© 2022

Yurah Kim

ALL RIGHTS RESERVED
SEGREGATION OF CHLOROGENIC ACIDS AND FLAVONOL GLYCOSIDES IN INTERSPECIFIC 
(VACCINIUM CORYMBOSUM X V. DARROWII) DIPLOID F2 BLUEBERRY POPULATION

By

YURAH KIM

A thesis submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Master of Science

Graduate Program in Plant Biology

Written under the direction of

Nicholi Vorsa

And approved by

__________________________________

__________________________________

__________________________________

New Brunswick, New Jersey

JANUARY 2022
ABSTRACT OF THE THESIS

Segregation of Chlorogenic Acids and Flavonol Glycosides in Interspecific (Vaccinium corymbosum x V. darrowii) Diploid F2 Blueberry Population

By YURAH KIM

Thesis Director:
Nicholi Vorsa

Blueberry (Vaccinium spp.), belonging to the family Ericaceae and subsection Cyanococcus, is one of the richest sources of antioxidant phytonutrients among fresh fruits and receives much attention for its phenolic compounds associated with human health-related benefits.

Chlorogenic acids, as predominant blueberry phenolics together with anthocyanins, also likely account for a large proportion of antioxidant activity. Flavonols, as one of six major flavonoid subclasses, are usually found in plants bound to sugars as O-glycosides and are the most abundant and widely distributed in nature. They have important roles as developmental regulators and/or signaling molecules in plants and have been associated with the prevention of various oxidative diseases.

The aim of this thesis is to investigate the variation of phenolic compound and organic acid content in two divergent blueberry species, Vaccinium corymbosum and V. darrowii, and their
F1 hybrids, and to evaluate segregation in the interspecific (*V. corymbosum* x *V. darrowii*) diploid blueberry F2 population for chlorogenic acid and flavonol glycoside phenotypic traits. The objectives of this thesis were to: (1) identify and quantify phenolic compounds and organic acids in blueberry using high-performance liquid chromatography, (2) evaluate the chlorogenic acid and flavonol glycoside profiles among six blueberry genotypes, (3) analyze the segregation pattern of phenolic phenotypes in a biparental F2 population derived from a cross between *V. corymbosum* and *V. darrowii*, and (4) classify the F2 individuals according to their relative composition of chlorogenic acids and flavonol glycosides.

The thesis will first focus on the analysis of different phenolic compounds and organic acids in clones of two blueberry species. Four phenolic acids and ten flavonol glycosides were identified using LC-MS-MS, while three organic acids were identified using HPLC. Genotype significantly affects the variation of average concentrations and profiles of phenolic compounds in blueberry species. *V. corymbosum* fruits presented mainly acetyl-caffeoylquinic acids in their phenolic acid composition, while *V. darrowii* fruits exhibited higher levels of 3-caffeoylquinic acid, and the interspecific F1 hybrid fruits showed intermediate levels of both compounds. Except for *V. corymbosum*, in which syringetin was present in the highest quantity, quercetin occurred as the most abundant aglycone. Galactosides and glucosides were the most abundant conjugate sugars in *V. corymbosum* and *V. darrowii* blueberries, whereas rhamnosides and galactosides were the most abundant conjugate sugars in F1 hybrid fruits. It was also found that genotype significantly affects the level of organic acids in blueberry species. Citric acid was present in low quantities in *V. darrowii*, while quinic acid was present in low quantities in *V. corymbosum*, and both were present in intermediate levels in the F1 hybrids. These observations suggest that both phenolic compound and organic acid biosynthesis are regulated differently in species.
Phenolic compounds in the segregating diploid F2 population showed largely continuous distribution for each compound value. All compounds segregated and most were approximately normally distributed, with one acetyl-caffeoylquinic acid isomer and syringetin-3-galactoside bimodally distributed. Many phenolic compounds also exhibited transgressive segregation. Gene models of blueberry phenolic compounds were determined in the F2 population using chi-square goodness of fit tests. For the studied compounds, results were found statistically non-significant, where the null hypothesis was Mendelian distribution in a single locus model. A principal component analysis showed that F2 individuals can be distinguished based on their phenolic compound profile. These findings suggest that most of the phenolic compounds examined in this study are quantitative in nature, likely involving multiple loci, and might be controlled by a single dominant gene.
ACKNOWLEDGEMENT

First and foremost, I would like to express my sincerest gratitude to my advisor, Dr. Nicholi Vorsa, for his continued guidance and direction during my graduate studies, his unending patience and understanding despite my constant late deadlines, and his genuine desire to help me produce the best research possible. I would also like to show my deepest appreciation to my graduate committee members, Dr. Daniel Ward and Dr. Thomas Gianfagna, for their generosity with their time and helpful advice. In addition, I would like to thank Dr. Ajay Singh, who guided me through hands-on instrument training.

