hellens R (2013) RNA interference silencing of *CHS* greatly alters the growth pattern of apple (*Malus* x *domestica*). Plant Signal Behav 8:e25033. doi:10.4161/psb.25033

Dare AP, Tomes S, Jones M, McGhie TK, Stevenson DE, Johnson RA, Greenwood DR, Hellens RP (2013) Phenotypic changes associated with RNA interference silencing of chalcone synthase in apple (*Malus* × *domestica*). Plant J 74:398–410. doi:10.1111/tpj.12140

Dare AP, Yauk Y-K, Tomes S, McGhie TK, Rebstock RS, Cooney JM, Atkinson RG (2017) Silencing a phloretin-specific glycosyltransferase perturbs both general phenylpropanoid

biosynthesis and plant development. Plant J Cell Mol Biol 91:237–250. doi:10.1111/tpj.13559

De Paepe D, Valkenborg D, Noten B, Servaes K, Diels L, Loose MD, Van Droogenbroeck B, Voorspoels S (2015) Variability of the phenolic profiles in the fruits from old, recent and new apple cultivars cultivated in Belgium. Metabolomics 11:739–752. doi:10.1007/s11306-014-0730-2

Dong H-Q, Li M, Zhu F, Liu F-L, Huang J-B (2012) Inhibitory potential of trilobatin from *Lithocarpus polystachyus* Rehd against α-glucosidase and α-amylase linked to type 2 diabetes. Food Chem 130:261–266. doi:10.1016/j.foodchem.2011.07.030

Duan N, Bai Y, Sun H, Wang N, Ma Y, Li M, et al (2017) Genome re-sequencing reveals the history of apple and supports a two-stage model for fruit enlargement. Nat Commun. doi:10.1038/s41467-017-00336-7

Dugé de Bernonville T, Gaucher M, Guyot S, Durel C-E, Dat JF, Brisset M-N (2011) The constitutive phenolic composition of two *Malus* × *domestica* genotypes is not responsible for their contrasted susceptibilities to fire blight. Environ Exp Bot 74:65–73. doi:10.1016/j.envexpbot.2011.04.019

Dugé de Bernonville T, Guyot S, Paulin J-P, Gaucher M, Loufrani L, Henrion D, et al (2010) Dihydrochalcones: Implication in resistance to oxidative stress and bioactivities against advanced glycation end-products and vasoconstriction. Phytochem 71:443–452. doi:10.1016/j.phytochem.2009.11.004

Eichenberger M, Lehka BJ, Folly C, Fischer D, Martens S, Simón E, Naesby M (2017) Metabolic engineering of *Saccharomyces cerevisiae* for de novo production of dihydrochalcones with known antioxidant, antidiabetic, and sweet tasting properties. Metab Eng 39:80–89. doi:10.1016/j.ymben.2016.10.019

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple Genotyping-by-Sequencing (GBS) approach for high diversity species. PLoS ONE 6:e19379. doi:10.1371/journal.pone.0019379

Fromm M, Bayha S, Carle R, Kammerer DR (2012) Characterization and quantitation of low and high molecular weight phenolic compounds in apple seeds. J Agric Food Chem 60:1232–1242. doi:10.1021/jf204623d

Fu M, Li C, Ma F (2013) Physiological responses and tolerance to NaCl stress in different biotypes of *Malus prunifolia*. Euphytica 189:101–109. doi:10.1007/s10681-012-0721-1

Gaucher M, Dugé de Bernonville T, Guyot S, Dat JF, Brisset M-N (2013a) Same ammo, different weapons: Enzymatic extracts from two apple genotypes with contrasted susceptibilities to fire blight (*Erwinia amylovora*) differentially convert phloridzin and phloretin *in vitro*. Plant Physiol Biochem 72:178–189. doi:10.1016/j.plaphy.2013.03.012

Gaucher M, Dugé de Bernonville T, Lohou D, Guyot S, Guillemette T, Brisset M-N, Dat JF (2013b) Histolocalization and physico-chemical characterization of dihydrochalcones: Insight into the role of apple major flavonoids. Phytochem 90:78–89. doi:10.1016/j.phytochem.2013.02.009

Gharghani A, Zamani Z, Talaie A, Oraguzie NC, Fatahi R, Hajnajari H, Wiedow C, Gardiner SE (2009) Genetic identity and relationships of Iranian apple (*Malus* × *domestica* Borkh.) cultivars and landraces, wild *Malus* species and representative old apple cultivars based on simple sequence repeat (SSR) marker analysis. Genet Resour Crop Evol 56:829–842. doi:10.1007/s10722-008-9404-0

Glaubitz JC, Casstevens TM, Lu F, Harriman J, Elshire RJ, Sun Q, Buckler ES (2014) TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. PLoS ONE 9:e90346. doi:10.1371/journal.pone.0090346

Gosch C, Halbwirth H, Kuhn J, Miosic S, Stich K (2009) Biosynthesis of phloridzin in apple (*Malus domestica* Borkh.). Plant Sci 176:223–231. doi:10.1016/j.plantsci.2008.10.011

Gosch C, Halbwirth H, Schneider B, Hölscher D, Stich K (2010a) Cloning and heterologous expression of glycosyltransferases from *Malus* x *domestica* and *Pyrus communis*, which convert phloretin to phloretin 2′-O-glucoside (phloridzin). Plant Sci 178:299–306. doi:10.1016/j.plantsci.2009.12.009

Gosch C, Halbwirth H, Stich K (2010b) Phloridzin: biosynthesis, distribution and physiological relevance in plants. Phytochem 71:838–843. doi:10.1016/j.phytochem.2010.03.003

Grempler R, Thomas L, Eckhardt M, Himmelsbach F, Sauer A, Sharp DE, et al (2012) Empagliflozin, a novel selective sodium glucose cotransporter-2 (SGLT-2) inhibitor: characterisation and comparison with other SGLT-2 inhibitors. Diabetes Obes Metab 14:83–90. doi:10.1111/j.1463-1326.2011.01517.x

Gross BL, Henk AD, Forsline PL, Richards CM, Volk GM (2012) Identification of interspecific hybrids among domesticated apple and its wild relatives. Tree Genet Genomes 8:1223–1235. doi:10.1007/s11295-012-0509-4

Gross BL, Volk GM, Richards CM, Reeves PA, Henk AD, Forsline PL, Szewc-McFadden A, Fazio G, Chao CT (2013) Diversity captured in the USDA-ARS National Plant Germplasm System apple core collection. J Am Soc Hortic Sci 138:375–381

Guo S, Guan L, Cao Y, Li C, Chen J, Li J, Liu G, Li S, Wu B (2016) Diversity of polyphenols in the peel of apple (*Malus* sp.) germplasm from different countries of origin. Int J Food Sci Technol 51:222–230. doi:10.1111/ijfs.12994

Höfer M, Ali M, Sellmann J, Peil A (2014) Phenotypic evaluation and characterization of a collection of *Malus* species. Genet Resour Crop Evol 61:943–964. doi:10.1007/s10722-014-0088-3

Hokanson SC, Lamboy WF, Szewc-McFadden AK, McFerson JR (2001) Microsatellite (SSR) variation in a collection of *Malus* (apple) species and hybrids. Euphytica 118:281–294. doi:10.1023/A:1017591202215

Hunter LD (1975) Phloridzin and apple scab. Phytochem 14:1519–1522. doi:10.1016/0031-9422(75)85343-X

Hutabarat OS, Flachowsky H, Regos I, Miosic S, Kaufmann C, Faramarzi S, et al (2016) Transgenic apple plants overexpressing the *chalcone 3-hydroxylase* gene of *Cosmos sulphureus* show increased levels of 3-hydroxyphloridzin and reduced susceptibility to apple scab and fire blight. Planta 243:1213–1224. doi:10.1007/s00425-016-2475-9

Jakobek L, Barron AR (2016) Ancient apple varieties from Croatia as a source of bioactive polyphenolic compounds. J Food Compos Anal 45:9–15. doi:10.1016/j.jfca.2015.09.007

Khan MA, Olsen KM, Sovero V, Kushad MM, Korban SS (2014a) Fruit quality traits have played critical roles in domestication of the apple. Plant Genome 7:1–18. doi:10.3835/plantgenome2014.04.0018

Khan SA, Tikunov Y, Chibon P-Y, Maliepaard C, Beekwilder J, Jacobsen E, Schouten HJ (2014b) Metabolic diversity in apple germplasm. Plant Breed 133:281–290. doi:10.1111/pbr.12134

Kviklys D, Liaudanskas M, Janulis V, Viškelis P, Rubinskienė M, Lanauskas J, Uselis N (2014) Rootstock genotype determines phenol content in apple fruits. Plant Soil Environ 60:234–240

Le Roux PM, Flachowsky H, Hanke M-V, Gessler C, Patocchi A (2012) Use of a transgenic early flowering approach in apple (*Malus* × *domestica* Borkh.) to introgress fire blight resistance from cultivar Evereste. Mol Breed 30:857–874. doi:10.1007/s11032-011-9669-4

Lee T-H, Guo H, Wang X, Kim C, Paterson AH (2014) SNPhylo: a pipeline to construct a phylogenetic tree from huge SNP data. BMC Genomics 15:162. doi:10.1186/1471-2164-15-162

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinforma Oxf Engl 25:1754–1760. doi:10.1093/bioinformatics/btp324

Maecheler M, Rousseeuw P, Struyf A, Hubert M, Hornik K (2016) Cluster: Cluster analysis basics and extensions. R package version 2.0.5.

Migicovsky Z, Gardner KM, Money D, Sawler J, Bloom JS, Moffett P, et al (2016) Genome to phenome mapping in apple using historical data. Plant Genome. doi:10.3835/plantgenome2015.11.0113

Moriya S, Iwanami H, Haji T, Okada K, Yamada M, Yamamoto T, Abe K (2015) Identification and genetic characterization of a quantitative trait locus for adventitious rooting from apple hardwood cuttings. Tree Genet Genomes 11:59. doi:10.1007/s11295-015-0883-9

Nikiforova SV, Cavalieri D, Velasco R, Goremykin V (2013) Phylogenetic analysis of 47 chloroplast genomes clarifies the contribution of wild species to the domesticated apple maternal line. Mol Biol Evol 30:1751–1760. doi:10.1093/molbev/mst092

Noiton DAM, Alspach PA (1996) Founding clones, inbreeding, coancestry, and status number of modern apple cultivars. J Am Soc Hortic Sci 121:773–782

Polat D, Yildirim F (2016) Variation of phenolic content in leaves of 14 apple rootstocks with varying growth vigour. Ziraat Fakültesi Derg - Süleyman Demirel Üniversitesi 11:63–70

Potts SM, Han Y, Khan MA, Kushad MM, Rayburn AL, Korban SS (2012) Genetic diversity and characterization of a core collection of *Malus* germplasm using simple sequence repeats (SSRs). Plant Mol Biol Report 30:827–837. doi:10.1007/s11105-011-0399-x

R Core Team (2016) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Robinson J, Harris S, Juniper B (2001) Taxonomy of the genus *Malus* Mill. (Rosaceae) with emphasis on the cultivated apple, *Malus domestica* Borkh. Plant Syst Evol 226:35–58

Rosenwasser RF, Sultan S, Sutton D, Choksi R, Epstein BJ (2013) SGLT-2 inhibitors and their potential in the treatment of diabetes. Diabetes Metab Syndr Obes 6:453–467

Schliep KP (2011) phangorn: phylogenetic analysis in R. Bioinformatics 27:592–593. doi:10.1093/bioinformatics/btq706

Tang J, Tang L, Tan S, Zhou Z (2015) The study of variation of phloridzin content in six wild *Malus* species. J Food Nutr Res 3:146–151

USDA (2014) National genetic resources program. Germplasm resources information network (GRIN). USDA, ARS Natl. Germplasm Resources Laboratory, Beltsville, MD. http://www.ars-grin.gov/npgs/index.html. Accessed 5 May 2017

Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, et al (2010) The genome of the domesticated apple (*Malus* × *domestica* Borkh.). Nat Genet 42:833–839. doi:10.1038/ng.654

Volk GM, Chao CT, Norelli J, Brown SK, Fazio G, Peace C, McFerson J, Zhong G-Y, Bretting P (2015a) The vulnerability of US apple (*Malus*) genetic resources. Genet Resour Crop Evol 62:765–794. doi:10.1007/s10722-014-0194-2

Volk GM, Henk AD, Baldo A, Fazio G, Chao CT, Richards CM (2015b) Chloroplast heterogeneity and historical admixture within the genus *Malus*. Am J Bot 102:1198–1208. doi:10.3732/ajb.1500095

Volz RK, McGhie TK (2011) Genetic variability in apple fruit polyphenol composition in *Malus*× *domestica* and *Malus sieversii* germplasm grown in New Zealand. J Agric Food Chem 59:11509–11521. doi:10.1021/jf202680h

Wagner I, Maurer WD, Lemmen P, Schmitt HP, Wagner M, Binder M, Patzak P (2014) Hybridization and genetic diversity in wild apple (*Malus sylvestris* (L.) Mill) from various regions in Germany and from Luxembourg. Silvae Genet 63:81–93. doi:10.1515/sg-2014-0012

Wan Y, Li D, Zhao Z, Mei L, Han M, Schwaninger H, Fazio G (2011) The distribution of wild apple germplasm in Northwest China and its potential application for apple rootstock breeding. Ix Int Symp Integrating Canopy Rootstock Environ Physiol Orchard Syst 903:123–141. doi:10.17660/ActaHortic.2011.903.12

Williams AH (1961) Dihydrochalcones of *Malus* species. J Chem Soc 155:4133–4136. doi:10.1039/JR9610004133

Williams AH (1966) Dihydrochalcones. In: Swain T (ed) Comparative Phytochemistry. Academic Press, New York, pp 291–307

Williams AH (1982) Chemical evidence from the flavonoids relevant to the classification of *Malus* species. Bot J Linn Soc 84:31–39. doi:10.1111/j.1095-8339.1982.tb00358.x

Xiao Z, Zhang Y, Chen X, Wang Y, Chen W, Xu Q, Li P, Ma F (2017) Extraction, identification, and antioxidant and anticancer tests of seven dihydrochalcones from *Malus* 'Red Splendor' fruit. Food Chem 231:324–331. doi:10.1016/j.foodchem.2017.03.111

Yahyaa M, Davidovich-Rikanati R, Eyal Y, Sheachter A, Marzouk S, Lewinsohn E, Ibdah M (2016) Identification and characterization of UDP-glucose:Phloretin 4′-O-glycosyltransferase from *Malus* x *domestica* Borkh. Phytochem 130:47–55. doi:10.1016/j.phytochem.2016.06.004

Zhang T-J, Liang J-Q, Wei X-Y, Wang P-X, Xu Y, Pen S, Fan M-T (2017) Development of an enzymatic synthesis approach to produce phloridzin using *Malus* x *domestica* glycosyltransferase in engineered *Pichia pastoris* GS115. Process Biochem. doi:10.1016/j.procbio.2017.05.009

Zheng X, Levine D, Shen J, Gogarten SM, Laurie C, Weir BS (2012) A high-performance computing toolset for relatedness and principal component analysis of SNP data. Bioinformatics 28:3326–3328. doi:10.1093/bioinformatics/bts606

Zhou K, Hu L, Li P, et al (2017) Genome-wide identification of glycosyltransferases converting phloretin to phloridzin in *Malus* species. Plant Science 265:131–145.

Zinman B, Wanner C, Lachin JM, Fitchett D, Bluhmki E, Hantel S, et al (2015) Empagliflozin, cardiovascular outcomes, and mortality in type 2 diabetes. N Engl J Med 373:2117–2128. doi:10.1056/NEJMoa1504720

CHAPTER 3

LINKAGE AND ASSOCIATION ANALYSIS OF DIHYDROCHALCONES PHLORIDZIN, SIEBOLDIN, AND TRILOBATIN IN *MALUS*

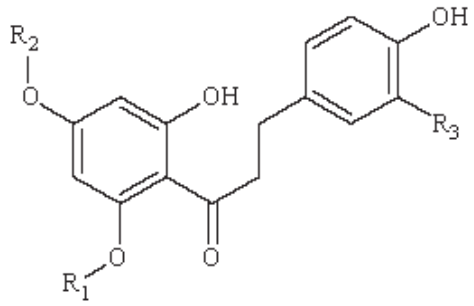
ABSTRACT

Dihydrochalcones (DHCs) are a distinctive characteristic of *Malus* species with phloridzin as the major DHC in most *Malus* species, including cultivated apple. DHCs in apple have unique chemical properties with commercial and nutritional value, and may yield important insights into the evolution and physiology of apple. A few species produce sieboldin and trilobatin instead of phloridzin, and interspecific hybridization may produce offspring with combinations of phloridzin, sieboldin, and trilobatin. Using *Malus prunifolia* PI 89816 as a common male parent, five F_1 populations were developed to understand the genetic basis of these DHCs in *Malus*. We measured DHC content in each population and observed segregation into five distinct DHC profiles, which fit a model for three independently segregating loci. QTL associated with DHC content were identified on linkage groups 7 and 8 of the *Malus* genome using linkage analysis with a cross of NY-152 by *M. prunifolia* PI 589816 and association mapping with a *Malus* germplasm collection. In addition to DHC segregation, we observed variation in the relative proportions of phloridzin, sieboldin, and trilobatin when found in combination. The QTL identified represent a critical step in understanding the genetic controllers of DHC content in *Malus*.

INTRODUCTION

Apples are a high value food crop with unique nutritional qualities, stemming primarily from production of phenolic compounds. Flavonoids, a class of phenolics, contribute significantly to the color and nutritional value of apple, of which dihydrochalcones (DHCs) are a minor subclass. Contrasted with the diversity of phenolic compounds (thousands described), approximately 256 DHCs have been described in a wide distribution of plant species, though few isolated from edible plants (Rivière 2016). Despite the rarity of DHCs, they are abundant in *Malus* species, and can represent up to 90% of the phenolic profile and 14% of the leaf dry weight (Gosch et al. 2010b). DHCs are absent or produced only in trace amounts in other food crops, including relatives within the Rosaceae. DHCs have diverse chemical properties valuable for industrial or nutritional purposes. Phloridzin (phloretin 2′-*O*-glucoside), derived from the aglycone precursor phloretin, is the prominent dihydrochalcone in *Malus* species, including cultivated apple, and has anti-diabetic effects through increased insulin sensitivity and inhibition of sodium glucose cotransporter-2 (SGLT-2) (Najafian et al. 2012). Additionally, phloridzin has antioxidant, anti-cancer, anti-inflammatory, and phytoestrogenic effects (Nair et al. 2014; Puel et al. 2005). In a few *Malus* species, phloridzin content is replaced entirely by phloretin derivatives, sieboldin (3-hydroxyphloretin-4′-*O*-glucoside) and trilobatin (phloretin-4′-*O*-glucoside), which also have unique chemical properties. Higher antioxidant activity was observed in sieboldin relative to phloridzin and trilobatin, with a capacity to inhibit vasoconstriction and formation of advanced glycation end-products associated with several degenerative diseases (Dugé de Bernonville et al. 2010; Xiao et al. 2017). Trilobatin, a sweet DHC, has anti-diabetic effects differing from phloridzin through α-glucosidase inhibition instead of SGLT-2 and reduction of chronic inflammation related to lipopolysaccharides (Dong et al. 2012; Fan et al. 2015).

Trilobatin is about 300 times sweeter than sucrose, with potential as a naturally occurring commercial sweetener (Rivière 2016). Phloretin, the precursor to these DHCs, also has potential industrial value and nutritional quality, although it is found in lower concentrations (Behzad et al. 2017). Figure 3.1 depicts the chemical structures of phloretin, phloridzin, sieboldin, and trilobatin.



Phloretin: R_1=H, R_2=H, R_3=H
Phloridzin: R_1=Glucose, R_2=H, R_3=H
Trilobatin: R_1=H, R_2=Glucose, R_3=H
Sieboldin: R_1=H, R_2=Glucose, R_3=OH

Figure 3.1 Chemical structure of the dihydrochalcone phloretin, and its derivatives, phloridzin, trilobatin, and sieboldin with various R-group substitutions.

DHCs have been studied for physiological roles in *Malus.* DHCs have been implicated in disease resistance in *Malus*, although there is no conclusive evidence directly relating dihydrochalcone concentrations to resistance (Dugé de Bernonville et al. 2011; Mikulič Petkovšek et al. 2008). Fluctuations in DHC and flavonoid biosynthesis in response to infection were reported (Gaucher et al. 2013a; Slatnar et al. 2012). Gosch et al. (2009) proposed that DHCs contribute to disease resistance through an oxidative cascade, and Gaucher et al. (2013a) observed variation in this cascade in susceptible and resistant genotypes. Conversely, DHCs have

been associated with apple replant disease by signaling and promoting pathogen growth (Yin et al. 2017). Phloridzin may act as an auxin regulator in *Malus* (Dare et al. 2017). Independent RNA interference silencing of chalcone synthase and phloretin-glycosyltransferases resulted in reduced auxin transport and aberrant tree architecture in apple (Dare and Hellens 2013; Dare et al. 2017). Whether or not sieboldin and trilobatin serve similar roles in *Malus* species needs to be researched.

While most *Malus* species produce phloridzin as the primary DHC, *Malus toringo* (Siebold) Siebold ex de Vriese and *Malus sargentii* Rehder replace phloridzin with sieboldin, and *Malus trilobata* (Poir.) Schneid. replaces phloridzin with trilobatin (Williams 1961). Interspecific hybridization with these species and other *Malus* results in codominant inheritance of phloridzin, sieboldin, and trilobatin (Hunter 1975; Williams and Jarrett 1975). Williams and Jarrett (1975) first suggested DHCs in apple were controlled by independently segregating genes. Unlike other phenolics with well characterized metabolic networks, little is known about DHC synthesis and regulation. Its unique presence in cultivated apple, has made it somewhat of a model for DHCs. However, genetic studies in *Malus* can be challenging. Most genetic studies in apple involve F_1 progeny from single or multiple families (McClure et al. 2016), though advances in genotyping are allowing for analysis of populations through association studies (Kumar et al. 2013). Association studies in *Malus* allow for inclusion of unrelated individuals and greater resolution of QTL, but are challenging in apple due to low linkage disequilibrium related to high heterozygosity and obligate outcrossing, and a large number of individuals are required for small effect QTL (McClure et al. 2014). The genetic variation of flavonoids in apple have been studied; with QTL for several major apple phenolics reported, including some related

to phloridzin content variation (Chagné et al. 2012; Khan et al. 2012; Verdu et al. 2014). More research is needed to understand the interactions of DHC metabolic pathways.

The aim of this study was to better characterize the genetic controllers of dihydrochalcone content in apple, in particular, those controlling sieboldin and trilobatin. We followed the segregation of major apple DHCs phloridzin, sieboldin, and trilobatin in five apple F_1 populations and confirmed that DHCs segregate independently using χ^2 analysis. Using linkage and association mapping, we identified marker-trait associations on apple linkage groups 7 and 8.

MATERIALS AND METHODS

Plant Material

Five F_1 populations were developed using *Malus prunifolia* (Willd.) Borkh. PI 589816 as the pollen parent. Seed parents included the ornamental cultivar 'Evereste' (population 16705) and selections from the Cornell University apple breeding program in Geneva, NY; NY-152 ('Wijcik McIntosh' × 'Suncrisp') (population 13427), and three 'Evereste' × 'Red Jade' selections (populations 16708, 16709, and 16710) (Table 3.1). Pedigrees are illustrated in Figure 3.2. *M. prunifolia* PI 589816, 'Evereste', and the 'Evereste' × 'Red Jade' progeny were selected based on dihydrochalcone profiles, containing phloridzin, sieboldin, and trilobatin. Population 13427 was initially developed to study the columnar trait of NY-152. Each cross had high fruit set and seed viability (> 95%), although both PI 589816 and NY-152 are pale green lethal carriers and 25% of the initial seedlings died at the cotyledon stage in population 13427 (Orcheski et al. 2015). Population 13427 was from a cross made in 2013, germinated in the summer of 2015, overwintered in a greenhouse, and planted in spring of 2016. Leaf samples of

13427 were collected from the greenhouse (fall 2015) and from the field (fall 2016). Populations 16705, 16708, 16709, and 16710 were developed in 2016, stratified in the winter of 2017, and planted in spring 2017. Samples were collected from the parents for HPLC and DNA extraction. Population 13427 was evaluated for morphological traits including plant height, architecture, and internode length, all related to the columnar phenotype. A major QTL for columnar was identified previously on linkage group 10 (Morimoto and Banno 2015) and validated in population 13427. All populations were planted in orchards maintained by the New York State Agricultural Experiment Station in Geneva, NY. PI 589816 and 'Red Jade' (PI 588786) are maintained on 'EMLA 7' rootstock in the USDA-ARS National Plant Germplasm System (NPGS) *Malus* collection in Geneva, NY. 'Evereste', NY-152, 'Evereste' × 'Red Jade' selections are maintained on 'Geneva 935' rootstock. Leaf samples from 377 accessions were collected from the NPGS *Malus* collection from 2013 to 2015, and included several wild and hybrid species (Gutierrez et al. 2017).

Table 3.1 Pedigree, parental dihydrochalcone classes, and population size for each F_1 population.

Pop.	Parents	DHC Profile	Count
13427	NY-152 ('Wijcik McIntosh' × 'Suncrisp') × PI 589816	P × SPT	445
16705	'Evereste' × PI 589816	SPT × SPT	150
16708	00121-005 ('Evereste' × 'Red Jade') × PI 589816	SPT × SPT	81
16709	00121-063 ('Evereste' × 'Red Jade') × PI 589816	SPT × SPT	44
16710	00121-016 ('Evereste' × 'Red Jade') × PI 589816	SPT × SPT	151

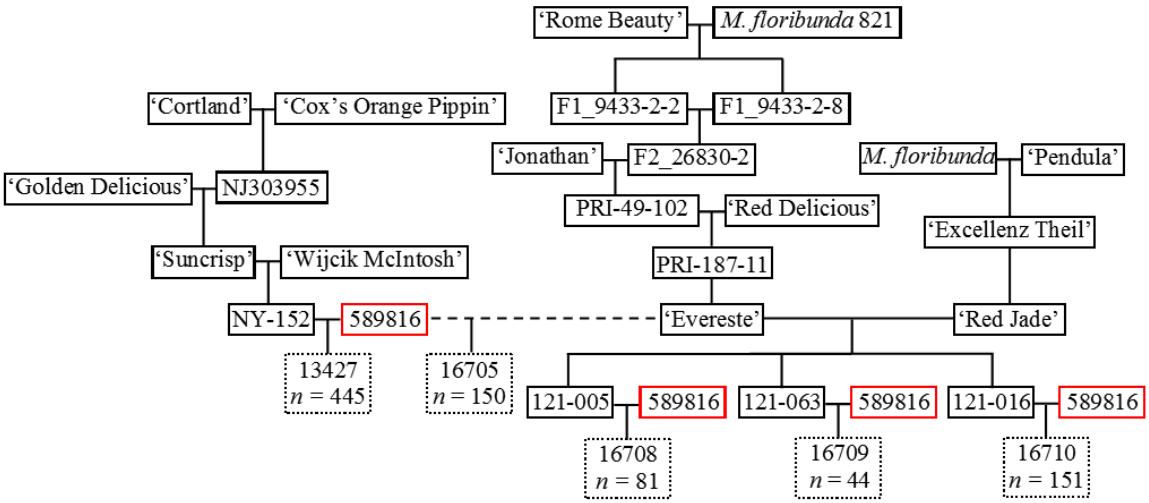


Figure 3.2 Pedigree of five F_1 populations, 13427, 16705, 16708, 16709, and 16710 (dotted lines). Red boxes denote common pollen parent, *M. prunifolia* PI 589816. Seed parent listed on left, except at dashed line.

High Performance Liquid Chromatography

Phenolics were extracted from leaf samples following the methods described by Gutierrez et al. (2017). Briefly, 25 mg of ground leaf tissue was suspended in 1.5 ml of 70% methanol (99.9%) with 2% formic acid (98–100%) for ten minutes then centrifuged, with two technical replicates per genotype. Supernatant was filtered with a 0.20 µm syringe filter, and compounds were separated with a 15% acetonitrile, 10% formic acid mobile phase and an Inertsil ODS-3 column (4.6×250 mm; 5 µm, GL Sciences Inc., Tokyo, Japan) using a Hitachi LaChrom Elite (San Jose, CA, USA) liquid chromatograph equipped with diode array detector. Standard curves for phloridzin, sieboldin, and trilobatin ranged from 7.8 µg to 1,000 µg with R^2 > 0.99, and sample values expressed in mg/g leaf fresh weight.

Genotyping

Individuals of population 13427 were scored using genotyping-by-sequencing (Elshire et al. 2011) following the methods described previously (Gutierrez et al. (2017). Briefly, sequence tags were aligned to the *M. ×domestica* Whole Genome Reference Assembly v.2

(https://www.rosaceae.org/) using Bowtie 2 (Langmead and Salzberg 2012) with parameters D, R, N, L and *i* set to 30, 5, 1, 15 and S, 1, 0.25 to reduce misalignment. The Tassel 5 (Glaubitz et al. 2014) GBS pipeline was used for SNP calling. Ten replicates of parents NY-152 and PI 589816 were merged into single genotypes. VCFtools (Danecek et al. 2011) was used to filter data; twenty-six individuals were removed for low mean depth; sites were filtered for read depth of 8 and minor allele frequency of 0.20; and genotypes were filtered for 80% missing data. Kinship analysis and PCA identified twenty-five potentially outcrossed individuals which were removed from subsequent analyses. Genotypes were filtered based on Mendelian error calculated in PLINK 1.9 (Purcell et al. 2007) and set to missing. SNP markers were filtered based on a χ^2 test for segregation distortion and >0.95 similarity using JoinMap4.1. Genotypes from 280 of *Malus* germplasm accessions included in this study were scored as previously described (Gutierrez et al. 2017).

Genetic Maps and QTL Analysis

GBS markers were divided into male (<nnxnp>) and female (<lmxll>) subsets based on JoinMap 4.1 (Kyazma, NL) conventions. Genetic maps for each parent were constructed with JoinMap 4.1, using regression mapping with linkages with recombination frequencies < 0.35 and LOD > 1.0, a jump threshold of 5.0 and Kosambi's function. Loci were iteratively added to linkage groups based on strongest cross-link value. Interval mapping with regression analysis was done for each map separately using MapQTL 6 (Kyazma, NL) software. Significance threshold for each trait was determined through permutations test (n=10,000). Association mapping was performed using the mixed linear model (MLM): $\mathbf{Y} = \mathbf{X\beta} + \mathbf{Zu} + \mathbf{e}$, where $\mathbf{Y}$ is a vector of DHC ratios; $\mathbf{X}$ and $\mathbf{Z}$ are design matrices; $\boldsymbol{\beta}$ is a fixed effects vector, which includes SNP markers, population structure (Q) determined by the first three principal components, and

the model intercept; **u** is a vector of random additive genetic effects with covariance matrix $\mathbf{2K}\sigma_a^2$, where **K** is VanRaden kinship matrix **(K)** determined by SNP markers; **X** and **Z** are design matrices; and **e** are the residual effects (Wang et al. 2014b). MLM was implemented in GAPIT (Lipka et al. 2012), with 51,348 SNPs filtered for < 80% missing data.

Data Analysis

Principal component analysis and statistical tests, including χ^2 tests for goodness of fit and independence, Pearson correlations, and ANOVA tests were calculated in R 3.3.2 (R Core Team 2016). Welch's ANOVA and Games-Howell post-hoc tests were used where sample size was unequal or equal variance could not be assumed, based on Bartlett's and Levene's tests for homogeneity of variance. Games-Howell tests were calculated in R package 'userfriendlyscience' (Peters, 2017). Mosaic plots and Pearson residuals were calculated using R package 'vcd' (Zeileis et al. 2007).

RESULTS

Dihydrochalcone Content in F_1 Populations

DHCs were measured in all populations and their parents. Mean retention times across all samples for sieboldin, phloridzin, and trilobatin were 15.0, 20.3, and 30.7 minutes, respectively. Mean DHC content in mg/g for parents of each population is presented in Table 3.2. In all populations, we observed DHC profiles differing from parents (Table 3.3). The ratios of progeny DHCs suggest they are controlled by independently segregating genes as determined by a χ^2 test for goodness of fit based on three-gene model for phloridzin, sieboldin, and trilobatin (Williams and Jarrett 1975) and inferred genotypes for each parent ($p > 0.05$) (Figure 3.3; Table 3.3). However, population 16708 deviated significantly from expected ratios (χ^2 (5) = 52.3, $p <$

0.001). Sieboldin was neither observed independent of trilobatin in the progeny nor in the *Malus* germplasm (Gutierrez et al. 2017), though the proportion of sieboldin to trilobatin can be large, and trilobatin detected in trace amounts. It was assumed that detection of sieboldin is dependent on trilobatin, and hypothetical profiles of sieboldin and phloridzin only produce phloridzin. Based on segregation in all five populations, PI 589816 is heterozygous for phloridzin, sieboldin, and trilobatin. Maternal parent NY-152 in population 13427 is homozygous for phloridzin, and recessive for sieboldin, and trilobatin, based on observed profiles (PT, P, and SPT) in 1:2:1 ratios. Population 16705 produced two DHC profiles (SPT and P) in 3:1 ratios, possible if 'Evereste' is homozygous for phloridzin and sieboldin, and heterozygous for trilobatin. Dugé de Bernonville et al. (2011) reported recovery of only P and SPT profiles in 46 F_1 progeny of 'MM106' (P) × 'Evereste' (SPT), which also fit this proposed genotype. 00121-063 and 00121-016 are homozygous for phloridzin and heterozygous for sieboldin and trilobatin, based on the 9:4:3 ratio of SPT, P and PT in populations 16709 and 16710. In population 16708 five DHC profiles were observed; P, T, PT, ST, and SPT. Selection 00121-005 was assumed to be heterozygous for phloridzin, sieboldin, and trilobatin, though profiles observed did not match the expected ratios.

Table 3.2 Mean leaf dihydrochalcone content (mg/g) in parents of apple F_1 populations.

Parent	DHC Profile	Phloridzin	Sieboldin	Trilobatin	Total DHC
PI 589816	SPT	53.8	17.2	7.4	78.3
'Evereste'	SPT	33.6	30.2	5.5	69.2
'Red Jade'	P	90.5	n.d	n.d	90.5
NY-152	P	37.3	n.d	n.d	37.3
00121-005	SPT	21.0	0.8	1.5	23.3
00121-063	SPT	27.1	20.7	20.4	68.3
00121-016	SPT	20.9	5.6	7.3	33.8

Not detected = n.d.

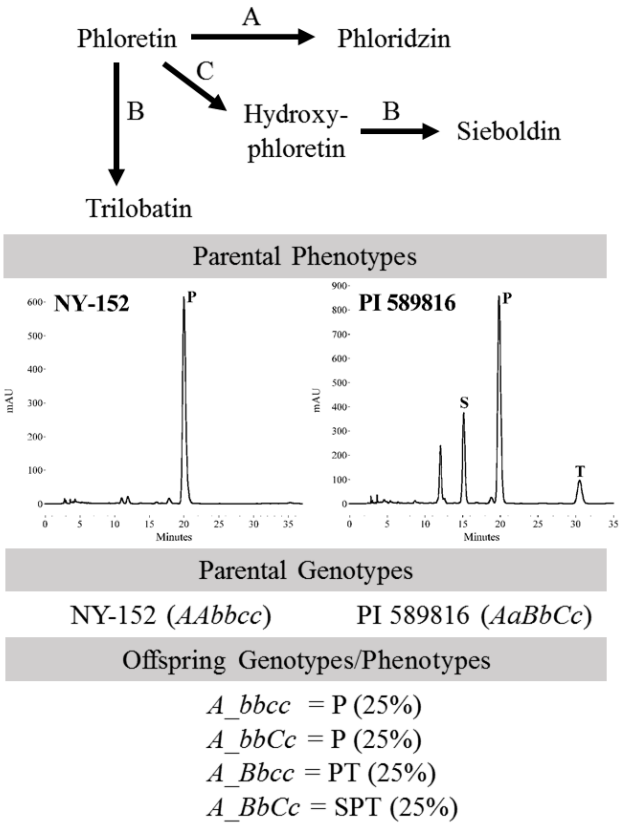


Figure 3.3 Three-gene co-dominant model for phloridzin, sieboldin, and trilobatin segregation in population 13427. Loci *A*, *B,* and *C* result in synthesis of phloridzin, trilobatin, and sieboldin, respectively, where sieboldin synthesis is dependent upon *B* allele (*A_bbC_* produces only phloridzin). Loci assumed to be unlinked based on segregation ratios. NY-152 produces only phloridzin, and PI 589816 produces phloridzin, sieboldin, and trilobatin (SPT).

Table 3.3 Number of offspring in each dihydrochalcone profile by population and accessions from the *Malus* germplasm collection in Geneva, NY with χ^2-test for goodness-of-fit based on inferred parental genotypes.

Pop.	DHC Profile					Hypothetical Genotypes	χ2 value
	P	T	PT	ST	SPT		(df, p-value)
13427	214	0	109	0	97	AAbbcc × AaBbCc	0.8 (2, 0.66)
16705	36	0	0	0	114	AABbCC × AaBbCc	0.1 (1, 0.77)
16708	39	2	3	6	31	AaBbCc × AaBbCc	52.3 (5, <0.0001)
16709	14	0	9	0	21	AABbCc × AaBbCc	1.5 (2, 0.48)
16710	41	0	34	0	76	AABbCc × AaBbCc	2.7 (2, 0.26)
Germplasm[a]	284	1	2	36	54		

[a] Gutierrez et al. 2017

DHC content varied significantly within and among populations, based on Welch's one-way analysis of variance and Games-Howell post hoc tests with Bonferroni correction.

Phloridzin content (11.5 to 144.3 mg/g) was significantly different among populations (F (4,170) = 9.35, $p < 0.001$). Sieboldin content (3.1 to 57.9 mg/g) varied among populations (F (4, 77) = 55.98, $p < 0.0001$), and was significantly lower in population 13427. Trilobatin content (0.1 to 46.5 mg/g) was different among populations (F (4, 103.9) = 138.6, $p < 0.0001$) and was significantly lower in population 16705. Total DHC (15.8 to 144.3 mg/g) and varied significantly among populations (F (4, 6.9) = 34.4, $p < 0.001$). Mean and range of individual and total DHC content are listed by population in Table 3.4 and depicted in Figure 3.4. In population 13427, mean DHC content was higher in 2016. This was likely due to environmental variation, as 2015 samples were from plants in the greenhouse, and 2016 samples were from field planted trees.