Furthermore, I very much appreciate the valuable assistance provided by the members of the Philip E. Marucci Center for Blueberry and Cranberry Research. I also wish to extend my thanks to my past and current lab members, Stephanie Fong, Yifei Wang, Mithu Chatterjee, Ira Herniter, and Konrad Wysocki, for offering their useful insight and lending a listening ear.

Finally, to my friends and family, I am not only grateful for their unwavering support, but for their encouragement that kept me going through those stressful days and nights.
# TABLE OF CONTENTS

ABSTRACT OF THE THESIS .................................................................................................................................................. ii

ACKNOWLEDGEMENT .......................................................................................................................................................... v

TABLE OF CONTENTS ......................................................................................................................................................... vi

LIST OF TABLES .................................................................................................................................................................. xi

LIST OF ILLUSTRATIONS .................................................................................................................................................. xii

CHAPTER 1 ........................................................................................................................................................................ 1

Literature Review ................................................................................................................................................................. 1

1.1 Blueberry (*Vaccinium* spp.) ........................................................................................................................................ 1

1.1.1 Introduction ............................................................................................................................................................... 1

1.1.2 History of cultivation ................................................................................................................................................. 3

1.1.3 Major world crop ......................................................................................................................................................... 4

1.1.4 Nutrition ................................................................................................................................................................... 6

1.1.5 Health benefits ......................................................................................................................................................... 6

1.2 Chlorogenic acids ........................................................................................................................................................... 9

1.2.1 Introduction ............................................................................................................................................................... 9

1.2.2 Structure .................................................................................................................................................................. 11

1.2.3 Biosynthesis ............................................................................................................................................................. 11

1.2.4 Biological functions of chlorogenic acids in plants ............................................................................................. 13
CHAPTER 2

Variation of phenolic compounds and organic acids in two diploid blueberry species, Vaccinium corymbosum and V. darrowii, and their F1 hybrids, using high-performance liquid chromatography
CHAPTER 3 ..................................................................................................................................... 98

Phenotypic evaluation of a segregating (*Vaccinium corymbosum* x *V. darrowii*) diploid F2
blueberry population for phenolic compounds............................................................................. 98

3.1 Introduction .................................................................................................................................. 98

3.2 Materials and methods .................................................................................................................. 100

3.2.1 Plant materials ...................................................................................................................... 100

3.2.2 Chemicals and reagents ..................................................................................................... 101

3.2.3 Extraction of blueberry phenolic compounds .................................................................... 101

3.2.4 HPLC apparatus and conditions ......................................................................................... 102

3.2.5 Compound identification with MS spectrometry of samples ............................................ 102

3.2.6 Chromatograms were used for characterization and quantification of phenolic compounds .................................................................................................................................. 103

3.2.7 Statistical analysis .............................................................................................................. 103

3.3 Results ..................................................................................................................................... 104

3.3.1 Segregation of phenolic traits in the F2 population ............................................................. 104

3.3.2 Principal component analysis (PCA) .................................................................................. 105

3.4 Discussion ................................................................................................................................. 106

3.4.1 Segregation of phenolic traits in the F2 population ............................................................. 106

3.4.2 Principal component analysis (PCA) .................................................................................. 107

3.5 Conclusions ............................................................................................................................... 108

3.6 References ............................................................................................................................... 109
3.7 Figures and tables ................................................................. 111
**LIST OF TABLES**

**Chapter 2:**

Table 2.1 List of abbreviations for all studied compounds ................................................................. 92

Table 2.2 Summary of chemical analysis methods for chlorogenic acids, flavonol glycosides, and organic acids .................................................................................................................................. 92

Table 2.3 Liquid chromatography retention time, [M-H]-, and fragmentation time of blueberry phenolic compounds ...................................................................................................................... 93

Table 2.4 Calibration curves used in the quantification of phenolic compounds and organic acids ....................................................................................................................................................... 93

Table 2.5 Mean and standard deviation values of all phenolic compounds among genotypes (mg/g FW) ...................................................................................................................................... 94

Table 2.6 Pearson's correlation matrix with coefficients (r) and p-values of phenolic compounds and average fruit weight in six wild diploid blueberry genotypes (N = 208) ......................................................... 95

Table 2.7 Kruskal-Wallis rank sum test results of total phenolic contents ................................................. 97

**Chapter 3:**

Table 3.1 Observed and expected segregation ratios and chi-square analysis for the F2 population segregating for fruit phenolic compound concentration .................................................. 116

Table 3.2 Phenotypic characteristics of the F2 parent species (V. corymbosum and V. darrowii) for the fruit phenolic compound traits to be analyzed (mg/g FW) .................................................. 117