Table 3.4 Mean (±SD) and range of leaf dihydrochalcone content (mg/g) in apple F_1 populations**.**

Pop.	Date	Phloridzin		Sieboldin		Trilobatin		Total DHC	
		Mean±sd	Range	Mean±sd	Range	Mean±sd	Range	Mean±sd	Range
13427	2015	50.6±14.1	18.1–106.2	10.7±3.0	3.9–19.7	12.4±6.2	3.2–27.4	58.7±13.8	20.3–106.2
13427	2016	69.2±20.5	25.6–133.3	10.0±3.8	4.1–20.1	15.0±6.9	2.7–33.9	78.3±16.4	33.5–133.3
16705	2017	51.3±23.6	16.3–144.3	20.4±9.7	5.1–46.0	3.2±3.4	0.1–19.4	69.1±19.9	34.0–144.3
16708	2017	58.2±26.3	2.5–130.1	22.9±12.1	5.7–57.9	10.9±10.5	0.3–37.9	68.6±22.4	19.8–130.1
16709	2017	55.2±22.1	24.8–104.5	19.4±8.7	5.5–34.0	15.7±12.2	2.9–46.5	75.1±17.8	39.6–104.5
16710	2017	47.5±19.6	11.5–97.9	14.9±7.9	3.1–39.1	11.5±8.4	1.0–36.7	63.5±16.3	15.8–111.7

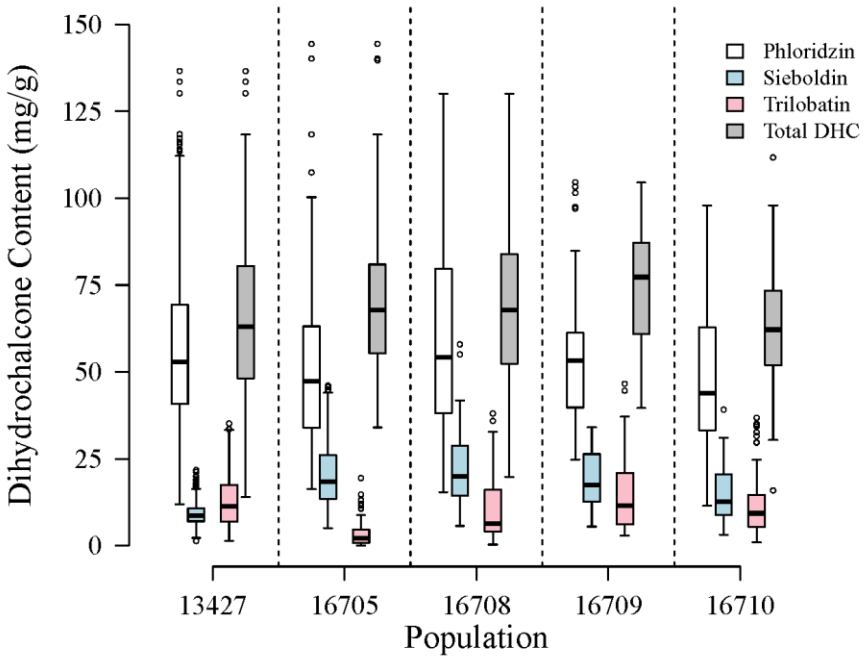


Figure 3.4 Boxplot of phloridzin (white), sieboldin (blue), trilobatin (pink), and total dihydrochalcone (gray) content (mg/g) in leaves, across five *Malus* F_1 populations.

Individual DHC content varied by DHC profiles. Mean phloridzin content was higher in P profiles than in PT or SPT (F (2, 1284.5) = 423.56, $p < 0.0001$) and mean trilobatin content was higher in PT profiles than in SPT (t (1), $p < 0.0001$), consistent across all populations (Figure 3.5). Correlations among individual DHCs varied by population. In 13427 there is a moderate correlation between sieboldin and phloridzin ($r = 0.54$), trilobatin and phloridzin ($r = 0.52$) and trilobatin and sieboldin ($r = 0.50$), where $p < 0.0001$. When subset by DHC profiles PT and SPT in population 13427, the correlations between trilobatin and phloridzin are 0.74 and 0.71, respectively. Plotting trilobatin against phloridzin reveals two distinct groups based on dihydrochalcone profile (Figure 3.6). This pattern was not observed in other populations. There was a moderate correlation between phloridzin and sieboldin ($r = 0.62$, $p < 0.001$) in population 16708 and between sieboldin and trilobatin ($r = 0.52$, $p < 0.0001$) in population 16710. Other

correlations among DHCs in these populations varied from weak to moderate but were not significant ($p > 0.01$).

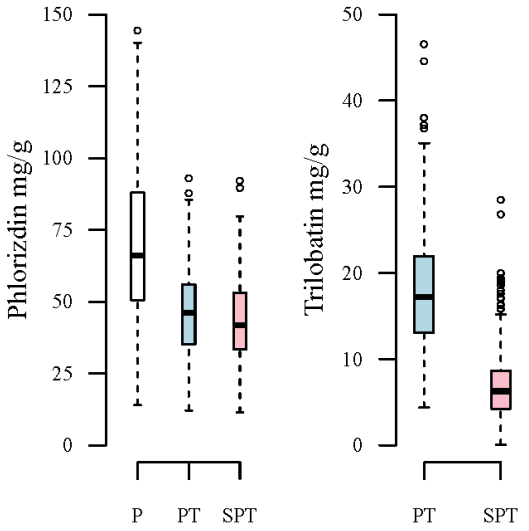


Figure 3.5 Leaf phloridzin and trilobatin content (mg/g) variation by dihydrochalcone profiles P, PT, and SPT across all populations.

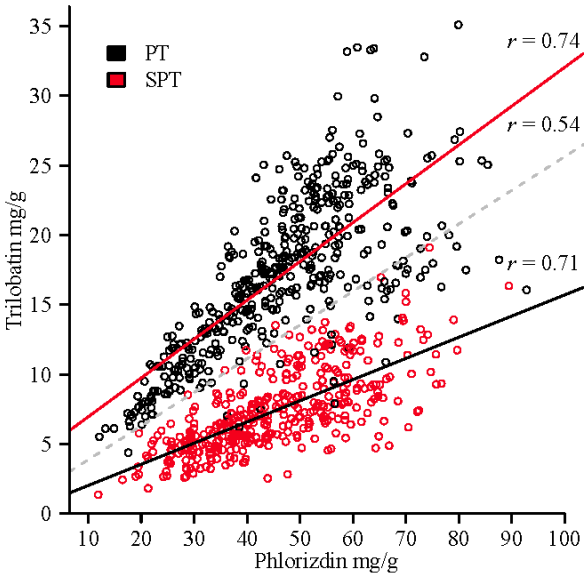


Figure 3.6 Scatterplot of leaf phloridzin (x-axis) and trilobatin (y-axis) content in population 13427. Dihydrochalcone profiles PT (black) and SPT (red) form two distinct groups with different slopes, and correlation of phloridzin and trilobatin increases when determined within each profile separately than combined.

The relative proportions of phloridzin, sieboldin, and trilobatin varied across populations. Phloridzin was the prominent DHC in all populations, with mean proportions > 60%. In population 13427 phloridzin was the prominent DHC in all individuals, ranging from 64 to 87% of the total content in SPT profiles. A few individuals in the other populations had proportions of sieboldin and trilobatin > 50% and as high as 64%. With the exception of rare T profiles, trilobatin was generally a minor component in all five populations. The proportions of phloridzin, sieboldin, and trilobatin in progeny with PT and SPT profiles are presented in Table 3.5.

Table 3.5 Mean (range) of leaf dihydrochalcone proportions in individuals of apple F1 populations with PT and SPT profiles.

Pop.	PT Profiles		SPT Profiles		
	P/Total DHC	T/Total DHC	P/Total DHC	S/Total DHC	T/Total DHC
13427	0.72 (0.63–0.88)	0.27 (0.12–0.37)	0.73 (0.64–0.87)	0.15 (0.06–0.23)	0.12 (0.05 – 0.20)
16705	n.d.	n.d.	0.63 (0.37–0.88)	0.31 (0.11–0.62)	0.05 (<0.01–0.25)
16708	0.60 (0.54-0.63)	0.40 (0.37–0.46)	0.60 (0.48–0.78)	0.31 (0.12–0.57)	0.11 (<0.01–0.43)
16709	0.60 (0.36–0.87)	0.40 (0.13–0.64)	0.61 (0.42–0.75)	0.25 (0.12–0.37)	0.13 (0.04 – 0.30)
16710	0.66 (0.37–0.84)	0.34 (0.15–0.63)	0.63 (0.38–0.88)	0.24 (0.05–0.50)	0.12 (0.03 – 0.29)

Not determined = n.d.

Genetic Mapping

GBS produced 1,115,484 reads, 37.9% of which aligned uniquely to the reference genome, and 15.7% aligned > 1 times. After filtering, ~ 3,100 and 1,500 GBS markers were available for the male and female maps respectively. To validate our genetic maps we confirmed a previously described locus for columnar architecture on LG10 (Morimoto and Banno 2015) using plant height, internode length, and tree architecture descriptors. Interval mapping using the male genetic map identified a QTL on the bottom of LG7 near marker LG7_27346079 at 67.1 cM, which explained 24.9, 22.4, and 50.2% of the variation of phloridzin, sieboldin, and trilobatin content, respectively. Another QTL on the bottom of LG8 near marker LG8_33717934

at 63.7 cM explained 19.8 % of the variation in sieboldin content. Male genetic maps of linkage groups 7 and 8 are highlighted in Figure 3.7. Figure 3.8 illustrates the proportions of DHCs with the two most significant SNP markers associated with trilobatin and sieboldin content. Marker LG7_27743464 follows a 1:2:1 segregation between homozygous G and profile P, and heterozygous A/G (R) and profiles PT and SPT. This marker is likely associated with trilobatin synthesis in this population. Marker LG8_31717934 is likely associated with sieboldin synthesis as nearly all SPT profiles are heterozygous C/T (Y), while most PT and over half P profiles are homozygous C/C. About 40% of individuals of profile P also are heterozygous for marker LG8_31717934, which suggests that theoretical profile SP manifests as P only.

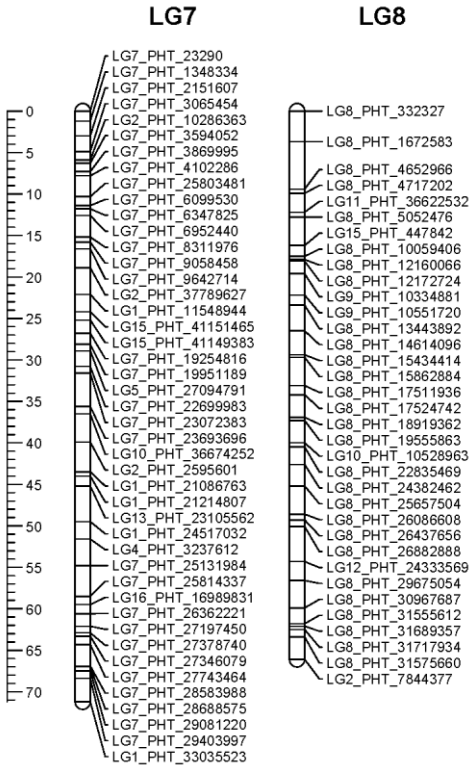


Figure 3.7 Male (*M. prunifolia* PI 589816) genetic maps of linkage groups 7 and 8 from population 13427 developed using genotyping-by-sequencing. SNP ID refers to chromosome (LG) and its physical position determined by sequence tag alignment to the 'Golden Delicious' Whole Genome Reference Assembly v.2.

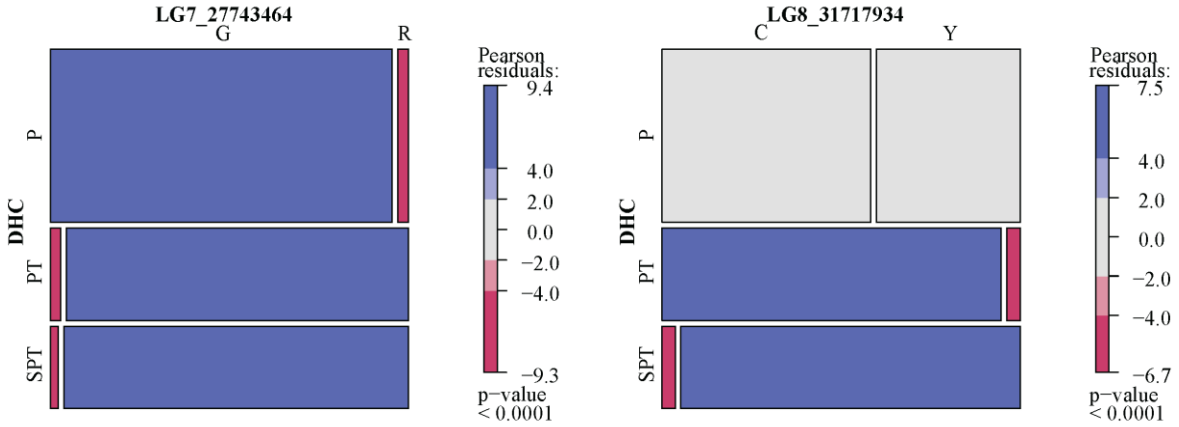


Figure 3.8 Mosaic plot of proportions of dihydrochalcone (DHC) profiles by most significant SNP markers on linkage groups 7 and 8 in population 13427. Box height represents proportion of individuals in DHC profiles and box width represents number of individuals with a particular genotype. Box colors represent deviation from expected observations assuming independence. Blue boxes have more individuals than expected, red boxes have fewer; and gray boxes are not significantly different based on Pearson residuals: $r_{ij} = (O_{ij} - E_{ij}) / \sqrt{E_{ij}}$.

Within the *Malus* germplasm collection, we identified an association (Figure 3.9) around the same position on LG 7 related to dihydrochalcone composition of 279 individuals. Marker LG7_27704634 homozygous G/G was tightly associated with profile P, whereas a significant portion of SPT, PT, and ST types were heterozygous G/T or homozygous T/T.

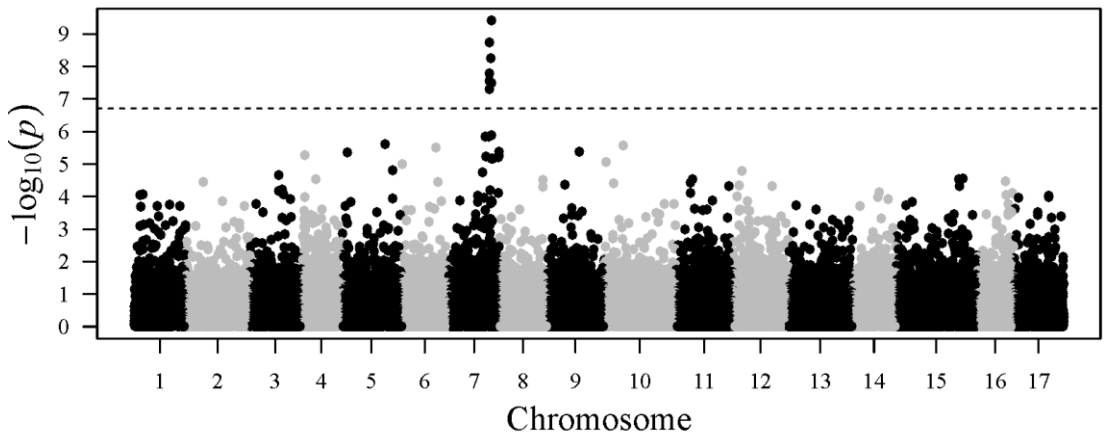


Figure 3.9 Manhattan plot based on mixed model of leaf dihydrochalcone composition in 279 *Malus* germplasm accessions controlling for population structure and relatedness with QTL on chromosome 7. Negative $\log_{10}$ *p*-values are plotted against physical position on 17 apple chromosomes. Horizontal line denotes Bonferroni adjusted significance threshold at $\alpha = 0.01$.

DISCUSSION

DHCs are structurally similar to narigenin chalcone (a precursor to flavonols, anthocyanins, and flavan-3-ols) but, deviate early through a Δ2–3 carbon double bond reduction of *p*-coumaroyl-CoA to *p*-dihydrocoumaroyl-CoA (Dare et al. 2013a; Ibdah et al. 2014). Chalcone synthase converts *p*-dihydrocoumaroyl-CoA and three molecules of malonyl-CoA to phloretin which is then glycosylated. A molecule of glucose is added to 2′-position of phloretin via phloretin 2′-O-glucosyltransferases and other glycosyltransferases forming phloridzin (Gosch et al. 2010a; Jugdé et al. 2008). Although, multiple enzymes synthesize phloridzin from phloretin *in vitro*, population analysis suggests individual DHCs are controlled by single loci. Because our genotyped population 13427 did not segregate for phloridzin, we did not identify a major determinant locus. *MdUGT88F1* and *MdUGT88F4*, on apple chromosome 15 substantially reduces phloridzin content when silenced (Dare et al. 2017; Zhou et al. 2017) and is a candidate for further investigation.

Trilobatin differs from phloridzin by the position of the glucose moiety, which could occur through a phloretin 4′-O-glucosyltransferase (Yahyaa et al. 2016). Our results suggest there is a major determinant for trilobatin synthesis on apple linkage group 7 near position 27704634 on version 2 of the apple genome. However, we did not identify potential gene candidates near this loci. A functional enzyme capable of synthesizing trilobatin was described by Yahyaa et al. (2016), but was located on apple chromosome 9. There is evidence that trilobatin synthesis enzymes are present in cultivated apples, though trilobatin is not typically detected. Yahyaa et al. (2016) identified *UGT75L17* from the ‘Golden Delicious’ genome, which is capable of synthesizing trilobatin from phloretin *in vitro*, and was expressed in ‘Golden

Delicious' tissues. Additionally, Dare et al. (2017) detected trilobatin synthesis in 'Royal Gala' *MdUGT88F1* silenced lines.

Sieboldin differs from trilobatin through C-3 hydroxylation. The biosynthesis of sieboldin has not been elucidated, but it may form through C-3 hydroxylation of phloretin, followed by glycosylation via phloretin 4′-O-glucosyltransferase (Hutabarat et al. 2016). Its dependence upon the same glycosyltransferase as trilobatin would explain why sieboldin is not detected independently. In population 13427 we observed about 40% of the P type individuals with the SNP LG8_32451169, associated with sieboldin.

In addition to phloridzin, sieboldin, and trilobatin, there are additional DHCs in *Malus*. These minor DHCs, including phloretin, do not occur in appreciable quantities, like, raising questions concerning the preferential formation of phloridzin, sieboldin, and trilobatin. However, in these populations we detected an unidentified peak around 13.1 minutes, with maximum absorbance at 280 nm. This peak was found in similar concentrations as sieboldin or trilobatin in these populations, and was strongly associated with sieboldin. Interval mapping identified a QTL for this unknown peak at the same position on LG 8 as sieboldin. We suggest that this unknown is a hydroxylated phloretin derivative, such as hydroxyphloridzin (3-hydroxyphloretin 2′-O-glucoside), but further investigation is needed.

Mean content for individual DHCs varied by DHC profile. Mean phloridzin content was significantly higher in profile P than in profiles PT and SPT, and mean trilobatin content was higher in profile PT than in SPT across all populations. Total DHC content did not vary between DHC profiles. These results suggest a partitioning effect between phloridzin, sieboldin, and trilobatin, though there was no difference between mean phloridzin content between PT and SPT profiles. However, we observed some significant positive linear relationships between the DHCs

in some, but not all of the F_1 populations, with correlations ranging from 0.00 to 0.74. From our observations, the correlations are stronger between trilobatin and phloridzin content when divided into PT and SPT profiles. Additionally, we observed significant variation in the proportion of DHCs relative to their parental types. While, phloridzin was the major DHC on average across populations, sieboldin or trilobatin were dominant in several individuals. Partitioning of these DHCs in apple merits further biochemical analysis.

Dihydrochalcone profiles observed in the progeny followed a pattern of three, co-dominantly expressed, segregating loci. These genes do not appear to be linked as ratios fit a χ^2-test for goodness of fit with large *p*-values in all populations but 16708. It is unclear why this population deviated significantly from the other populations. The maternal parent, 00121-005, was unique among the other 'Evereste' × 'Red Jade' progeny. Although, its dihydrochalcone profile was SPT, it produced reduced levels of sieboldin and trilobatin (Table 3.2), and low levels of total leaf dihydrochalcone content relative to its 16 full siblings, ranging from 23.3 to 129 mg/g. Additionally, it has unique morphological characteristics, including linear to lanceolate leaf shape, generally < 7 mm wide and 28 mm long leaves, shortened internodes (< 12.5 mm), and small fruit, similar to those described by Dare et al. (2017) in reduced phloridzin transgenic lines. Figure 3.10 highlights leaf variation between population founders. There are potentially other genetic factors affecting dihydrochalcone concentrations in this population, which merit further investigation.

Figure 3.10 Leaf variation between select founders of F_1 populations on 2.54 cm^2 grids: from top left to bottom right, 'Evereste', 'Red Jade', PI 589816, and 'Evereste' × 'Red Jade' parents of 16708, 16709, and 16710.

The evolution of dihydrochalcone production in *Malus* is unique within the Rosaceae. Although no phloridzin is detected in pear (*Pyrus communis*), a close relative, phloretin and phloridzin are synthesized from p-dihydrocoumaroyl-CoA with malonyl-CoA with enzymatic extracts from pear leaves, suggesting biochemical divergence in *Malus* from pear at an earlier step, which is likely double bond reduction of *p*-coumaroyl CoA (Gosch et al. 2009, 2010a). Apple-pear intergeneric hybrids are able to produce both phloridzin and arbutin, a pear-specific metabolite absent in apple (Fischer et al. 2014). Although phloridzin was reported in other rosaceous plants (rose and strawberry), it is not found in high concentrations (Gosch et al. 2010b). Phloretin is the precursor to each compound, though sieboldin may be formed following an additional hydroxylation step. Phloretin was not detected in this study, and is generally a minor component of dihydrochalcone content or not reported in *Malus* phenolic studies involving various tissue types (De Paepe et al. 2015; Tang et al. 2015; Verdu et al. 2014);

presumably, phloretin is nearly completely glycosylated *in planta*. Although, the total dihydrochalcone content may represent an estimate of phloretin content, no QTL were identified in association with the observed variation. Phloretin is more reactive than its glycosylated forms, resulting in cytotoxicity (Gosch et al. 2009). Silencing of phloretin 2′-*O*-glucosyltransferase UGT88F1 in 'Royal Gala' resulted in both reduced phloridzin and phloretin production (Dare et al. 2017). There may be some feedback mechanism which inhibits the accumulation of phloretin in healthy plants.

There were no obvious genes related to dihydrochalcone synthesis in relation to the loci identified on LGs 7 and 8. Biochemical analysis has focused on a few key points in the dihydrochalcone pathway: double bond reduction of *p*-coumaroyl CoA to dihydro-*p*-coumaroyl CoA; formation of phloretin from dihydro-*p*-coumaroyl and malonyl-CoA via chalcone synthase; and glycosylation of phloretin to form phloridzin and other phloretin glycosides. A list of key genes investigated for dihydrochalcone synthesis both *in vitro* and *in vivo* is presented in Table 3.6. Dihydrochalcone content can be regulated via other enzymes outside its biosynthetic pathway (Khan et al. 2012; Rihani et al. 2017).

There is evidence that phloridzin concentrations may affect other upstream pathways (Dare et al. 2017). Other QTL for DHCs have been reported in apple populations. Verdu et al. (2014) identified QTL on LGs 1 and 5 for phloridzin content in fruit and juice which explained 6.4 to 11.6% of the variation, and on LGs 3, 5, 12, and 15 for phloretin xyloglucoside content in fruit and juice which explained 17.4 to 23.6% of the variation. Chagné et al. (2012) identified a QTL for phloridzin xyloglucoside on LG 17 explaining 10.3% of the variation. No candidate genes were proposed for these QTL. Other traits mapped to LGs 7 and 8, include aphid (*Dysaphis devecta* Walker and *Eriosoma lanigerum* Hausm.) resistance has been associated with

loci on LGs 7 and 8 (Bus et al. 2008; Cevik and King 2002). Woolly apple aphid resistance allele *Er3* from 'Aotea 1' is located on LG 8 (Bus et al. 2008). 'Aotea 1' originates from *M. toringo*, a source of sieboldin. Additionally, marker-trait associations for powdery mildew (from *Malus* hybrid 'White Angel') and fire blight resistance (from 'Fiesta'), fruit weight, and fruit acidity were identified on LGs 7 and 8 (Devoghalaere et al. 2012; Evans and James 2003; Khan et al. 2007; Zhang et al. 2012).

Table 3.6 List of genes related to dihydrochalcone biosynthesis.

Gene	Function	References
EU872155	Chalcone synthase	Gosch et al. 2009; Yahyaa et al. 2017
EU872156	Chalcone synthase	Gosch et al. 2009; Yahyaa et al. 2017
EU872158	Chalcone synthase	Gosch et al. 2009; Yahyaa et al. 2017
FJ216429	Chalcone 3-hydroxylase	Hutabarat et al. 2016
FJ216426	Flavonoid 3′-hydroxylase	Hutabarat et al. 2016
UGT88F1	Phloretin 2′-*O*-glucosyltransferase	Dare et al. 2017; Gosch et al. 2010a
UGT88F2	Phloretin 2′-*O*-glucosyltransferase	Gosch et al. 2010a
UGT88F4	Phloretin 2′-*O*-glucosyltransferase	Gosch et al. 2010a; Zhou et al. 2017
UGT71A15	Phloretin 2′-*O*-glucosyltransferase	Gosch et al. 2012; Jugdé et al. 2008
UGT75L17	Phloretin 4′-*O*-glucosyltransferase	Yahyaa et al. 2016

There were more markers (<nnxnp>) for the male parent than for the female in population 13427. This bias for PI 589816 may stem from its hybrid origin. Though classified as *M. prunifolia*, PI 589816 does not fit the chemical or morphological description of other accessions of *M. prunifolia* (Gutierrez et al. 2017, Chapter 2). For pseudo-testcross mapping two maps are estimated for both parents with markers homozygous in one parent and heterozygous in the other. There appears to be increased heterozygosity in PI 589816, making it difficult to find heterozygous markers in the female and homozygous markers in the male. Groupings based on the LOD independence score resulted in 17 linkage groups for each parent, representing the 17 apple chromosomes. Though it was possible to relate the linkage groups to their corresponding chromosome, several GBS markers from other genomic locations were tightly linked. For

example, on our LG7 we observed several groups of markers corresponding to chromosome 1. This is likely due to genomic misalignment during the GBS pipeline, related to the genomic structure of apple (Velasco et al. 2010); chromosomes 1 and 7 are homeologs from a whole genome duplication. This was also reported by others using GBS in apple (McClure et al. 2016; Norelli et al. 2017). There was a strong correlation between a marker's physical position and map distance.

CONCLUSIONS

Dihydrochalcones phloridzin, sieboldin, and trilobatin are unique compounds in *Malus* species. We observed significant DHC content variation among five F1 populations and segregation of DHCs in apple, following a model of three, independently segregating loci. Additionally, we identified the first genomic loci associated with trilobatin and sieboldin segregation in population 13427 and *Malus* germplasm. Markers on LG7 and LG8 were strongly associated with trilobatin and sieboldin production, respectively. We posited that sieboldin is biochemically linked to trilobatin, based on the loci on LG7. Most cultivated apples do not have appreciable quantities of these DHCs and most only produce phloridzin. Loci identified in this study require further investigation but may facilitate the utilization of genetic resources to broaden the nutritional quality of new apple cultivars.

REFERENCES

Behzad S, Sureda A, Barreca D, Nabavi SF, Rastrelli L, Nabavi SM (2017) Health effects of phloretin: From chemistry to medicine. Phytochem Rev 1–7. doi:10.1007/s11101-017-9500-x

Bus VGM, Chagné D, Bassett HCM, Bowatte D, Calenge F, Celton J-M, et al (2008) Genome mapping of three major resistance genes to woolly apple aphid (*Eriosoma lanigerum* Hausm.). Tree Genet Genomes 4:223–236. doi:10.1007/s11295-007-0103-3

Cevik V, King GJ (2002) Resolving the aphid resistance locus *Sd-1* on a BAC contig within a sub-telomeric region of *Malus* linkage group 7. Genome 45:939–945

Chagné D, Krieger C, Rassam M, Sullivan M, Fraser J, André C, et al (2012) QTL and candidate gene mapping for polyphenolic composition in apple fruit. BMC Plant Biol 12:12. doi:10.1186/1471-2229-12-12

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al (2011) The variant call format and VCFtools. Bioinformatics 27:2156–2158. doi:10.1093/bioinformatics/btr330

Dare A, Hellens R (2013) RNA interference silencing of *CHS* greatly alters the growth pattern of apple (*Malus* x *domestica*). Plant Signal Behav 8:e25033. doi:10.4161/psb.25033

Dare AP, Tomes S, Cooney JM, Greenwood DR, Hellens RP (2013) The role of enoyl reductase genes in phloridzin biosynthesis in apple. Plant Physiol Biochem 72:54–61. doi:10.1016/j.plaphy.2013.02.017

Dare AP, Yauk Y-K, Tomes S, McGhie TK, Rebstock RS, Cooney JM, Atkinson RG (2017) Silencing a phloretin-specific glycosyltransferase perturbs both general phenylpropanoid biosynthesis and plant development. Plant J Cell Mol Biol 91:237–250. doi:10.1111/tpj.13559

De Paepe D, Valkenborg D, Noten B, Servaes K, Diels L, Loose MD, Van Droogenbroeck B, Voorspoels S (2015) Variability of the phenolic profiles in the fruits from old, recent and new apple cultivars cultivated in Belgium. Metabolomics 11:739–752. doi:10.1007/s11306-014-0730-2

Devoghalaere F, Doucen T, Guitton B, Keeling J, Payne W, Ling TJ, et al (2012) A genomics approach to understanding the role of auxin in apple (*Malus* x *domestica*) fruit size control. BMC Plant Biol 12:7. doi:10.1186/1471-2229-12-7

Dong H-Q, Li M, Zhu F, Liu F-L, Huang J-B (2012) Inhibitory potential of trilobatin from *Lithocarpus polystachyus* Rehd against α-glucosidase and α-amylase linked to type 2 diabetes. Food Chem 130:261–266. doi:10.1016/j.foodchem.2011.07.030

Dugé de Bernonville T, Gaucher M, Guyot S, Durel C-E, Dat JF, Brisset M-N (2011) The constitutive phenolic composition of two *Malus* × *domestica* genotypes is not responsible

for their contrasted susceptibilities to fire blight. Environ Exp Bot 74:65–73. doi:10.1016/j.envexpbot.2011.04.019

Dugé de Bernonville T, Guyot S, Paulin J-P, Gaucher M, Loufrani L, Henrion D, et al (2010) Dihydrochalcones: Implication in resistance to oxidative stress and bioactivities against advanced glycation end-products and vasoconstriction. Phytochem 71:443–452. doi:10.1016/j.phytochem.2009.11.004

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple Genotyping-by-Sequencing (GBS) approach for high diversity species. PLoS ONE 6:e19379. doi:10.1371/journal.pone.0019379

Evans K, James C (2003) Identification of SCAR markers linked to Pl-w mildew resistance in apple. Theor Appl Genet 106:1178–1183. doi:10.1007/s00122-002-1147-2

Fan X, Zhang Y, Dong H, Wang B, Ji H, Liu X (2015) Trilobatin attenuates the LPS-mediated inflammatory response by suppressing the NF-κB signaling pathway. Food Chem 166:609–615. doi:10.1016/j.foodchem.2014.06.022

Fischer TC, Malnoy M, Hofmann T, Schwab W, Palmieri L, Wehrens R, et al (2014) F1 hybrid of cultivated apple (*Malus* ×*domestica*) and European pear (*Pyrus communis*) with fertile F_2 offspring. Mol Breed 34:817–828. doi:10.1007/s11032-014-0077-4

Gaucher M, Dugé de Bernonville T, Guyot S, Dat JF, Brisset M-N (2013) Same ammo, different weapons: Enzymatic extracts from two apple genotypes with contrasted susceptibilities to fire blight (*Erwinia amylovora*) differentially convert phloridzin and phloretin *in vitro*. Plant Physiol Biochem 72:178–189. doi:10.1016/j.plaphy.2013.03.012

Glaubitz JC, Casstevens TM, Lu F, Harriman J, Elshire RJ, Sun Q, Buckler ES (2014) TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. PLoS ONE 9:e90346. doi:10.1371/journal.pone.0090346

Gosch C, Flachowsky H, Halbwirth H, Thill J, Mjka-Wittmann R, Treutter D, Richter K, Hanke M-V, Stich K (2012) Substrate specificity and contribution of the glycosyltransferase UGT71A15 to phloridzin biosynthesis. Trees 26:259–271. doi:10.1007/s00468-011-0669-0

Gosch C, Halbwirth H, Kuhn J, Miosic S, Stich K (2009) Biosynthesis of phloridzin in apple (*Malus domestica* Borkh.). Plant Sci 176:223–231. doi:10.1016/j.plantsci.2008.10.011

Gosch C, Halbwirth H, Schneider B, Hölscher D, Stich K (2010a) Cloning and heterologous expression of glycosyltransferases from *Malus* x *domestica* and *Pyrus communis*, which convert phloretin to phloretin 2′-O-glucoside (phloridzin). Plant Sci 178:299–306. doi:10.1016/j.plantsci.2009.12.009

Gosch C, Halbwirth H, Stich K (2010b) Phloridzin: biosynthesis, distribution and physiological relevance in plants. Phytochem 71:838–843. doi:10.1016/j.phytochem.2010.03.003

Gutierrez BL, Zhong G-Y, Brown SK (2017) Genetic diversity of dihydrochalcone content in *Malus* germplasm. Manuscript submitted for publication to Genetic Resources and Crop Evolution on 8/15/17. Manuscript submitted for publication

Hunter LD (1975) Phloridzin and apple scab. Phytochem 14:1519–1522. doi:10.1016/0031-9422(75)85343-X

Hutabarat OS, Flachowsky H, Regos I, Miosic S, Kaufmann C, Faramarzi S, et al (2016) Transgenic apple plants overexpressing the *chalcone 3-hydroxylase* gene of *Cosmos sulphureus* show increased levels of 3-hydroxyphloridzin and reduced susceptibility to apple scab and fire blight. Planta 243:1213–1224. doi:10.1007/s00425-016-2475-9

Ibdah M, Berim A, Martens S, Valderrama ALH, Palmieri L, Lewinsohn E, Gang DR (2014) Identification and cloning of an NADPH-dependent hydroxycinnamoyl-CoA double bond reductase involved in dihydrochalcone formation in *Malus* × *domestica* Borkh. Phytochem 107:24–31

Jugdé H, Nguy D, Moller I, Cooney JM, Atkinson RG (2008) Isolation and characterization of a novel glycosyltransferase that converts phloretin to phlorizin, a potent antioxidant in apple. FEBS J 275:3804–3814. doi:10.1111/j.1742-4658.2008.06526.x

Khan MA, Durel C-E, Duffy B, Drouet D, Kellerhals M, Gessler C, Patocchi A (2007) Development of molecular markers linked to the 'Fiesta' linkage group 7 major QTL for fire blight resistance and their application for marker-assisted selection. Genome 50:568–577. doi:10.1139/g07-033

Khan SA, Chibon P-Y, de Vos RCH, Schipper BA, Walraven E, Beekwilder J, et al (2012) Genetic analysis of metabolites in apple fruits indicates an mQTL hotspot for phenolic compounds on linkage group 16. J Exp Bot 63:2895–2908. doi:10.1093/jxb/err464

Kumar S, Garrick DJ, Bink MC, Whitworth C, Chagne D, Volz RK (2013) Novel genomic approaches unravel genetic architecture of complex traits in apple. BMC Genomics 14:393-2164-14–393

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. doi:10.1038/nmeth.1923

Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhang Z (2012) GAPIT: genome association and prediction integrated tool. Bioinformatics 28:2397–2399. doi:10.1093/bioinformatics/bts444

McClure KA, Gardner KM, Toivonen PM, Hampson CR, Song J, Forney CF, DeLong J, Rajcan I, Myles S (2016) QTL analysis of soft scald in two apple populations. Hortic Res 3:16043. doi:10.1038/hortres.2016.43

McClure KA, Sawler J, Gardner KM, Money D, Myles S (2014) Genomics: A potential panacea for the perennial problem. Am J Bot 101:1780–1790

Mikulič Petkovšek M, Stampar F, Veberic R (2008) Increased phenolic content in apple leaves infected with the apple scab pathogen. J Plant Pathol 90:49–55

Money D, Gardner K, Migicovsky Z, Schwaninger H, Zhong G-Y, Myles S (2015) LinkImpute: fast and accurate genotype imputation for non-model organisms. G3 Genes Genomes Genet 5:2383–2390. doi:10.1534/g3.115.021667

Morimoto T, Banno K (2015) Genetic and physical mapping of *Co*, a gene controlling the columnar trait of apple. Tree Genet Genomes 11:1–11

Nair S, Ziaullah Z, Rupasinghe HV (2014) Phloridzin fatty acid esters induce apoptosis and alters gene expression in human liver cancer cells (261.2). FASEB J 28:261.2

Najafian M, Jahromi MZ, Nowroznejhad MJ, Khajeaian P, Kargar MM, Sadeghi M, Arasteh A (2012) Phloridzin reduces blood glucose levels and improves lipids metabolism in streptozotocin-induced diabetic rats. Mol Biol Rep 39:5299–5306

Norelli JL, Wisniewski M, Fazio G, Burchard E, Gutierrez B, Levin E, Droby S (2017) Genotyping-by-sequencing markers facilitate the identification of quantitative trait loci controlling resistance to *Penicillium expansum* in *Malus sieversii*. PLoS ONE 12:e0172949. doi:10.1371/journal.pone.0172949

Orcheski B, Parker R, Brown S (2015) Pale green lethal disorder in apple (*Malus*) is caused by a mutation in the *PHYLLO* gene which is essential for phylloquinone (vitamin K1) biosynthesis. Tree Genet Genomes 11:131. doi:10.1007/s11295-015-0956-9

Puel C, Quintin A, Mathey J, Obled C, Davicco MJ, Lebecque P, Kati-Coulibaly S, Horcajada MN, Coxam V (2005) Prevention of bone loss by phloridzin, an apple polyphenol, in ovariectomized rats under inflammation conditions. Calcif Tissue Int 77:311–318. doi:10.1007/s00223-005-0060-5

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al (2007) PLINK: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559–575. doi:10.1086/519795

R Core Team (2016) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rihani KAL, Jacobsen H-J, Hofmann T, Schwab W, Hassan F (2017) Metabolic engineering of apple by overexpression of the *MdMyb10* gene. J Genet Eng Biotechnol. doi:10.1016/j.jgeb.2017.01.001

Rivière C (2016) Dihydrochalcones: Occurrence in the plant kingdom, chemistry and biological activities. In: Atta-ur-Rahman (ed) Studies in Natural Products Chemistry, 1st edn. Elsevier, pp 253–381

Slatnar A, Mikulič Petkovšek M, Halbwirth H, Stampar F, Stich K, Veberic R (2012) Polyphenol metabolism of developing apple skin of a scab resistant and a susceptible apple cultivar. Trees 26:109–119

Tang J, Tang L, Tan S, Zhou Z (2015) The study of variation of phloridzin content in six wild *Malus* species. J Food Nutr Res 3:146–151

Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, et al (2010) The genome of the domesticated apple (*Malus* × *domestica* Borkh.). Nat Genet 42:833–839. doi:10.1038/ng.654