Table 3.3 Eigenanalysis of the correlation matrix for phenolic compounds found in the F2 population ........................................................................................................................................................................... 118
LIST OF ILLUSTRATIONS

Chapter 1:

Figure 1.1 Chemical structure of 5-caffeoylquinic acid and 3-caffeoylquinic acid ................. 61
Figure 1.2 Basic skeleton structures of flavonoid compound and flavonol aglycones ............. 61

Chapter 2:

Figure 2.1 Pedigree chart of blueberry plants ........................................................................ 86
Figure 2.2 HPLC chromatogram of blueberry phenolic compounds .................................... 86
Figure 2.3 Bar chart of linear mixed effects model estimated average concentrations of phenolic compounds for six wild diploid blueberry genotypes (mg/g FW) ......................................................... 87
Figure 2.4 Effect of harvest year on concentrations of total phenolic compounds (mg/g FW) .... 89
Figure 2.5 Average composition of different compound groups ............................................. 90
Figure 2.6 Effect of localization in epidermis or flesh of blueberry fruit on concentrations of total phenolic compounds (mg/g FW) .................................................................................................. 91
Figure 2.7 Average composition of different organic acids in genotypes ............................... 91

Chapter 3:

Figure 3.1 Frequency distribution histograms of phenolic compounds in the F2 population (mg/g FW) ........................................................................................................................................ 111
Figure 3.2 Principal component analysis of phenolic compounds in the F2 population .......... 114
CHAPTER 1

Literature Review

1.1 Blueberry (*Vaccinium* spp.)

1.1.1 Introduction

Blueberries (*Vaccinium* spp.) are a diverse group of perennial flowering shrubs that belong in the family Ericaceae (Syn. Health) and produce clusters of small blue or purple fruit (Vander Kloet, 1988). There are about 450 known species in this family geographically distributed in the Northern Hemisphere and in the mountains of tropical Asia and Central and South America (Luby et al., 1991). Approximately 35% of the species are native to America, and of those, only the 26 species in subsection *Cyanococcus* that occur naturally in eastern and northcentral North America are referred to as true blueberries (Hancock & Draper, 1989; Retamales & Hancock, 2012; Vander Kloet, 1988). The genus *Vaccinium* contains other agriculturally important subsections such as *Oxycoccus* (*V. macrocarpon* Ait.; American cranberry), *Vitis-Idaea* (*V. vitis-idaea*; lingonberry), and *Myrtillus* (*V. myrtillus* L.; bilberry, whortleberry) (Song & Hancock, 2011).

Cultivars are categorized into three main varieties: highbush, lowbush, and rabbiteye. Highbush blueberries can be further divided into northern, southern, and half-high types, depending on their chilling requirements and winter hardiness (Retamales & Hancock, 2012). Most commercial production comes from species in the subsection *Cyanococcus*, which includes the cultivars *Vaccinium corymbosum* L. (highbush), *V. ashei* Reade (rabbiteye; syn. *V. virgatum* Ait.), and native stands of *V. angustifolium* Ait. (lowbush) and *V. myrtilloides* Michx. (lowbush) (Retamales & Hancock, 2012; Song & Hancock, 2011). Other species are utilized as sources of
desirable traits in cultivar development, such as *V. darrowii* Camp (lowbush), *V. constablaei* Gray (highbush), and *V. elliotii* (highbush) (Retamales & Hancock, 2012).

Blueberries naturally occur as diploids, tetraploids, and hexaploids (Song & Hancock, 2011). This polyploidy, plus overlapping morphologies, continuous introgression through hybridization, and a general lack of chromosome differentiation, makes the taxonomy of *Vaccinium* subsections difficult to resolve (Retamales & Hancock, 2012). Furthermore, blueberries demonstrate some degree of self-sterility. Meader and Darrow observed a reduction in seed viability and size in self-pollinations of highbush blueberries compared to those of cross pollinations (Meader & Darrow, 1947). In addition, a marked reduction in fruit set was demonstrated in self-pollinations of native tetraploid and hexaploid *V. corymbosum*, as well as a decrease in plump seeds per berry by two-thirds or more at all ploidy levels (Vander Kloet & Lyrene, 2011). Similarly, a reduction in fruit set was also noted among self-pollinations of 19 wild rabbiteye clones, albeit with substantial variability among clones (Garvey & Lyrene, 1987). Consequently, self-pollinations are rarely used in breeding due to reduced seed set and germination, and blueberries are most often asexually propagated through cuttings and tissue culture (Retamales & Hancock, 2012).

Prior to 1916, all blueberries were harvested from the wild (Song & Hancock, 2011). Indeed, many wild edible *Vaccinium* species have been harvested for thousands of years by indigenous peoples (Kalt & Dufour, 1997; Moerman, 1998). Native Americans not only used *Vaccinium* fruit as a major food source but the plant parts for different medicinal preparations (Hummer, 2013; Vander Kloet, 1988). In western and eastern North America, Native Americans intentionally burned native stands of blueberries to encourage their spread and productivity (Gough, 1997;
Song & Hancock, 2011). Domestication of highbush and rabbiteye blueberries occurred at the end of the 19th century, with plants initially dug from the wild and transplanted into New England and Florida fields (Song & Hancock, 2011). It was not until the mid-1990s, however, that blueberry production truly launched in the U.S. Studies on antioxidant capacity conducted around that time demonstrated that blueberries had the highest antioxidant activity among over 30 fruits and vegetables tested (Howell et al., 2001; Prior et al., 1998; Wang et al., 1996). These findings popularized blueberries as a “superfood” and brought attention to the growing body of research on the health potential of this fruit.

1.1.2 History of cultivation

Tetraploid northern highbush blueberry breeding began in the early 1990s in New Jersey with the work of Frederick Coville of the U.S. Department of Agriculture (USDA). Coville was the first to establish the fundamental requirements of blueberry cultivation. He determined that blueberries require moist but well-drained acidic soil, had low nutrient requirements, and need low-temperature winter chilling periods (Coville, 1916; Hancock et al., 2003; Mainland, 2012; Song & Hancock, 2011). In addition, he developed propagation procedures by cuttings, grafting, and budding and discovered that some genotypes are self-sterile (Coville, 1927; Mainland, 2012; Retamales & Hancock, 2012).

Coville made his first wild selection for use in crosses in 1908 (Retamales & Hancock, 2012), and in 1910, published his findings in *Experiments in Blueberry Culture*, USDA Bulletin 193 (Mainland, 2012). In 1911, he began a collaborative breeding project with private cranberry grower Elizabeth C. White of the J.J. White Company in Whitesbog, NJ (Ehlenfeldt, 2009). Over a span of 26 years, they collected elite wild clones of *V. corymbosum* and *V. angustifolium* and used them
in the breeding of improved types. Their first hybrid ‘Pioneer’ was released in 1920 (Mainland, 2012; Retamales & Hancock, 2012).

George Darrow made important contributions on interfertility and phylogeny on native Vaccinium species in cooperation with taxonomist W.H. Camp (Hancock, 2006b). He also initiated rabbiteye blueberry breeding. Arlen Draper focused on incorporating the genes of most wild Vaccinium species into the cultivated highbush background (Hancock, 2006a; Hancock & Galletta, 1995). In addition, he released an outstanding number of northern and southern highbush cultivars with improved fruit quality and higher productivity (Hancock & Galletta, 1995; Retamales & Hancock, 2012).

Ralph Sharpe began work on the development of low-chill highbush varieties (Retamales & Hancock, 2012; Sharpe & Darrow, 1959). His southern highbush blueberries combined the fruit quality and productivity of northern highbush V. corymbosum with the low chilling requirement of V. darrowii, an evergreen species native to Florida. Stanley Johnson developed varieties with improved cold tolerance (Retamales & Hancock, 2012). He hybridized northern highbush V. corymbosum with lowbush V. angustifolium, and out of his work came the half-high cultivar ‘Northland,’ which possessed higher yields and larger fruit than lowbush blueberries but with a low enough stature to be protected by snow in areas of extreme cold (Finn et al., 1990; Hancock et al., 2008).

1.1.3 Major world crop

Highbush blueberry culture gradually spread across the U.S. from 1910 onwards (Retamales & Hancock, 2012). Blueberries are now grown in 37 states in the U.S. and six provinces in Canada
and are a major commercial crop in North America (Strik, 2005; Strik & Yarborough, 2005), with most production coming from highbush and lowbush varieties (Retamales & Hancock, 2012). The largest acreages of northern highbush blueberries are in Michigan, New Jersey, North Carolina, and Oregon, whereas the greatest acreages of southern highbush blueberries are in Florida, California, and Georgia (Song & Hancock, 2011). For lowbush varieties, commercial cultivation is centered in Maine, Quebec, New Brunswick, and Nova Scotia (Hancox et al., 2008; Strik, 2005). Blueberries are also cultivated internationally in Australia, Chile, Argentina, New Zealand, and parts of Asia and Europe where highbush and rabbiteye varieties make up most commercial production (Retamales & Hancock, 2012; Strik, 2005; Strik & Yarborough, 2005).