Verdu CF, Guyot S, Childebrand N, Bahut M, Celton J-M, Gaillard S, et al (2014) QTL analysis and candidate gene mapping for the polyphenol content in cider apple. PLOS ONE 9:e107103. doi:10.1371/journal.pone.0107103

Wang Q, Tian F, Pan Y, Buckler ES, Zhang Z (2014) A SUPER Powerful Method for Genome Wide Association Study. PLOS ONE 9:e107684. doi:10.1371/journal.pone.0107684

Williams A, Jarrett J (1975) Hybridization of *Malus*. In: Report - Long Ashton Research Station 1974. University of Bristol, Bristol, Eng, p 44

Williams AH (1961) Dihydrochalcones of *Malus* species. J Chem Soc 155:4133–4136. doi:10.1039/JR9610004133

Xiao Z, Zhang Y, Chen X, Wang Y, Chen W, Xu Q, Li P, Ma F (2017) Extraction, identification, and antioxidant and anticancer tests of seven dihydrochalcones from *Malus* 'Red Splendor' fruit. Food Chem 231:324–331. doi:10.1016/j.foodchem.2017.03.111

Yahyaa M, Ali S, Davidovich-Rikanati R, Ibdah M, Shachtier A, Eyal Y, Lewinsohn E, Ibdah M (2017) Characterization of three chalcone synthase-like genes from apple (*Malus* x *domestica* Borkh.). Phytochem 140:125–133. doi:10.1016/j.phytochem.2017.04.022

Yahyaa M, Davidovich-Rikanati R, Eyal Y, Sheachter A, Marzouk S, Lewinsohn E, Ibdah M (2016) Identification and characterization of UDP-glucose:Phloretin 4′-O-glycosyltransferase from *Malus* x *domestica* Borkh. Phytochem 130:47–55. doi:10.1016/j.phytochem.2016.06.004

Yin C, Xiang L, Wang G, Wang Y, Shen X, Chen X, Mao Z (2017) Phloridzin promotes the growth of *Fusarium moniliforme* (*Fusarium verticillioides*). Sci Hortic 214:187–194. doi:10.1016/j.scienta.2016.11.035

Zeileis A, Meyer D, Hornik K (2007) Residual-based Shadings for Visualizing (Conditional) Independence. J Comput Graph Stat 16:507–525

Zhang Q, Ma B, Li H, Chang Y, Han Y, Li J, et al (2012) Identification, characterization, and utilization of genome-wide simple sequence repeats to identify a QTL for acidity in apple. BMC Genomics 13:537. doi:10.1186/1471-2164-13-537

Zhou K, Hu L, Li P, et al (2017) Genome-wide identification of glycosyltransferases converting phloretin to phloridzin in *Malus* species. Plant Science 265:131–145 . doi: 10.1016/j.plantsci.2017.10.003

CHAPTER 4

INCREASED PHLORIZIN CONTENT ASSOCIATED WITH RUSSETING IN APPLE FRUIT

ABSTRACT

Phloridzin is a phenolic compound, unique to apple (*Malus ×domestica* Borkh.) and its wild relatives. Since its discovery, phloridzin has been researched for its nutraceutical properties, including anti-diabetic, anti-cancer, and antioxidant activities, making phloridzin a potential target for nutritional improvement in new apple cultivars. However, phloridzin accumulates at significantly lower concentrations in fruit than in vegetative tissues and seeds. In 'Golden Delicious' and its sports, we observed higher phloridzin content in peels of sports with a cuticle disorder termed russet. In russeted apples, the smooth waxy fruit cuticle is partially or entirely replaced by a corky layer, induced through environmental and genetic effects. To understand the variation of phloridzin content and fruit russet in apple fruit, we surveyed 108 accessions with variation in russeting from the USDA-ARS *Malus* germplasm collection in Geneva, NY. Russeting in apple fruit ranged from 0 to 100%, and phloridzin content ranged from 24.3 to 825.0 μg/g in peels. Mean phloridzin content varied significantly between russeting groups; in groups with light (0-5%), medium-high (70-80%), and high (90-100%) russeting mean phloridzin content was 115.2, 591.2, and 378.8 μg/g. We observed that genetic factors and russeting are strong predictors of phloridzin content in peels, but not fruit flesh or leaves. Other peel phenolics are negatively associated with fruit russeting in apple. Following changes in phloridzin content during fruit development in 'Golden Delicious' (low to medium russet) and its

sports, ‘Empress Spur’ (low russet), ‘Razor’ (complete russet), and ‘Sergeant Russet’ (medium to high russet), we observed variable phloridzin content related to russet incidence.

INTRODUCTION

Consumption of apples and apple products is inversely correlated to several degenerative diseases, including cancer, cardiovascular disease, and diabetes (Hyson 2011). Antioxidants from fresh fruits and vegetables mediate effects of harmful dietary advanced glycation end products associated with modern diets and chronic disease (Liu 2013; Luévano-Contreras et al. 2017). Cultivated apple, *M.* ×*domestica*, and its wild relatives are diverse in nutraceutical content, including carotenoids, phenolics, triterpenes, and vitamin C (Ampomah-Dwamena et al. 2012; De Paepe et al. 2015; Fang et al. 2017). Some studies have suggested that phenolic content is higher and more diverse in heirloom apple cultivars, making phenolics a potential target for improvement in apple breeding. Apples contribute a significant portion of dietary phenolics in the United States (Liu 2013).

The dihydrochalcone phloridzin belongs to the flavonoid class of phenolics, and is unique to apple. Phloridzin was studied for medicinal applications, in particular, for its effects on glucose absorption in the intestines and kidneys (Najafian et al. 2012). Phloridzin is found in nearly all *Malus* species and is absent in closely related genera such as *Prunus* and *Pyrus*. The synthesis of phloridzin was elucidated (Dare et al. 2013), but there is more to discover about its metabolic pathways (Dare et al. 2017; Yahyaa et al. 2017). Initial research indicates phloridzin may regulate auxin transport in apple (Dare et al. 2017), and phloridzin may contribute to disease resistance (Gaucher et al. 2013; Gosch et al. 2009). Phloridzin is disproportionately distributed in the apple tissues; it is more abundant in vegetative tissues (66–90% of total phenolic content),

and lower in mature fruit (2–6% of total phenolic content), where it concentrates primarily in the peel (Gosch et al. 2010). Phenolics and other apple nutraceuticals are also more abundant in the peel.

The apple peel, or cuticle, serves as a protective barrier against biotic and abiotic factors, prevents water loss, and is a key component of fruit aesthetics. Russeting is a response to damage to the cuticle, and restores some of the lost functionality of the cuticle, though with increased water loss and reduced elasticity (Khanal et al. 2014; Lashbrooke et al. 2015). Russeting occurs in a variety of forms in apple with complete and semi-russet cultivars. While many cultivars have some russeting around the pedicle and calyx, russeting on the fruit body is unfavorable in US markets. Due to consumer preference for waxy types, the market value of russeted apples is reduced. However, many russet and semi-russet apples are popular heirloom cultivars, valued for their flavor, fruit quality, and unusual appearance. Russet is thought to occur during fruit expansion, where rapid sigmoidal cell expansion in the fruit during a short period, strains the cuticle and leads to microcracking (Ginzberg and Stern 2016). The damaged cuticle is shed, and the peridermal layer suberizes, sealing the fruit (Lashbrooke et al. 2015). Cultivars with reduced incidence of russeting tend to have smaller fruit epidermal cells, of uniform size and arrangement, and uniform cuticle thickness. Conversely, russet cultivars have irregularly sized fruit epidermal cells, with irregular arrangements, and variable cuticle thickness (Knoche and Lang 2017). Genetic factors are a major determinant of russet formation, while biotic and abiotic factors can exacerbate seasonal russeting occurrence (Falginella et al. 2015). This is highlighted in sport families, such as ‘Golden Delicious’. ‘Golden Delicious’ is prone to seasonal russeting, whereas its genetic sports ‘Smoothee’ and ‘Empress Spur’ show reduced seasonal russeting, and

‘Sergeant Russet’ has higher seasonal russeting and ‘Razor’ is consistently russeted (Falginella et al. 2015; Wang et al. 2014). Russeting is also associated with anatomical features of the cultivar.

Cuticle disorders in apple have been the focus of several recent research articles, and progress has been made in understanding the cuticle from a genetic and biochemical perspective. The genetic basis for russeting is being explored through QTL analysis and transcriptomics, and new headway is being made to understand suberin biosynthesis and deposition and other metabolic changes (Lashbrooke et al. 2016; Legay et al. 2016, 2017). Suberin is a complex lipophilic polyester, composed of fatty acids and glycerol, with interlinked phenolic acids, relying on several pathways. The changes in the fruit peel due to suberin synthesis and deposition have been associated with changes in the nutraceutical profile of apples. Andre et al. (2012) observed distinct triterpenoid profiles associated with russeting in a survey of 109 apple cultivars. Standard cultivars (cuticle intact) primarily contain ursolic and oleanolic acids, whereas five additional triterpenes in suberized tissues, with health promoting benefits (Andre et al. 2013). To explore other changes associated with russeting in apple, we surveyed the phenolic content in 108 apple cultivars from the USDA National Plant Germplasm System (NPGS) *Malus* collection in Geneva, NY, with an emphasis on phloridzin content. We observed that phloridzin content increases in response to russeting in apple peels but not leaves or flesh, and propose it is likely upregulated by suberin synthesis. We also observed that other flavonoids tend to decrease with russeting in apple fruit.

MATERIALS AND METHODS

Plant Material

Fruit and leaf samples were collected late September 2015 and 2016, from 108 accessions in the USDA-NPGS *Malus* germplasm collection in Geneva, NY, based on prior russeting incidence.

All accessions were grafted on ‘EMLA 7’ rootstock. These accessions included five sport families with variation for russeting: ‘Baldwin’ (*n* = 3), ‘Boskoop’ (*n* = 4), ‘Bouteille’ (*n* = 2), ‘Court Pendu’ (*n* = 3), and ‘Golden Delicious’ (*n* = 14). Sport families are listed in Table 4.1. Sports were identical for seven SSR markers (Gross et al. 2012). Cultivars ‘Arkad Zimnyi’, ‘Bouteille’, ‘Grise Dieppose’, and ‘M.4’ were only collected in 2015, and cultivars ‘Calville du Mont d’Or’, ‘Chestnut Crab’, ‘Dermen Delicious 4x’ and ‘Feys Record’ were only collected in 2016. Peel and flesh were divided, frozen, and stored at -80 ℃. In 2017, fruit and leaf samples were collected from Golden Delicious (mild russet) and three of its sports; ‘Razor’ (complete russet), ‘Sergeant Russet’ (semi-russet), and ‘Empress Spur’ (low russet). Tissues were collected on 6/2/2017, 6/20/17, 7/17/17, and 10/3/17 (maturity). Samples were frozen and stored at -80 ℃. Maturity was determined following the Cornell starch-iodine staining pattern (Blanpied and Silsby 1992).

Table 4.1 List of apple cultivar in five sport families with plant introduction number (PI No.).

Plant.ID	PI No.	Sport Family
‘Colby Baldwin’	589103	‘Baldwin’
‘Loop Russet Baldwin’	589102	‘Baldwin’
‘Tohoku 4’	255900	‘Baldwin’
‘Belle de Boskoop’	300258	‘Boskoop’
‘Bogo Belle de Boskoop’	589237	‘Boskoop’
‘Rouge Belle de Boskoop’	589143	‘Boskoop’
‘St. Anna Rode Boskoop’	223666	‘Boskoop’
‘Bouteille’	187296	‘Bouteille’
‘Bramtot’	158731	‘Bouteille’
‘Court Pendu’	589601	‘Court Pendu’
‘Court Pendu Gris’	589602	‘Court Pendu’
‘Court Pendu Plat’	123960	‘Court Pendu’
‘Clear Gold’	589224	‘Golden Delicious’
‘Empress Spur’	589316	‘Golden Delicious’
‘Golden Delicious’	590184	‘Golden Delicious’
‘Smoothee’	589041	‘Golden Delicious’
‘Golden Delicious SE-69’	589535	‘Golden Delicious’
‘Golden Precoce’	589448	‘Golden Delicious’
‘Goldspur Golden Delicious’	644190	‘Golden Delicious’
‘Puregold’	589966	‘Golden Delicious’
‘Razor’	589136	‘Golden Delicious’

Table 4.1 (Continued)

Plant.ID	PI No.	Sport Family
'Sergeant Russet'	589125	'Golden Delicious'
'Smoothgold'	589904	'Golden Delicious'
'Starkspur'	589210	'Golden Delicious'
'Woodward'	589528	'Golden Delicious'
'Yellow Delicious'	589856	'Golden Delicious'

Russet intensity was scored visually in 2015 and 2016, based on the percent of fruit surface covered, and categorized into one of five categories (0-5%, 10-20%, 30-40%, 50-60%, 70-80%, 90-100%). Russet evaluations for 2,012 *Malus* accessions, collected from 1999 to 2014, was retrieved from the Germplasm Resources Information Network (GRIN; USDA 2014). Though *M.* ×*domestica* was predominant (*n* = 1,174), data included 43 other *Malus* species and hybrids. Variables included percent fruit surface russet (0-100%), russet location (pedicel end, calyx end, pedicel and calyx ends, entire), and russet type (fine, medium, and heavy).

***High Performance Liquid Chromatograph*y**

Fruit was divided into peel and flesh samples and analyzed separately. Fruit phenolics were extracted from 500 mg frozen tissue, and leaf phenolics were extracted from 25 mg frozen tissue. Samples were ground and extracted in 1.5 ml 70% methanol, 2% formic acid and shaken for 15 minutes and then centrifuged for 15 minutes 15,000 rpm. Fruit was divided by peel and flesh and analyzed separately. Fruit phenolics were separated following the method of Zhang et al. (2010), using a Hitachi LaChrom Elite (San Jose, CA, USA) HPLC with (L-2450) 2-channel diode array detector, equipped with an Inertsil ODS-3 column (4.6×250 mm; 5 µm, GL Sciences Inc., Tokyo, Japan) at 30℃ and a Nova-Pack C_{18} guard column (3.9×20mm, 4 µm, Waters, Milford, MA, USA). Leaf phenolics were separated using an isocratic mobile phase of 15% acetonitrile and 10% (98%) formic acid, following the methods of Gutierrez et al (2017).

Data Analysis

Conversion of HPLC data was done using EZChrom Elite 3.1.4 with standard curves for both fruit and leaves for major apple flavonoids and chlorogenic acid with R^2 values > 0.99. Sample values were expressed in mg/g fresh weight. Data analysis was performed using R 3.3.2 (R Core Team 2016) and JMP® (SAS Institute Inc., Cary, NC) version 12.0.1. Pearson's correlation coefficients were calculated for phenolics, and principal component analysis was performed with the correlation matrix. Welch's ANOVA and Games-Howell post-hoc tests were used to account for unbalanced sample size between groups. Data were log transformed to meet statistical assumptions. Data were fit using the following mixed linear model: $Y_{ijk} = \mu + G_i + Y_j + R_k + GY_{ij} + e_{ijk}$, where Y_{ijk} is the phenolic measurement on, genotype *i*, year *j* and russet group *k*; μ is the overall mean; G_i is the random effect of genotype *i*; Y_j is the random effect of year *j* and R_k the random effect of russet group *k*; GY_{ij}, and YR_{jk}, are the random interaction effects of genotype *i*, year *j* and russet group *k*; and e_{ijk} is the residual error. The model was fitted with the following assumptions:

$G_i \sim N(0, \sigma_G^2), Y_j \sim N(0, \sigma_Y^2), R_k \sim N(0, \sigma_R^2), GY_{ij} \sim N(0, \sigma_{GY}^2), YR_{jk} \sim N(0, \sigma_{YR}^2)$, and $e_{ijk} \sim N(0, \sigma_e^2)$.

Repeatability estimates were calculated using the formula:

$r = \frac{\sigma_G^2}{(\sigma_G^2 + \sigma_{GY}^2 + \sigma_e^2)}$ (Volz and McGhie 2011).

RESULTS

Russet Intensity

Fruit russet intensity varied from 0-100%, with the majority of GRIN accessions between 0 to 5%. Only 24 of 2,012 accessions were completely russet (90-100%). Most fruit russet type was 'fine', and occurred more frequently at the pedicel end (Figure 4.1). Seasonal russet

intensity from 2015 to 2016 varied (Table 4.2), but was strongly correlated (r = 0.87, p <0.0001). Correlations were lower between seasons and GRIN data (r = 0.78 for 2015 and 0.70 for 2016, p <0.0001).

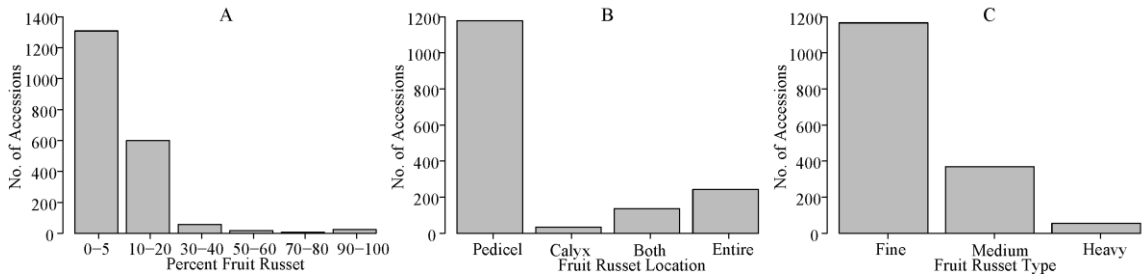


Figure 4.1 Bar plots of GRIN fruit russet data for 2,012 accessions of *Malus* species from the USDA National Plant Germplasm System in Geneva, NY. Percent of fruit russet on fruit surface, sorted into ranges (A). Location of russeting on fruit (B); pedicel or calyx ends, both pedicel and calyx, or entire fruit. Type of fruit russet (C); fine, medium, or heavy.

Table 4.2 Fruit russet intensity across *Malus* germplasm and selected accessions.

Plant Material	Fruit Russet Intensity					
	0-5%	10-20%	30-40%	50-60%	70-80%	90-100%
2015 (n=104)	53	5	6	8	5	27
2016 (n=104)	53	11	5	12	11	12
GRIN (n=106)	48	18	7	8	5	20
GRIN (n=2,012)	1306	599	58	18	7	24

Flavonoid Content Variation

Phenolic compounds chlorogenic acid, cyanidin 3-*O*-galactoside, phloridzin, and quercetin glycosides (flavonols) were detected in fruit and leaf samples. There was significant variation of total and individual phenolic content in fruit peel and leaf tissue. In fruit, peel phloridzin content represented from 3 to 93% of total measured phenolics, ranging from 24.3 to 825.0 μg/g with a mean of 233.5 μg/g. Correlation between peel phloridzin content from 2015 to 2016 was high (r = 0.75, p < 0.0001). Consistent with other reports, phloridzin was the primary phenolic in leaf samples, representing 58 to 93% of total measured phenolics, ranging from 15.3 to 81.64 mg/g with a mean of 48.5 mg/g. Peel phenolic content is listed in Table 4.3. There were weak but significant (p < 0.001) correlations between leaf and peel chlorogenic acid (r = 0.17),

hyperin (r = 0.13), isoquercitrin (r = 0.20), and quercitrin (r = 0.13) content. There was no correlation between peel phloridzin content and leaf phloridzin content (r = 0). There were some correlations between phenolic compounds: phloridzin and chlorogenic acid were moderately correlated in peels (r = 0.46, p <0.0001) and leaves (r = 0.28, p < 0.0001). Quercetin glycosides were all correlated with each other (r = 0.30 to 0.80), and negatively correlated to phloridzin (r = -0.46 to -0.14) in leaf samples.