In the U.S., cultivated blueberries are the second most important berry crop and the single most important commercial crop in the Vaccinium genus (Kramer et al., 2020; Song & Hancock, 2011). 324,000 tons were produced in 2020 on over 91,000 acres that included 18,200 acres in Washington, 17,200 acres in Michigan, 16,000 acres in Georgia, 13,500 acres in Oregon, 8,400 acres in New Jersey, 7,300 acres in North Carolina, 6,400 acres in California, and 4,400 acres in Florida (USDA, 2021). Production of 319,000 tons of cultivated blueberries were utilized at a value of 904 million dollars, and 24,000 tons of wild blueberries were utilized at a value of 29 million dollars (USDA, 2021). Furthermore, annual fresh blueberry per capita consumption has grown over 510%, climbing in value from 460 million dollars in 2010 to 760 million dollars in 2019 (Kramer et al., 2020). In 2020, approximately 55% of harvested blueberries were intended for the fresh market, while the remaining 45% were processed into various pie fillings, yogurts, ice cream, prepared muffin and pancake mixes, syrups, jams, and preserves (Eck, 1988; Song & Hancock, 2011; USDA, 2021).
1.1.4 Nutrition

Blueberries are natural sources of food, beverage, and nutraceutical ingredients due to their rich nutritional and bioactive compounds. The nutrient levels of blueberries are affected by many factors such as variety, growing conditions, maturity at harvest, and processing carried out during post-harvest storage (Häkkinen & Törrönen, 2000; Song & Hancock, 2011). In general, blueberries are composed of 84% water, 0.7% protein, 0.3% fat, 2.4% dietary fiber, 0.5% ash, 14.5% carbohydrates, and 11% sugars (Hancock et al., 2003; USDA, 2019). Blueberry fruits are also intermediate to low sources of vitamin A (81-100 IU/100 g), vitamin C (7-16 mg/100 g), amino acids (most prominently arginine), minerals, potassium, phosphorus, and calcium (Eck, 1988; Hancock et al., 2003; Padmanabhan et al., 2016; Prior et al., 1998). In addition, ripe blueberries contain 1-2% organic acids (primarily citric acid) (Padmanabhan et al., 2016; Song & Hancock, 2011), with titratable acids measuring 0.2-1.4% (Zhang et al., 2020a).

1.1.5 Health benefits

Blueberries have received considerable attention for their abundant phenolic compounds, strong antioxidant capacity, and potential health benefits (Cao et al., 1996; Giovanelli & Buratti, 2009; Howard et al., 2003; Sellappan et al., 2002; Sinelli et al., 2008; Vrhovsek et al., 2012; Wang & Jiao, 2000). The high phenolic content of blueberries is thought to be associated with their health-promoting properties (Bujor et al., 2019). In addition, blueberries are known as one of the richest sources of antioxidant phytonutrients among fresh fruits (Song & Hancock, 2011), with a total antioxidant capacity ranging from 13.9-45.9 μmol Trolox equivalents per gram of fresh fruit (Connor et al., 2002b, 2002a; Ehlenfeldt & Prior, 2001; Prior et al., 1998; Zheng & Wang, 2003). Anthocyanins, phenolic compounds responsible for the pigmentation of blueberries, have been shown to be among the most powerful of antioxidants (Ehlenfeldt &
Prior, 2001). Accumulation of anthocyanins occurs during fruit development and ripening (Brown & Shipley, 2011), and total anthocyanins can range from 85-270 mg per 100 g of blueberries (Song & Hancock, 2011). Anthocyanins have been postulated to protect against cancer, diabetes, heart and vascular diseases, and neurodegenerative diseases in humans (Ehlenfeldt & Prior, 2001; Seeram et al., 2006).

Chlorogenic acids are also predominant polyphenolic compounds in blueberries and likely account for a large proportion of their antioxidant activity (Drewnowski & Gomez-Carneros, 2000; Song & Hancock, 2011). In highbush and lowbush varieties, chlorogenic acid is the major phenolic acid and constitutes 10-16% of total acids in the fruit (Kalt & McDonald, 1996; Kang et al., 2015; Markakis et al., 1963). It is present in concentrations of 98-208 mg per 100 g FW (fresh weight) V. corymbosum cultivars and 59-110 mg per 100 g FW V. angustifolium cultivars (Gao & Mazza, 1994; Markakis et al., 1963; Schuster & Herrmann, 1985). Other phenolic acids such as caffeic acid, p-coumaric acid, and ferulic acid are present in less than 1% concentrations in blueberries (Kalt & McDonald, 1996; Schuster & Herrmann, 1985), as well as minor amounts of free gallic acid and its derivatives (Herrmann, 1989).