Table 4.3 Mean fruit russet intensity (%) and peel phenolic (μg/g) composition in 108 *M.* ×*domestica* accessions from the USDA National Plant Germplasm *Malus* collection. Abbreviations: plant introduction number (PI No.), phloridzin (Phl), chlorogenic acid (Chl), hyperin (Hyp), isoquercitrin (Iso), reynoutrin (Rey), avicularin (Avi), quercitrin (Que), cyanidin 3-galactoside (Cya)

Table 4.3 (Continued)

Plant.ID	PI No.	Russet	Phl	Chl	Hyp	Iso	Rey	Avi	Que	Cya
'Aargauer Jubilaums'	134819	0	99.4	19.0	65.0	22.4	64.9	127.1	28.2	2.4
'Abbondanza'	206542	0	169.2	46.0	249.5	64.6	101.9	195.1	53.9	142.2
'Antonovka Ottawa'	590149	0	167.0	28.5	192.7	70.1	91.8	164.5	72.7	16.5
'Arkad Zimnyi'	259446	5	197.0	9.7	32.7	34.2	15.4	35.0	10.0	31.2
'Arkansas Black'	589117	2	753.2	19.9	501.8	156.1	211.4	310.9	99.4	229.0
Ashmead's Kernel'	589654	90	656.5	22.6	13.8	8.2	14.1	31.2	20.6	4.7
'Badami Golden Delicious'	589139	2	177.3	8.0	553.0	198.4	136.3	226.7	157.5	25.7
'Belle de Boskoop'	300258	22	262.4	18.7	352.6	120.6	135.7	217.2	16.5	23.9
'Belle de Crollon'	162544	2	130.3	29.8	79.0	82.1	38.9	104.3	58.2	15.3
'Bellefleur Kitaika'	588777	0	68.7	15.5	177.5	151.2	79.2	130.4	33.8	55.8
'Bellefleur Rekord'	259447	2	95.5	24.4	329.4	76.1	129.4	189.5	60.8	109.5
'Bemali'	589863	15	57.1	25.3	79.6	50.9	25.7	60.7	6.2	2.8
'Bitterforest Apple'	160387	2	62.0	30.1	60.6	17.2	18.5	50.2	22.1	9.8
'Bogo Belle de Boskoop'	589237	48	324.5	15.8	129.0	39.4	56.1	105.1	36.5	10.5
'Boiken'	589134	2	314.4	11.6	306.7	129.4	122.3	256.3	69.0	42.7
'Boskoop Rouge'	224547	48	278.7	10.3	166.3	78.6	63.1	126.9	19.1	67.5
'Bouteille'	187296	50	303.9	69.3	157.1	95.5	44.9	116.4	44.4	9.1
'Bramtot'	158731	80	469.3	130.3	15.0	30.4	13.5	38.7	19.2	2.7
'Calville Du Mont d'Or'	162720	50	596.7	27.4	53.8	25.4	45.1	70.9	10.7	4.7
'Canavial-14'	183961	18	411.5	23.3	65.9	32.3	59.7	86.1	24.2	1.3
'Cap of Liberty'	161830	0	195.3	48.2	193.7	112.6	60.4	116.1	39.7	153.0
'Chenango Strawberry'	589239	2	182.6	9.4	135.9	51.3	85.2	182.5	56.5	300.3
'Chestnut Crab'	588803	50	251.5	40.7	16.2	12.7	17.9	35.8	5.7	21.7
'Clear Gold'	589224	20	181.8	11.1	227.5	101.4	76.2	148.3	49.9	7.3
'Colby Baldwin'	589103	5	181.3	13.6	341.2	113.4	154.0	253.8	44.1	249.4
'Court Pendu'	589601	5	64.5	38.9	98.8	26.8	39.4	75.2	12.9	84.1
'Court Pendu Gris'	589602	95	403.2	45.0	13.7	8.2	12.2	24.8	14.0	8.6
'Court Pendu Plat'	123960	8	81.5	43.5	159.3	39.7	52.2	99.9	30.2	55.6
'Dermen Delicious 4X'	589110	5	112.1	3.2	44.9	15.7	36.1	53.0	3.0	54.8
'Doucet Rouge'	161760	5	162.3	103.6	308.9	180.8	143.6	273.8	81.6	47.6
'Drap dOr Guemene'	131823	68	239.2	13.4	110.4	87.2	43.7	94.2	23.7	68.8
'Dunkerton Late Sweet'	589666	5	66.5	16.2	209.7	39.9	84.4	192.0	33.7	13.3
'Egremont Russet'	260000	100	538.1	53.7	18.4	25.6	41.4	84.1	30.9	2.2
'Empress Spur’	589316	20	145.6	10.4	193.5	72.2	64.6	97.5	41.7	8.9
'Fall Russet'	589012	72	153.6	28.1	39.4	23.0	14.9	30.7	5.9	11.2
'Fenouillet de Ribours'	590126	90	394.6	3.5	30.2	16.0	21.7	31.9	17.3	3.9
'Feys Record'	125686	5	51.3	6.5	69.9	34.0	35.5	67.3	9.8	123.7
'Frequin Lacaille'	247314	45	322.5	115.2	149.2	104.3	59.3	144.8	24.2	10.7
'Gilliflower'	590023	18	346.7	6.3	340.1	192.2	106.4	246.9	74.7	131.3
'Glockenapfel'	134808	5	225.7	9.4	145.1	55.0	57.5	112.2	46.1	15.0
'Golden Delicious'	590184	5	104.3	8.6	105.3	39.1	51.3	96.2	30.6	2.0
‘Smoothee’	589041	5	81.3	6.8	87.4	31.7	41.8	72.4	48.2	0.3
'Golden Delicious SE-69'	589535	30	189.7	13.0	92.8	31.6	39.3	72.0	20.1	2.4
'Golden Harvey'	590128	35	88.6	5.7	32.0	21.0	22.6	34.0	10.8	2.8
'Golden Nugget'	589126	58	65.3	7.0	159.0	88.7	77.4	139.1	6.2	32.3
'Golden Precoce'	589448	12	133.3	9.3	206.8	83.0	67.9	125.1	37.4	2.6
'Golden Russet'	589892	78	291.3	13.1	14.9	16.6	18.0	36.3	13.2	1.4

Table 4.3 (Continued)

Plant.ID	PI No.	Russet	Phl	Chl	Hyp	Iso	Rey	Avi	Que	Cya
'Goldspur'	644190	18	114.8	9.0	135.1	60.8	52.8	98.3	48.5	9.8
'Graue Renette Von Zabergau'	589233	95	664.9	21.4	10.0	12.2	12.6	28.2	18.7	1.1
'Grimes Golden'	588791	0	201.9	11.3	477.4	302.6	191.8	303.5	79.1	14.6
'Grise Dieppoise'	125746	95	404.1	33.9	72.5	34.2	28.3	60.4	21.6	4.2
'Grosse Launette'	161761	80	260.0	80.3	67.5	74.5	47.5	95.8	21.2	21.8
'Honeycrisp'	644174	2	38.9	8.3	88.4	13.0	46.9	46.5	25.9	98.8
'Hubbardston Nonsuch'	589202	2	104.3	12.9	182.6	234.0	115.1	171.5	33.8	62.1
'Hudsons Golden Gem'	590157	85	147.6	28.1	31.8	22.2	22.3	37.9	13.9	0.9
'Isle of Wight Pippin'	590131	88	368.2	31.3	232.4	208.8	109.1	238.3	48.3	5.4
'Keukelaar Greening'	589248	2	68.6	9.0	87.2	43.6	60.8	117.1	30.0	39.8
'Lemoen'	161407	30	226.8	13.7	78.2	25.0	41.9	97.5	24.2	14.6
'Loop Russet Baldwin'	589102	78	355.0	21.8	99.2	46.6	41.8	81.8	37.5	23.5
'Lorna Doone'	161840	5	58.2	6.1	144.7	48.5	62.4	147.0	9.2	169.5
'Lowry'	589150	2	79.2	15.3	211.1	74.6	89.1	153.0	38.6	77.3
'M.4'	589347	60	202.9	40.6	217.2	69.2	83.6	190.0	48.5	10.6
'Malinda'	644176	5	93.8	16.5	117.2	70.9	59.2	111.7	46.3	16.8
'Martha'	589061	2	73.2	19.7	124.1	45.5	55.8	94.5	17.3	207.0
'McNicholas Greening'	589106	5	63.6	10.6	234.7	230.7	113.7	174.9	32.4	27.7
'Merton No. 789'	133101	5	138.5	2.8	119.6	38.8	59.5	111.0	71.0	89.3
'Metais'	173983	60	692.2	139.3	135.2	53.2	62.4	119.3	29.4	46.5
'Montgomery'	589256	2	130.6	2.8	436.6	107.8	150.8	276.9	92.9	227.4
'Muz-alma'	30326	2	69.5	13.2	129.6	52.5	37.6	84.2	23.4	4.0
'Oxbo'	127703	2	81.8	17.7	206.6	59.5	84.3	154.5	57.7	13.3
'Padleys Pippin'	199420	92	307.7	23.1	10.0	8.4	9.3	15.4	11.9	3.1
'Paradiso'	589323	2	108.5	5.9	339.7	214.3	130.5	303.4	39.9	122.5
'Pepinka Litowska'	589116	0	103.8	26.9	108.3	101.8	46.9	107.3	30.5	29.1
'Pigeonnet Rouge'	132273	2	779.1	15.5	720.5	312.5	184.8	437.4	114.5	139.3
'Pine Golden Pippin'	157734	100	267.2	38.8	17.4	20.9	23.6	41.0	19.0	3.1
'Pitmaston Pineapple'	280029	72	53.3	5.7	20.6	15.2	11.1	17.9	9.7	1.9
'Pitmaston Pineapple'	279323	80	79.1	11.4	61.9	36.5	22.0	37.0	14.1	6.1
'Pomme Cloche'	134668	2	203.0	8.5	137.8	50.4	48.5	97.5	22.3	42.0
'Pomme Grise'	589242	88	174.5	14.9	38.7	34.3	16.4	29.0	10.7	0.5
'Primate'	589229	0	24.3	8.0	9.8	8.0	8.7	22.3	14.4	2.4
'Puregold'	589966	2	119.4	7.0	184.2	106.0	61.5	120.6	73.8	63.3
'Razor Golden Delicious'	589136	100	426.1	21.4	13.4	24.4	22.7	32.8	16.8	1.4
'Red Cinnamon'	207636	0	112.2	14.8	258.5	62.6	114.9	191.1	31.8	59.5
'Reine des Pommes'	132571	62	825.0	178.2	44.0	30.4	28.2	49.5	36.1	53.3
'Reine des Reinettes x1700'	279326	12	282.8	10.3	175.2	144.6	63.5	149.6	16.8	101.7
'Reinette Grise de Portugal'	590138	92	527.6	70.7	17.1	19.7	19.6	28.7	15.4	2.3
'Reinette Tres Tardive'	162741	98	672.8	29.8	15.0	24.7	27.8	45.1	25.0	0.6
'Renetta Dorata'	104034	85	165.4	21.2	8.5	9.4	7.9	18.7	8.3	3.1
'Rouge Belle de Boskoop'	589143	95	463.2	12.6	56.8	19.1	31.9	65.0	29.8	7.0
'Roxbury Russet'	588971	75	659.9	22.2	46.7	31.5	45.0	66.2	32.2	3.5
'Sam Young'	264559	75	389.0	36.5	126.9	64.6	40.4	106.8	29.5	29.7
'Sergeant Russet’	589125	82	271.9	9.1	24.2	18.1	16.2	26.1	12.7	0.7
'Severnii Bujon'	107245	5	118.3	21.0	429.3	104.3	156.7	250.1	82.6	212.5

Table 4.3 (Continued)

Plant.ID	PI No.	Russet	Phl	Chl	Hyp	Iso	Rey	Avi	Que	Cya
'Severny Sinap K-21.39'	589511	0	77.1	13.2	102.1	42.6	36.0	66.3	45.7	18.1
'Sir Prize'	589961	2	160.3	14.2	612.9	247.4	180.5	310.3	163.9	2.2
'Smoothgold'	589904	2	98.9	7.6	193.2	68.1	59.7	112.9	70.3	7.1
'St. Anna Rode Boskoop'	223666	60	460.3	12.1	93.7	31.3	50.9	68.7	28.8	21.1
'St. Edmunds Russet'	589658	100	389.6	18.0	5.7	12.2	7.7	12.7	10.6	1.6
'Starkspur'	589210	28	130.4	11.5	269.7	106.3	87.0	155.3	29.5	8.7
'Stone Pippin'	133750	2	91.6	5.5	276.0	160.1	132.2	180.2	102.8	35.7
'Tohoku 4'	255900	5	139.7	7.8	263.5	95.2	118.4	191.6	25.9	350.8
'Vogelcalville'	188525	5	109.8	3.1	21.8	23.4	13.1	27.8	10.2	1.5
'Wealthy'	588788	2	114.8	11.0	154.1	39.5	49.6	100.5	44.2	265.2
'Whitney Russet King'	589194	90	319.2	20.7	96.8	42.4	77.2	179.1	25.6	10.3
'Winston'	589675	2	161.2	17.3	471.7	214.0	149.0	329.0	61.3	75.2
'Woodward'	589528	18	78.5	6.8	101.6	43.8	31.8	58.8	16.6	6.1
'Yellow Delicious'	589856	10	149.4	7.7	323.6	115.7	98.9	183.9	77.0	3.8
'Zlatna Resistenta'	589539	2	113.5	15.2	413.7	190.0	148.9	264.2	101.6	13.7
Mean		32.9	233.5	24.5	157.7	73.4	64.8	121.1	37.6	47.4
Min		0	24.3	2.8	5.7	8.0	7.7	12.7	3.0	0.3
Max		100	825.0	178.2	720.5	312.5	221.4	437.4	163.9	350.8

Principal component analysis revealed a relationship between quercetin glycosides, phloridzin, and fruit russet intensity in peel samples (Figure 4.2). This relationship was not apparent in leaf or flesh samples. PC1 separates accessions based on quercetin glycoside content and fruit russeting, and PC 2 separates accessions based on phloridzin and chlorogenic acid content and fruit russeting. Together, PCs 1 and 2 explain 65.9% of the observed variation. Eigenvectors, denoted as arrows in Figure 4.2, move in opposite directions on PC1 for russeting and quercetin glycosides, and in the same direction for phloridzin and russeting on PC2. Compounds avicularin, hyperin, isoquercitrin, and reynoutrin explain more of the variance in PC1 than quercitrin and cyanidin-3-O-galactoside, based on the length of their arrows. Accessions in the top-right quadrant tend to have high quercetin glycoside and phloridzin content, with low russeting. Likewise, accessions in the top-left quadrant tend to have high phloridzin content and higher russeting intensity, but lower quercetin glycoside content. PCA

suggests that quercetin glycosides are highly correlated with each other, likewise with phloridzin and chlorogenic acid. Phloridzin content was moderately correlated with fruit russet intensity in apple peels (r = 0.53, p <0.0001). Chlorogenic acid content had a weak but significant correlation to russeting (r = 0.21, p < 0.0001). Other flavonoids, including, avicularin, cyanidin-3-O-galactoside, hyperin, isoquercitrin, and reynoutrin were negatively correlated with russeting, with r = -0.42, -0.32, -0.44, -0.33, -0.43, respectively, with p < 0.0001. Figure 4.3 shows a trend of increasing log transformed phloridzin and chlorogenic acid content in peel based on russet groups. There were significant differences between russeting groups in the log transformed phloridzin (F (5,92.2) = 64.6, p < 0.0001) and chlorogenic acid (F (5,88.2) = 11.2, p < 0.0001) peel content.. Differences between groups were determined with Games-Howell post-hoc tests (Table 4.4) Cultivars with 70–80% had the highest overall mean phloridzin content. Within the 0 to 5% russet group, there were outliers for high and low phloridzin content. Conversely, some 90 to 100% russeted material had low phloridzin content.

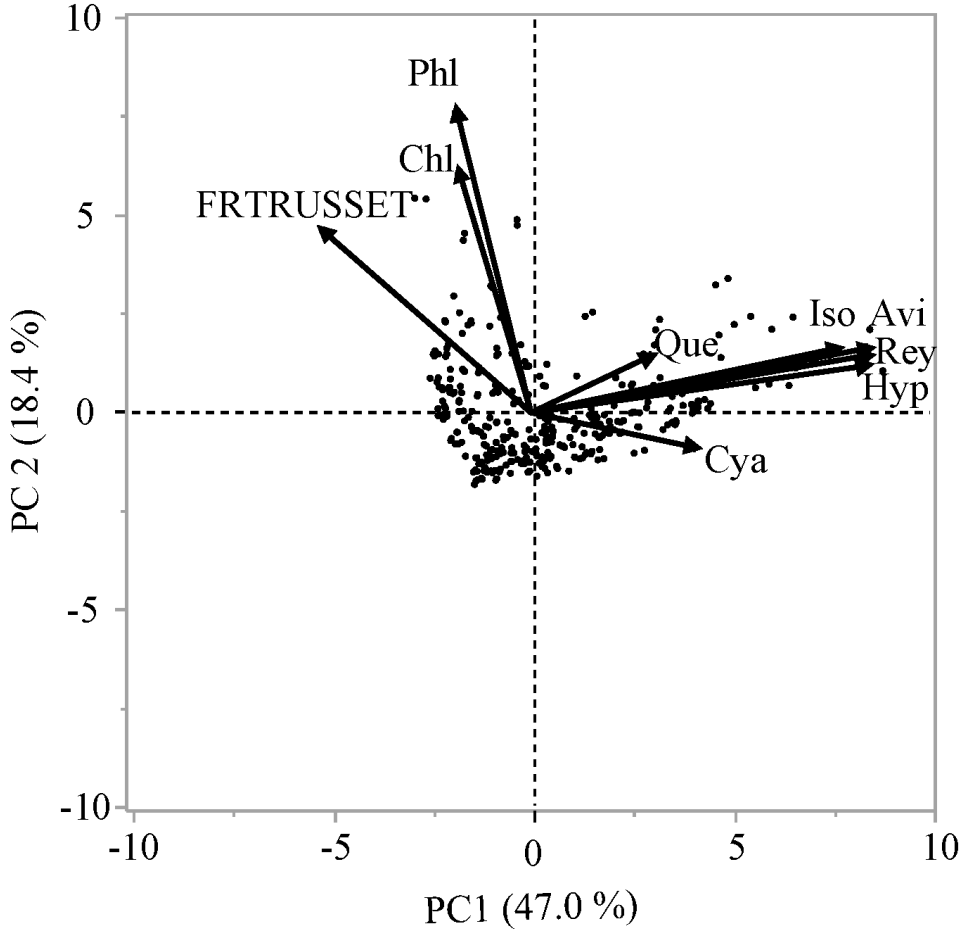


Figure 4.2 Principal component analysis of individual phenolics and russet intensity from 108 *Malus* accessions. Phloridzin and chlorogenic acid are associated with fruit russet intensity. PCs 1 and 2 explain 65.9% of the variation among accessions. Arrows show direction and weight of variables. Abbreviations: fruit russet intensity (FRTRUSSET), phloridzin (Phl), chlorogenic acid (Chl), hyperin (Hyp), isoquercitrin (Iso), reynoutrin (Rey), avicularin (Avi), quercitrin (Que), cyanidin 3-galactoside (Cya).