Blueberries also possess a diverse profile of other flavonoids. Diets replete in flavonoid-rich foods are associated with increased longevity and decreased incidence of cardiovascular diseases (Formica & Regelson, 1995). The most frequently studied flavonoid is quercetin, which belongs to the flavonol subclass of flavonoid compounds. Quercetin has been shown to have biological properties consistent with its sparing effect on the cardiovascular system (Formica & Regelson, 1995; Shabrova et al., 2011). It has been exhibited to modify antiprostanoid and anti-inflammatory responses, prevent atherosclerotic plaque formation, and have antithrombic,
antihypertensive, and antiarrhythmic effects (Formica & Regelson, 1995). One of the most concentrated sources of dietary flavonols are cultivated blueberries, which can provide an average of 196 mg/kg of a complex mixture of up to 23 different glycosides (Vrhovsek et al., 2012). While blueberries carry the same predominant aglycones and glycosides, relative proportions vary between different varieties and genotypes (Ballington et al., 1988; Cho et al., 2005; Vrhovsek et al., 2012). Quercetin can be present from 1.7-2.9 mg per 100 g FW and myricetin can be present from 2.3-2.6 mg per 100 g FW (Bilyk & Sapers, 1986; Häkkinen et al., 1999). The dominant sugar conjugates are galactosides and glucosides, with percentages of glycosides dependent on the blueberry variety (Vrhovsek et al., 2012).

Recognition of the human health benefits of blueberries is supported by a growing body of observational evidence, along with research from human clinical studies and from animal and in vitro models. Regular, moderate intake of blueberries is correlated with reduced biomarkers and risk of a multitude of non-communicable human diseases (Kalt et al., 2020). Vendrame et al. (2013) demonstrated that a wild blueberry-enriched diet improves the proinflammatory status associated with metabolic syndrome in obese Zucker rats. Furthermore, chronic inflammation is not only a critical component of metabolic syndrome but can also lead to increased risk of type 2 diabetes and cardiovascular disease, which suggests blueberry intake might have cardiometabolic benefits (Vendrame et al., 2013). In individuals with increased risk for type 2 diabetes, short term consumption of wild blueberry juice was shown to promote cardioprotective effects by improving systolic blood pressure (Stote et al., 2017).

In addition, blueberry consumption may indirectly contribute to host health via modulating the gut microbiome (Kalt et al., 2020). Jiao et al. (2019) reported that blueberry polyphenols
enhanced the growth of specific beneficial bacteria, such as Proteobacteria, *Bifidobacterium*, and *Helicobacter*, in the gastrointestinal tract of C57BL/6J mice. Thus, the evidence reported by Jiao et al. (2019) indicates that blueberry polyphenols may, as a potential prebiotic agent, influence the gut microbiota to positively affect high-fat diet-induced obesity in C57BL/6J mice. Also, another study in C57BL/6J mice demonstrated that blueberry feeding modulated cecal microbial communities, decreasing α-diversity, altering β-diversity, and lowering the *Firmicutes:Bacteroidetes* ratio (Wankhade et al., 2019). However, it should be noted that sexually dimorphic responses to blueberry supplementation were observed, such as a stronger association with microbes only in male mice, which were associated with predicted metabolic pathways (Wankhade et al., 2019).

### 1.2 Chlorogenic acids

#### 1.2.1 Introduction

The term “chlorogenic acid (CGA)” classically refers to a group of ester compounds formed between *trans*-cinnamic acid and quinic acid (Lu et al., 2020; Upadhyay & Mohan Rao, 2013). CGAs are divided into groups depending on the identity, number, and position of their acyl moiety (Clifford, 2000). The most common groups are *p*-coumaroylquinic acids, caffeoylquinic acids, feruloylquinic acids, and dicaffeoylquinic acids (Lallemand et al., 2012). There are also several isomeric forms in each group (Clifford, 2000; Lu et al., 2020). The most abundant isomer present in dietary and plant sources is 5-*O*-caffeoylquinic acid (5-CQA) (Figure 1.1A) (Tajik et al., 2017), and it is the most extensively studied due to its earlier commercial availability than other isomers (Naveed et al., 2018). Currently, the term “chlorogenic acid (CGA)” refers to 5-CQA (Lu et al., 2020). It should be noted, however, that problems exist concerning the nomenclature of isomers 5-CQA and 3-CQA (Figure 1.1B). In 1976, the International Union of Pure and Applied
Chemistry (IUPAC) reversed the order of the numbering of atoms on the quinic acid ring (Panico et al., 1993). Consequently, the previously identified 3-CQA was renamed as 5-CQA (Panico et al., 1993), and the current 3-CQA refers to neochlorogenic acid in accordance with the new numbering system (Lu et al., 2020). Further contributing to the problem, many chemical suppliers and researchers still use pre-IUPAC nomenclature (Panico et al., 1993).

CGAs are one of the most available phenolic acids in foods, such as coffee and tea (Moeenfard et al., 2014; Naveed et al., 2018). They are found in nearly all plant species and are widespread dietary components present in many plant-based foods, including berries, apples, eggplants, tomatoes, spearmint, chicory, sunflower seeds, potatoes, and sweet potato leaves (Alcázar Magaña et al., 2021; Clifford et al., 2017; Kuhnert et al., 2014; Kumar et al., 2020; Lu et al., 2020). They are also found in some commonly used medicinal herbs such as chrysanthemum flowers, artemisia leaves, burdock roots, dandelion roots, and echinacea roots (Clifford et al., 2020; Upadhyay & Mohan Rao, 2013). CGAs are the major hydroxycinnamic acids present in blueberries (Kader et al., 1996), with 5-CQA being the major component (Schuster & Herrmann, 1985). Hydroxycinnamic acids, derived from cinnamic acid, are one of two main subgroups of phenolic acids (Clifford, 2000). There are four most common hydroxycinnamic acids (p-coumaric, caffeic, ferulic, and sinapic acids) (Kumar & Goel, 2019), and among them, caffeic acid is the predominant acid found in blueberries (Murkovic, 2003). Typically, the free forms of hydroxycinnamic acids are very rare in fruits (Murkovic, 2003). They are more widely present in bound forms, and the most abundant soluble bound derivative is CGA (Kumar & Goel, 2019; Murkovic, 2003).
1.2.2 Structure

Hydroxycinnamic acids, derived from phenylalanine and tyrosine (Erukainure et al., 2018), are the hydroxyl metabolites of cinnamic acid (Martinez et al., 2017). They possess a common structure of a 9-carbon skeleton with a side chain double bond and are predominantly in trans configuration (Clifford et al., 2017; Martinez et al., 2017). CGAs are a large group of phenolic compounds synthesized by the esterification of C6-C3-trans-hydroxycinnamic acid with 1L-(-)-quinic acid (Clifford et al., 2017). CGAs have a common structure of quinic acid (tetrahydroxy-cyclohexane carboxylic acid) acylated with one to four caffeic acid (3,4-dihydroxycinnamic acid) moieties (Alcázar Magaña et al., 2021; Oestreich-Janzen, 2019). 15 structure combinations are possible and include mono-, di-, tri-, and tetra-caffeoylquinic acids (Alcázar Magaña et al., 2021). In addition, isomers and epimers in the cyclohexane part and substitutions at the aromatic ring exist and produce a whole family of CGAs (Oestreich-Janzen, 2019). Isomers in quinic acid are 3-CQA and 4-CQA, which are found in amounts of about 10% of 5-CQA (Oestreich-Janzen, 2019). Substitutions at the aromatic ring change the hydroxycinnamic moiety. Examples are p-coumaroylquinic acid and feruloylquinic acid, which are present in concentrations orders of magnitude lower than caffeoylquinic acids (Oestreich-Janzen, 2019).

1.2.3 Biosynthesis

CGAs are synthesized via the phenylpropanoid pathway (Clifford et al., 2017). The pathway initiates from aromatic amino acids phenylalanine, tryptophan, and tyrosine, which are converted from simple carbohydrate molecules in the shikimic pathway (Kumar & Goel, 2019). Phenylalanine synthesized in the shikimic pathway is sequentially catalyzed by phenylalanine ammonia lyase (PAL), cinnamate 4’-hydroxylase (C4H), and 4-cinnamoyl-CoA ligase (4CL) to yield p-coumaroyl-CoA, with trans-cinnamate and p-coumaric acid acting as intermediates (Clifford et
al., 2017). The final conversion to 5-CQA involves the enzymes hydroxycinnamoyl-CoA quinate hydroxycinnamoyl transferase (HQT) and p-coumaroyl-3’-hydroxylase (C3H), a cytochrome P450 oxidase (Clifford et al., 2017). Hydroxycinnamoyl transferases reversibly shuttle hydroxycinnamoyl units between their CoA and esterified forms (Lallemand et al., 2012). HQT has a stronger preference for quinate, whereas hydroxycinnamoyl-CoA shikimate hydroxycinnamoyl transferase (HST) has a stronger preference for shikimate (Lallemand et al., 2012). In addition, both enzymes accept a range of acyl donors that include p-coumaroyl-CoA, caffeoyl-CoA, and feruloyl-CoA (Lallemand et al., 2012).

In CGA synthesis, HQT catalyzes p-coumaroyl-CoA to produce 5-O-p-coumaroylquinic acid, which is subsequently hydroxylated by C3H to yield 5-CQA (Clifford et al., 2017). Alternatively, p-coumaroyl-CoA is hydroxylated by C3H to yield caffeoyl-CoA, and HQT catalyzes the formation of 5-CQA (Clifford et al., 2017). It is currently believed that the primary route for 5-CQA biosynthesis is via p-coumaroyl-CoA and quinic acid (Niggeweg et al., 2004), where caffeic acid is generated by the combined action of HQT and HST and a P450 3’-hydroxylase (Abdulrazzak et al., 2006). Considerably less information is known regarding biosynthesis of other CQA isomers. The perceived wisdom is that other CQAs are derived from 5-CQA, but there is little data on the isomerases that would be involved in such conversions (Clifford et al., 2017).