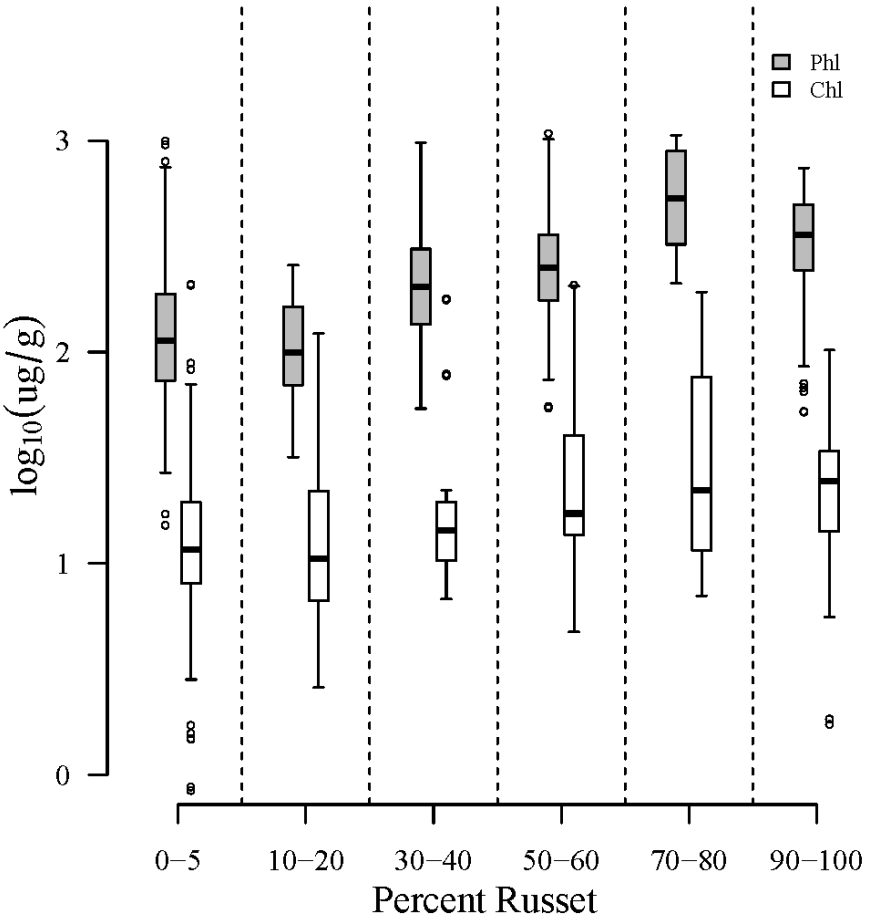


Figure 4.3 Boxplot of log transformed phloridzin (gray) and chlorogenic acid (white) content from fruit peels by russet intensity in 108 apple cultivars

Table 4.4 Mean log transformed apple peel phenolic content across russet intensity groups. Content was significantly different across groups based on ANOVA. Tukey's HSD post-hoc tests showed differences between specific groups; groups not connected by a letter are significantly different. Abbreviations: phloridzin (Phl), chlorogenic acid (Chl), hyperin (Hyp), isoquercitrin (Iso), reynoutrin (Rey), avicularin (Avi), quercitrin (Que), cyanidin 3-galactoside (Cya).

Russet	Phl	Chl	Hyp	Iso	Rey	Avi	Que	Cya
0-5%	2.07 a	1.10 a	2.13 a	1.76 a	1.80 a	2.08 a	1.13 a	1.49 a
10-20%	2.00 a	1.09 a	2.04 a	1.68 ab	1.69 ab	1.98 ab	0.69 b	0.95 b
30-40%	2.32 b	1.26 ab	1.87 ab	1.58 ab	1.58 abc	1.93 abc	1.08 ab	1.13 ab
50-60%	2.36 b	1.36 b	1.69 b	1.48 bc	1.50 cd	1.69 bc	1.17 b	1.09 ab
70-80%	2.71 b	1.49 b	1.57 bc	1.44 bc	1.45 bcd	1.79 cd	1.28 b	1.13 ab
90-100%	2.51 c	1.35 b	1.29 c	1.22 c	1.29 d	1.56 d	1.21 b	0.69 b

Linear modeling of phloridzin content in peel, controlling for genotype (G), date (D), russet (R), and genotype × date (G×D), explained a high percentage of the variation in phloridzin content, where R and G explained 35% and 41% of phloridzin content variation, respectively. For chlorogenic acid, R and G explained 5% and 77% of the content variation, respectively. For total flavonols and cyanidin 3-*O*-galactoside, R explained 19% and 13% respectively, and G explained 30 and 35%, respectfully. Repeatability estimates for phloridzin, chlorogenic acid, total flavonols, and cyanidin 3-galactoside were 0.63, 0.80, 0.37, and 0.42, respectively. Leaf phenolic content was fit using this same model. R was not significant for any of the compounds analyzed. Variance component estimates are given in Table 4.5.

Table 4.5 Variance component estimates of and leaf peel phenolic data on 108 apple cultivars.

Tissue	Source	Phl	Chl	Flavonols	Cya[a]
Peel	Russet (R)	0.056 ± 0.24	0.007 ± 0.08	0.036 ± 0.19	0.071 ± 0.27
	Genotype (G)	0.065 ± 0.25	0.117 ± 0.34	0.056 ± 0.24	0.192 ± 0.44
	Year (Y)	0.003 ± 0.06	n.s.[b]	0.003 ± 0.06	0.026 ± 0.16
	G×Y	0.008 ± 0.09	0.009 ± 0.10	0.058 ± 0.24	0.193 ± 0.44
	Residual	0.056 ± 0.24	0.007 ± 0.08	0.036 ± 0.19	0.071 ± 0.27
	Total variation explained by R (%)	35	5	19	13
	Total variation explained by G (%)	41	77	30	35
Leaf	Russet (R)	n.s.	n.s.	n.s.	
	Genotype (G)	0.0037 ± 0.06	0.021 ± 0.14	0.01 ± 0.1	
	Year (Y)	0.0003 ± 0.02	0.025 ± 0.16	0.003 ± 0.05	
	G×Y	0.0047 ± 0.07	0.032 ± 0.18	0.008 ± 0.09	
	Residual	0.0055 ± 0.07	0.009 ± 0.09	0.006 ± 0.08	
	Total variation explained by R (%)	0	0	0	
	Total variation explained by G (%)	26	24	37	

[a] Cya was not measured in leaf samples
[b] N.S = Not significant

To further isolate the effect of russeting on phloridzin content, the data were subset to only include sport families with russeting variation. Five sport families were identified within

this subset; ‘Bouteille’, ‘Boskoop’, ‘Baldwin’, ‘Court Pendu’, and ‘Golden Delicious’. Russeting variation and among genetic sports is illustrated in Figure 4.4. Russeting intensity ranged from 0 to 100% and phloridzin ranged from 44.9 to 1062 μg/g across sports (Fig 5). The same mixed linear model was fit with these sport families, with sport representing the genotype of its members. The effect of R, G, and G×Y explained 59%, 13% and 12% of the peel phloridzin variation, respectively. R, G, and G×Y explained 4.5% and 84% and 2.5% of the chlorogenic acid variation, and R and G explained 31 and 5% of the variation of total flavonols. Repeatability estimates were 0.31 and 0.88 for phloridzin and chlorogenic acid.

Figure 4.4 Apple russet variation between ‘Golden Delicious’ (mild to no russet) and its sports, ‘Sergeant Russet’ (semi-russet) and ‘Razor’ (complete russet) on 1 cm^2 grids.

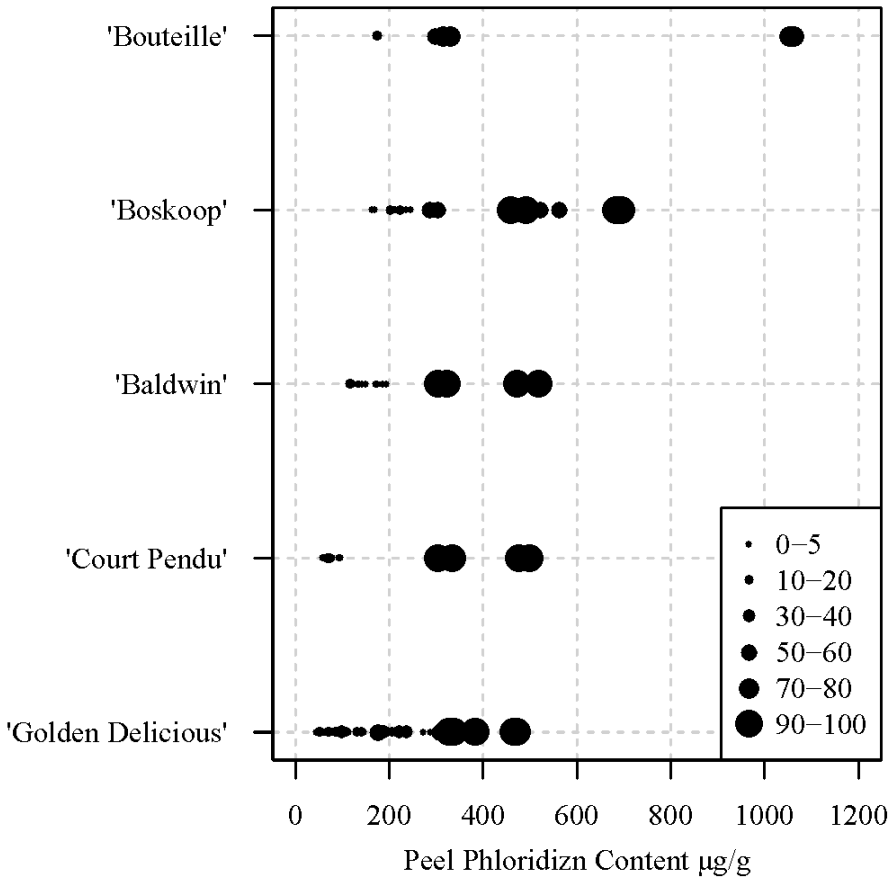


Figure 4.5 Peel phloridzin content (µg/g) variation from 2015 to 2016 across five sport families and percent of russet incidence. Point circumference represents russet incidence as a percentage of fruit surface.

Phenolic content and russet intensity was measured in ‘Golden Delicious’ and three of its sports (Tables 6 and 7). On 6/2/17 (several weeks after full bloom), no russeting was evident across accessions. By June 20, 2017 ‘Razor’ was 80 to 90% russet and ‘Sergeant Russet’ was 20 to 30%. By maturity (10/3/2017), ‘Sergeant Russet’ was 80 to 90% russeted. Slight russeting occurred around the pedicel of ‘Golden Delicious’ and ‘Empress Spur’ throughout the season. Peel and flesh phloridzin content was highest on June 2, 2017 in all accessions and decreased throughout the season. Likewise, chlorogenic acid and total flavonols decreased throughout the season in peel and flesh samples. Peel phloridzin concentrations decreased at a lower rate in ‘Razor’ and ‘Sergeant Russet’ than in ‘Golden Delicious’ and ‘Empress Spur’. Conversely, peel

total flavonols decreased at a higher rate in these russet sports. Among accessions, phloridzin content was highest in ‘Razor’ and ‘Sergeant Russet’, even prior to russet development. Leaf phloridzin content and total flavonols were stable throughout the season across sports. Leaf chlorogenic acid decreased throughout the season, but increased significantly on 10/3/17. Season increase of leaf chlorogenic acid was reported in apple (Mikulič Petkovšek et al. 2003). Figure 4.6 illustrates the metabolite changes in leaves and fruit samples in these sports, throughout the season.

Table 4.6 Russet intensity in ‘Golden Delicious’ and its sports on four dates throughout fruit development.

Plant ID	6/2/17	6/20/17	7/17/17	10/3/17
‘Empress Spur’	0	0	0-5	0-5
‘Golden Delicious’	0	0	0-5	0-5
‘Sergeant Russet’	0	20-30	70-80	80-90
‘Razor’	0	80-90	100	100

Table 4.7 Phenolic content (μg/g) in peel, flesh, and leaf tissues of 'Golden Delicious' and its sports across four dates during fruit development.

Plant.ID	Tissue	Phenolic	6/2/2017	6/20/2017	7/17/2017	10/3/2017
'Empress Spur'	Peel	Phl	3745.0	812.6	241.2	71.1
		Chl	3178.8	1492.8	415.4	115.8
		Flav	5006.5	3908.0	3000.2	1301.7
	Flesh	Phl	2794.9	261.8	51.9	10.8
		Chl	5538.7	1987.0	571.8	161.7
		Flav	194.8	192.2	34.7	17.5
	Leaves	Phl†	37.8	47.7	43.4	49.4
		Chl	937.9	586.1	230.9	4173.0
		Flav†	13.3	12.2	15.1	14.0
'Golden Delicious'	Peel	Phl	3487.0	695.0	270.8	84.9
		Chl	3650.0	1828.0	527.8	156.1
		Flav	5528.2	4255.4	3347.8	3156.3
	Flesh	Phl	1314.0	220.2	43.3	12.9
		Chl	5318.3	2842.1	549.7	241.6
		Flav	203.1	263.6	39.5	19.3
	Leaves	Phl†	50.2	51.8	58.7	59.3
		Chl	1058.2	438.0	360.4	3169.2
		Flav†	17.9	16.3	16.2	20.3
'Razor'	Peel	Phl	4871.6	2370.8	1779.0	641.7
		Chl	4536.8	4199.2	1179.0	270.0
		Flav	2982.4	1871.0	1020.4	312.8
	Flesh	Phl	1730.5	311.9	62.1	11.8
		Chl	6032.4	3852.0	917.08	195.8
		Flav	181.7	232.5	37.2	21.1
	Leaves	Phl†	49.5	55.2	60.4	56.8
		Chl	1229.1	495.2	390.8	2773.3
		Flav†	15.6	21.4	16.5	15.8
'Sergeant Russet'	Peel	Phl	5132.4	1524.0	664.3	401.1
		Chl	3809.8	2561.8	882.2	322.1
		Flav	3297.8	2288.2	852.4	437.4
	Flesh	Phl	3130.7	395.5	42.5	9.6
		Chl	5619.8	3328.6	694.2	214.2
		Flav	202.9	244.0	30.0	12.1
	Leaves	Phl†	39.8	56.6	52.3	46.4
		Chl	1632.2	668.2	314.7	2316.4
		Flav†	10.3	10.6	9.0	12.9

† Concentrations reported in mg/g

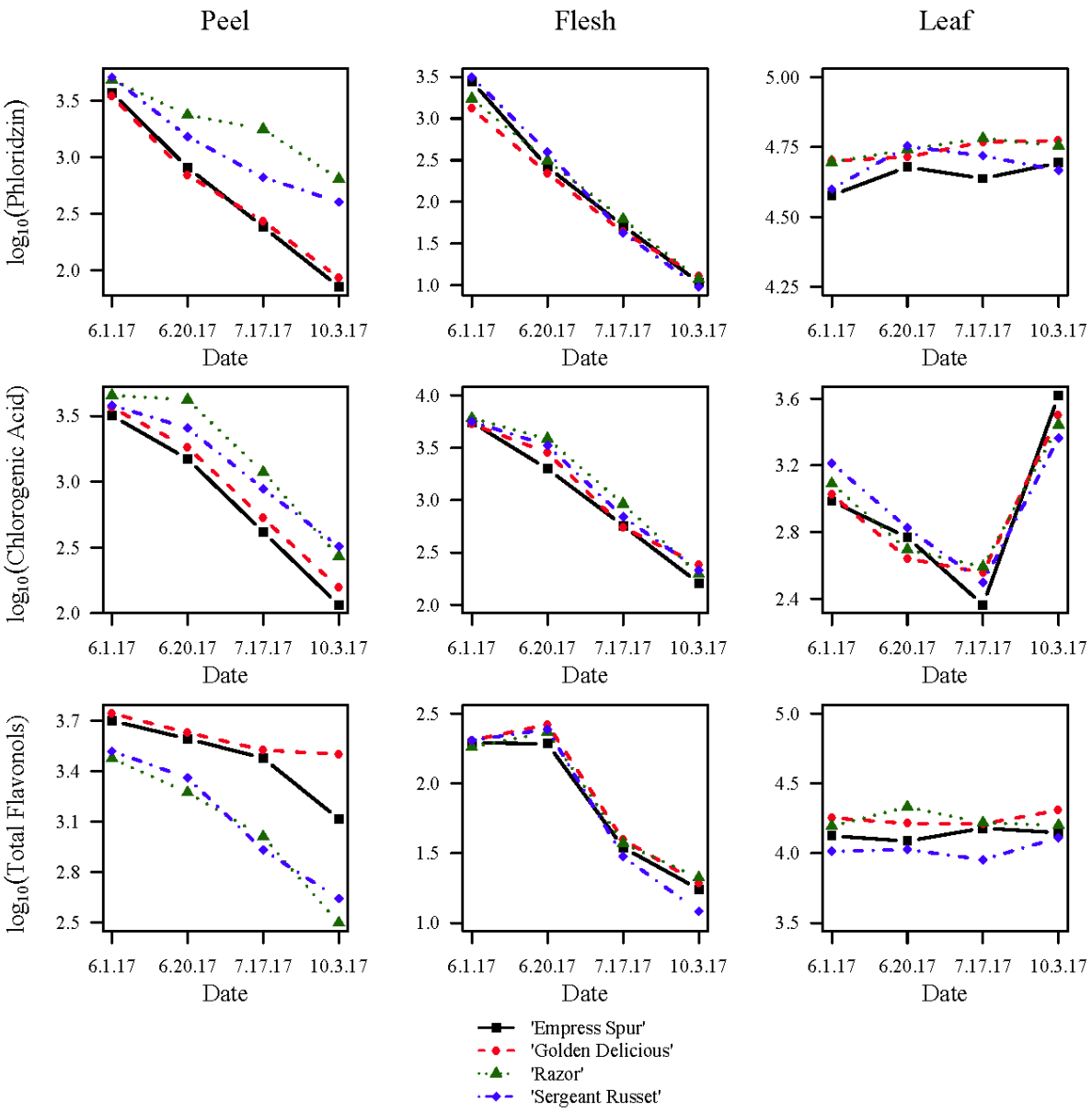


Figure 4.6 Changes in log transformed phloridzin, chlorogenic acid, and total flavonols content fruit (peel and flesh) and leaves, in 'Golden Delicious' and its sports across four dates.

DISCUSSION

Fruit russet is an important disorder in apple fruits that significantly affects commercial value. Though seasonal russeting is prevalent in certain cultivars such as 'Golden Delicious', evaluations across 2,012 germplasm accessions suggest prevalence >20% surface russeting is rare in apple. Seasonal correlations of fruit russeting were high among the 108 accessions used in

this study. Peel phenolic content was significantly associated with russeting in apple across germplasm accession, russet sports, and throughout fruit development. Russeting had the highest impact on phloridzin content. Linear modeling revealed that russet and genotype explained 38 and 36% of the variation in peel phloridzin content, repetively. Though there was an association with russeting, we also identified high phloridzin content in select non-russet accessions. Chlorogenic acid was also associated with russeting intensity, but to a lesser degree than phloridzin. Conversely, fruit russeting had a negative impact on flavonols and cyanidin 3-galactoside content in the peel.

Wang et al. (2014) observed decreased chlorogenic acid content in 'Golden Delicious' and a non-russeting sport 'Feng Shuai' in early fruit development and about equal at maturity. However, in our study, chlorogenic acid content was higher in the russet sports 'Razor' and 'Sergeant Russet' than 'Golden Delicious' throughout development and maturity. High phloridzin content was also reported by Andre et al. (2012) in peels of russet cultivars 'Merton Russet' (107.6 μg/g), 'Wilmont Russet' (107.6 μg/g), and Belle de Boskoop (183.2 μg/g), and by Anastasiadi et al. (2017) in 'Egremont Russet', and 'Wheeler's Russet'.

Phenolic content, including phloridzin, decreased in apple fruit throughout the season, consistent with the research of Baldi et al. (2017) and Stanger et al. (2017). However, phloridzin decreased at a slower rate in russeted apples. Additionally, at early fruit development (6/2/17) there was no apparent russeting, though phloridzin content was much higher in sports 'Razor' and 'Sergeant Russet'. It is likely that changes in gene expression and regulation of biochemical pathways were occurring before russeting was evident on June 6, 2017. Only 18 days later (6/20/17) 'Razor' was nearly completely russeted, and 'Sergeant Russet' had mild russeting. While the development of russet in apple is only partially understood, microcracking of the

cuticle is the first sign of russet development, and can be observed within 20 days after petal fall (Knoche et al, 2011).

Environmental conditions play a significant role in russeting, and horticultural practices are used to prevent russet incidence in commercial orchards (Barceló-Vidal et al. 2013; Ginzberg et al. 2014). Genetic factors also contribute to fruit russet; a major QTL was identified on apple linkage group 12, and apple transcriptomics identified key genes related to russeting (Falginella et al. 2015; Legay et al. 2016; Wang et al. 2014).

Phenolic compounds, including chlorogenic acid are key components of lignin and suberin synthesis. However, expression of lignin/suberin pathways may inhibit specific classes of phenolics (Baldi et al. 2017). Phloridzin is unique to apple and its relationship to suberin and lignin synthesis has not been previously described. Lashbrooke et al. (2016) described suberin biosynthesis and upregulation of phenylpropanoid pathway genes in apple. 4-coumarate:COA ligase (4CL) is one of the initial steps toward suberin and lignin synthesis and may affect phloridzin content (Baldi et al. 2017). Chagné et al. (2012) and Khan et al. (2012) identified a QTL controlling multiple phenolics, including phloridzin, on apple linkage group 16, and proposed leucoanthocyanin reductase as a candidate gene. Although downstream of phloridzin synthesis, leucoanthocyanin reductase is a feedback regulator of 4CL, a critical step for *p*-coumaroyl-CoA synthesis which is a branch point, separating phloridzin from other flavonoids. If 4CL upregulates phloridzin synthesis, it may partially explain why phloridzin concentrations are much higher in vegetative tissues and are correlated with russeting in fruit. Legay et al. (2017) observed differential gene expression and metabolic profiles in waxy and russeted portions of ‘Cox’s Orange Pippin’. Their study emphasized key genes related to suberin synthesis and deposition, some of which may be related to phloridzin synthesis.

CONCLUSIONS

Phloridzin is an important phenolic in apple. However, it is partitioned disproportionately in vegetative and fruit tissues, with nearly a 1000 fold difference between fruit and leaves. Fruit peel phloridzin content was positively correlated to fruit russeting in *Malus* germplasm accessions. Chlorogenic acid was also associated positively with fruit russeting, while flavonols were associated negatively. Genetic factors and russeting explained a high percentage of variation of fruit peel phloridzin content in 108 accessions, and russeting explained the majority of the variation in genetic sports. We propose that enzymes associated with lignin and suberin synthesis, such as 4CL, in apple are key regulators of phloridzin synthesis in apple.

REFERENCES

Ampomah-Dwamena C, Dejnoprat S, Lewis D, Sutherland P, Volz RK, Allan AC (2012) Metabolic and gene expression analysis of apple (*Malus*× *domestica*) carotenogenesis. J Exp Bot 63:4497–4511. doi:10.1093/jxb/ers134

Anastasiadi M, Mohareb F, Redfern SP, Berry M, Simmonds MSJ, Terry LA (2017) Biochemical profile of heritage and modern apple cultivars and application of machine learning methods to predict usage, age, and harvest season. J Agric Food Chem. doi:10.1021/acs.jafc.7b00500

Andre CM, Greenwood JM, Walker EG, Rassam M, Sullivan M, Evers D, Perry NB, Laing WA (2012) Anti-inflammatory procyanidins and triterpenes in 109 apple varieties. J Agric Food Chem 60:10546–10554. doi:10.1021/jf302809k

Andre CM, Larsen L, Burgess EJ, Jensen DJ, Cooney JM, Evers D, Zhang J, Perry NB, Laing WA (2013) Unusual immuno-modulatory triterpene-caffeates in the skins of russeted varieties of apples and pears. J Agric Food Chem 61:2773–2779. doi:10.1021/jf305190e

Baldi P, Moser M, Brilli M, Vrhovsek U, Pindo M, Si-Ammour A (2017) Fine-tuning of the flavonoid and monolignol pathways during apple early fruit development. Planta 245:1021–1035. doi:10.1007/s00425-017-2660-5

Barceló-Vidal C, Bonany J, Martín-Fernández JA, Carbó J (2013) Modelling of weather parameters to predict russet on 'Golden Delicious' apple. J Hortic Sci Biotechnol 88:624–630. doi:10.1080/14620316.2013.11513016

Dare AP, Tomes S, Cooney JM, Greenwood DR, Hellens RP (2013) The role of enoyl reductase genes in phloridzin biosynthesis in apple. Plant Physiol Biochem 72:54–61. doi:10.1016/j.plaphy.2013.02.017

Dare AP, Yauk Y-K, Tomes S, McGhie TK, Rebstock RS, Cooney JM, Atkinson RG (2017) Silencing a phloretin-specific glycosyltransferase perturbs both general phenylpropanoid biosynthesis and plant development. Plant J Cell Mol Biol 91:237–250. doi:10.1111/tpj.13559

De Paepe D, Valkenborg D, Noten B, Servaes K, Diels L, Loose MD, Van Droogenbroeck B, Voorspoels S (2015) Variability of the phenolic profiles in the fruits from old, recent and new apple cultivars cultivated in Belgium. Metabolomics 11:739–752. doi:10.1007/s11306-014-0730-2

Falginella L, Cipriani G, Monte C, Gregori R, Testolin R, Velasco R, Troggio M, Tartarini S (2015) A major QTL controlling apple skin russeting maps on the linkage group 12 of 'Renetta Grigia di Torriana'. BMC Plant Biol 15:150

Fang T, Zhen Q, Liao L, Owiti A, Zhao L, Korban SS, Han Y (2017) Variation of ascorbic acid concentration in fruits of cultivated and wild apples. Food Chem 225:132–137. doi:10.1016/j.foodchem.2017.01.014

Gaucher M, Dugé de Bernonville T, Guyot S, Dat JF, Brisset M-N (2013) Same ammo, different weapons: Enzymatic extracts from two apple genotypes with contrasted susceptibilities to fire blight (*Erwinia amylovora*) differentially convert phloridzin and phloretin *in vitro*. Plant Physiol Biochem 72:178–189. doi:10.1016/j.plaphy.2013.03.012

Ginzberg I, Fogelman E, Rosenthal L, Stern RA (2014) Maintenance of high epidermal cell density and reduced calyx-end cracking in developing 'Pink Lady' apples treated with a combination of cytokinin 6-benzyladenine and gibberellins A_4+A_7. Sci Hortic 165:324–330

Ginzberg I, Stern RA (2016) Strengthening fruit-skin resistance to growth strain by application of plant growth regulators. Sci Hortic 198:150–153

Gosch C, Halbwirth H, Kuhn J, Miosic S, Stich K (2009) Biosynthesis of phloridzin in apple (*Malus domestica* Borkh.). Plant Sci 176:223–231. doi:10.1016/j.plantsci.2008.10.011

Gosch C, Halbwirth H, Stich K (2010) Phloridzin: biosynthesis, distribution and physiological relevance in plants. Phytochem 71:838–843. doi:10.1016/j.phytochem.2010.03.003

Gross BL, Volk GM, Richards CM, Forsline PL, Fazio G, Chao CT (2012) Identification of 'duplicate' accessions within the USDA-ARS National Plant Germplasm System *Malus* collection. J Am Soc Hortic Sci 137:333–342

Gutierrez BL, Zhong G-Y, Brown SK (2017) Genetic diversity of dihydrochalcone content in *Malus* germplasm. Manuscript submitted for publication to Genetic Resources and Crop Evolution on 8/15/17

Hyson DA (2011) A comprehensive review of apples and apple components and their relationship to human health. Adv Nutr Int Rev J 2:408–420. doi:10.3945/an.111.000513

Khanal BP, Knoche M, Bussler S, Schlueter O (2014) Evidence for a radial strain gradient in apple fruit cuticles. Planta 240:891–897. doi:10.1007/s00425-014-2132-0

Knoche M, Khanal BP, Stopar M (2011) Russeting and microcracking of 'Golden Delicious' apple fruit concomitantly decline due to gibberellin A_{4+7} application. J Am Soc Hortic Sci 136:159–164

Knoche M, Lang A (2017) Ongoing growth challenges fruit skin integrity. Crit Rev Plant Sci 36:190–215. doi:10.1080/07352689.2017.1369333

Lashbrooke J, Aharoni A, Costa F (2015) Genome investigation suggests MdSHN3, an APETALA2-domain transcription factor gene, to be a positive regulator of apple fruit cuticle formation and an inhibitor of russet development. J Exp Bot erv366

Lashbrooke J, Cohen H, Levy-Samocha D, Tzfadia O, Panizel I, Zeisler V, Massalha H, Stern, A, Trainotti L, Schreiber L, Costa F, Aharoni A (2016) MYB107 and MYB9 homologs regulate suberin deposition in angiosperms. Plant Cell 28:2097–2116

Legay S, Cocco E, André CM, Guignard C, Hausman J-F, Guerriero G (2017) Differential lipid composition and gene expression in the semi-russeted 'Cox Orange Pippin' apple variety. Front Plant Sci 8. doi: 10.3389/fpls.2017.01656

Legay S, Guerriero G, André C, Guignard C, Cocco E, Charton S, Boutry M, Rowland O, Hausman J-F (2016) *MdMyb93* is a regulator of suberin deposition in russeted apple fruit skins. New Phytol 212:977–991. doi:10.1111/nph.14170

Liu RH (2013) Health-promoting components of fruits and vegetables in the diet. Adv Nutr Int Rev J 4:384S–392S. doi:10.3945/an.112.003517

Luévano-Contreras C, Gómez-Ojeda A, Macías-Cervantes MH, Garay-Sevilla ME (2017) Dietary advanced glycation end products and cardiometabolic risk. Curr Diab Rep 17:63. doi:10.1007/s11892-017-0891-2

Mikulič Petkovšek M, Usenik V, Štampar F (2003) The role of chlorogenic acid in the resistance of apples to apple scab (*Venturia inaequalis* (Cooke) G. Wind. Aderh.). Res Rep Biotech Fac Univ Ljubl 81:233–242

Najafian M, Jahromi MZ, Nowroznejhad MJ, Khajeaian P, Kargar MM, Sadeghi M, Arasteh A (2012) Phloridzin reduces blood glucose levels and improves lipids metabolism in streptozotocin-induced diabetic rats. Mol Biol Rep 39:5299–5306

R Core Team (2016) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Stanger MC, Steffens CA, Soethe C, Moreira MA, do Amarante CVT (2017) Phenolic content and antioxidant activity during the development of 'Brookfield' and 'Mishima' apples. J Agric Food Chem 65:3453–3459. doi:10.1021/acs.jafc.6b04695

USDA (2014) National genetic resources program. Germplasm resources information network (GRIN). USDA, ARS Natl. Germplasm Resources Laboratory, Beltsville, MD. http://www.ars-grin.gov/npgs/index.html. Accessed 5 May 2017

Volz RK, McGhie TK (2011) Genetic variability in apple fruit polyphenol composition in *Malus*× *domestica* and *Malus sieversii* germplasm grown in New Zealand. J Agric Food Chem 59:11509–11521. doi:10.1021/jf202680h

Wang L, Li J, Gao J, Feng X, Shi Z, Gao F, Xu X, Yang L (2014) Inhibitory effect of chlorogenic acid on fruit russeting in 'Golden Delicious' apple. Sci Hortic 178:14–22

Yahyaa M, Ali S, Davidovich-Rikanati R, Ibdah M, Shachtier A, Eyal Y, Lewinsohn E, Ibdah M (2017) Characterization of three chalcone synthase-like genes from apple (*Malus* x *domestica* Borkh.). Phytochem 140:125–133. doi:10.1016/j.phytochem.2017.04.022

CHAPTER 5

DIRECTIONS FOR DIHYDROCHALCONE STUDIES IN APPLE

Throughout our studies of dihydrochalcones (DHCs) in *Malus*, we uncovered new and exciting questions on these compounds and their significance to *Malus* evolution and taxonomy. Additionally, we identified and developed promising genetic resources to address these questions. The aim of this chapter is to provide some conclusions and insight into potential studies to further our understanding of DHCs in apple and its wild relatives.

Although all *Malus* species accumulate DHCs, there is significant variation in DHC content among and within species and tissue types. From our germplasm surveys, leaf DHC content ranged from 25.6 to 113.7 mg/g. Within the cultivated apple, leaf DHC content ranged from 34.1 to 66.8 mg/g and from 24.3 to 779.1 μg/g in the peels of non-russeted fruit. DHC content in leaves and peel in germplasm studies had high broad-sense heritability estimates, 0.76 and 0.63, respectively. Key genes related to phloridzin synthesis have been described. Zhou et al. (2017) screened the 'Golden Delicious' genome for glycosyltransferases related to phloridzin synthesis and measured differences in gene expression across species and tissue types. Yahyaa et al. (2017) measured phloretin content and changes in expression of chalcone synthase genes in 'Golden Delicious' vegetative tissues and fruit. From these studies, it is clear that additional genes are also associated with DHC content variation.

Though the 'Golden Delicious' genome has been leveraged for many genetic studies, genomic variation among *Malus* species (Höfer and Meister 2010) could be explored with whole-genome re-sequencing (Duan et al. 2017), particularly for species with exceptional DHC accumulation, or with trilobatin or sieboldin. Sieboldin and trilobatin are rare dihydrochalcones, found in a few species and hybrids. These DHCs were useful in distinguishing interspecific

hybrids in germplasm collections, in conjunction with genetic and morphological data. Based on their nutritional properties, they could be used in breeding programs to diversify the nutritional profile of modern apples. However, many of the genetic resources for these DHCs have poor fruit quality, such as *Malus floribunda* 821, but may contain other valuable traits, such as resistance to apple scab (Brown 2012). Through genetic mapping, we identified loci on LGs 7 and 8 associated with the presence of trilobatin and sieboldin, respectively, which could be used in genomics-assisted breeding. While an initial search near these loci did not identify any candidate genes associated with DHC synthesis, this is an area for future studies. Zhou et al. (2017) identified a glycosyltransferase on LG 8 associated with DHC content in sieboldin and trilobatin-producing cultivar, *Malus* spp. 'Adams', but this was about 3 Mb from the locus described in chapter 3. Although a three-gene model was proposed, segregation of phloridzin was not observed in genotyped population 13427, so no locus was proposed for phloridzin. The most informative population would be one in which both parents are heterozygous for phloridzin, sieboldin, and trilobatin. As genetic control for sieboldin and trilobatin is elucidated, their physiological significance could be investigated (Dare et al. 2017). Biochemical analysis is required to confirm that sieboldin originates from 4′-glycosylation of hydroxyphloretin.

While, the DHC composition of phloridzin, sieboldin, and trilobatin follow single gene inheritance, there is significant variation in individual content and proportions of these DHC in the germplasm and F_1 populations. In many accessions, phloridzin is the prominent form in accessions of profile SPT, however, there are cases of sieboldin or trilobatin-dominant accessions, or all three being in equal proportions. Though the biosynthetic pathway of DHCs has been proposed, it is unclear how they are regulated, or what affects their equilibrium. There is a strong effect of DHC profile (P, PT, or SPT) on individual DHCs. Among the five F_1

populations used in our studies (chapter 3), phloridzin content was higher in profile P, than in PT or SPT, but there was no difference between phloridzin content in profile PT or SPT. Similarly, trilobatin content was higher in profile PT than in SPT. The locus on LG 7 was associated with phloridzin content variation, and the locus on LG8 was associated with trilobatin. However, DHCs were positively correlated in some of our F_1 populations, suggesting they are not competitors.

Due to the unique properties of phloridzin, and its presence in apple, it is a target for selective breeding in new apple cultivars. Phloridzin is distributed disproportionately in vegetative tissues relative to fruit. Within the fruit, phloridzin is more abundant in the peel. We identified genetic resources with high peel phloridzin content, relative to the popular cultivars 'Golden Delicious' (104.3 μg/g) and 'Honeycrisp' (38.9 μg/g). However, many sources of high peel phloridzin also had poor fruit quality, such as 'Pigeonnet Rouge' (779.1 μg/g). Cultivar 'Arkansas Black' had high phloridzin content (753.2 μg/g), as well as high flavonol content (1279.6 μg/g) and moderate fruit quality. We observed that that russeting in apple fruit is associated with increased phloridzin content and decreased flavonol content. Although russeted apples possess unique qualities, and are valued as heirloom cultivars, consumer and industry preference for waxy fruit would make them challenging to market. However, they provide a unique model in which to study DHC regulation. While, suberin pathways, upregulated during fruit russeting, are interconnected to flavonoid synthesis, phloridzin is the sole flavonoid to be upregulated. Legay et al. (2016, 2017) identified key genes differentially expressed in waxy and russeted portions of apple peels, which may be regulators of phloridzin. A transcriptomic approach, comparing contrasting russet phenotypes, may yield insights into the association of russeting and phloridzin accumulation.

Apples are a high value crop, with world-wide distribution. Breeding will continue to create new cultivars with superior fruit quality, but also must meet the challenges of evolving consumer preferences, horticultural practices, climate change, and should include the selection for new traits, such as nutritional quality. Utilization of diverse genetic resources will aid this effort. Advances in phenotyping, genotyping, and genomic information for apple are leading to increased understanding of *Malus* species and apple domestication, both of which will facilitate the use of diverse genetic resources in apple improvement (Duan et al. 2017, Kumar et al. 2012).

REFERENCES

Brown S (2012) Apple. In: Badenes ML, Byrne DH (eds) Fruit Breeding. Springer, New York, pp 329–367

Dare AP, Yauk Y-K, Tomes S, McGhie TK, Rebstock RS, Cooney JM, Atkinson RG (2017) Silencing a phloretin-specific glycosyltransferase perturbs both general phenylpropanoid biosynthesis and plant development. Plant J 91:237–250. doi: 10.1111/tpj.13559

Duan N, Bai Y, Sun H, et al (2017) Genome re-sequencing reveals the history of apple and supports a two-stage model for fruit enlargement. Nat Commun 8. doi: 10.1038/s41467-017-00336-7

Höfer M, Meister A (2010) Genome size variation in *Malus* species. J Botany 2010: e480873 . doi: 10.1155/2010/480873

Kumar S, Chagné D, Bink MC, Volz, RK, Whitworth, C, Carlisle C. (2012). Genomic selection for fruit quality traits in apple (*Malus*× *domestica* Borkh.). PLoS ONE, 7(5), e36674.

Legay S, Cocco E, André CM, Guignard C, Hausman J-F, Guerriero G (2017) Differential lipid composition and gene expression in the semi-russeted 'Cox Orange Pippin' apple variety. Front Plant Sci 8. doi: 10.3389/fpls.2017.01656

Legay S, Guerriero G, André C, Guignard C, Cocco E, Charton S, Boutry M, Rowland O, Hausman J-F (2016) *MdMyb93* is a regulator of suberin deposition in russeted apple fruit skins. New Phytol 212:977–991. doi: 10.1111/nph.14170

Yahyaa M, Ali S, Davidovich-Rikanati R, et al (2017) Characterization of three chalcone synthase-like genes from apple (*Malus* x *domestica* Borkh.). Phytochem 140:125–133. doi: 10.1016/j.phytochem.2017.04.022

Zhou K, Hu L, Li P, et al (2017) Genome-wide identification of glycosyltransferases converting phloretin to phloridzin in *Malus* species. Plant Science 265:131–145. doi: 10.1016/j.plantsci.2017.10.003