A study in fresh-cut carrot demonstrated that CGA synthesis and PAL activity are increased when exposed to UV light, particularly that of UVB and UVC, indicating that CGA formation is induced by UV radiation (Surjadinata et al., 2017). Another study focused on blueberries showed that a higher degree of UV radiation in the growth period of two subsequent harvest years resulted in significantly different CGA content among fruit (Pertuzatti et al., 2021). Furthermore,
CGA synthesis was found to be upregulated by blue light, which acts by inducing expression of HST (Chen et al., 2020).

1.2.4 Biological functions of chlorogenic acids in plants

Plants in nature are constantly under environmental stresses that stimulate the production of reactive oxygen species (ROS) (Akula & Ravishankar, 2011; Mei et al., 2020). Excess ROS produced under stress disrupts normal metabolism by causing oxidative damage to lipids, proteins, and nucleic acids and hampers plant growth and development (Akula & Ravishankar, 2011; McCord, 2000). CGAs are a group of phenolic secondary metabolites that scavenge ROS and metals and provides protection against a broad range of insect herbivores and different pathogens (Kundu & Vadassery, 2019; Meng et al., 2013; Sung & Lee, 2010).

Antioxidant activity

Mei et al. (2020) reported that CGA alleviated induced oxidative stress in detached apple leaves by decreasing rapid chlorophyll loss, reducing membrane damage and lipid oxidation, and stimulating antioxidant enzyme activity. Similarly, Niggeweg et al. (2004) demonstrated that plants with higher levels of CGA, induced by overexpression of its synthesis enzyme, exhibited improved antioxidant capacity and resistance to infection by a bacterial pathogen.

Anti-microbial, anti-fungal, anti-viral activity

CGAs are one of the initial products formed during transcriptional activation of the phenylpropanoid pathway in response to pathogen infection and abiotic stress events (Grace & Logan, 2000; Lee & Bostock, 2007; Rahman et al., 2015). A study in potatoes showed that higher concentrations of CGA were present in plants responding to bacterial invasion, which may
support the association of CGAs with imparting resistance to bacterial infection (Ngadze et al., 2011). Furthermore, another study reported that CGA effectively inhibited bacterial growth of all tested pathogens by significantly increasing outer and plasma membrane permeability, leading to loss of barrier function and cell death (Lou et al., 2011). Also, CGA inhibited spore germination and reduced hyphal growth of plant necrotrophic fungi (Martínez et al., 2017).

Protection against herbivores and pests

A metabolome comparison of thrips-resistant and -susceptible chrysanthemums by Leiss et al. (2009) showed that higher amounts of CGA are present in thrips-resistant cultivars. Moreover, an examination of the sensitivity of cabbage butterfly larvae to feeding deterrents identified CGA as one of the most prominent deterrents in nasturtium (Huang & Renwick, 1995).

Protection against abiotic stressors

Recently, Gharibi et al. (2019) investigated the effects of drought stress on polyphenolic compounds and reported that CGA was the most abundant phenolic acid among those that continued to accumulate with increasing duration of stress. In addition, Jiang et al. (2020) demonstrated that accumulation of CGA is positively correlated with overexpression of a synthase gene involved in reducing the negative effects of abiotic stress. Furthermore, a study in lettuce showed that CGA levels are lower in cultivated lettuce in comparison to those in wild lettuce, which may suggest a role for CGA in plant adaptability to environmental stresses (Zhang et al. 2020b).
Signaling molecule effects

Phenolic acids have been reported to act as signaling molecules in the establishment of arbuscular mycorrhizal and legume-rhizobia symbioses and play multifunctional roles in rhizospheric plant-microbe interactions (Mandal et al., 2010; Martens, 2002). Mhlongo et al. (2014) revealed that structurally and functionally diverse priming agents of plant defense all trigger the accumulation of CGAs in tobacco cells. More recently, the authors found 20 cinnamic acid derivatives, including CGAs, that were implicated as discriminatory biomarkers of priming in tobacco cells (Mhlongo et al., 2016).

1.2.5 Health benefits in humans

CGA is one of the most abundant beneficial polyphenols in the human diet and is well known as a nutritional antioxidant in plant-based foods (Kundu & Vadassery, 2019; Meng et al., 2013). Dietary consumption of CGA is associated with the prevention of certain oxidative and degenerative, age-related diseases (Rodriguez de Sotillo et al., 2006; Rodriguez de Sotillo & Hadley, 2002; Suzuki et al., 2006; Upadhyay & Mohan Rao, 2013; Wang et al., 2007). Compelling evidence indicates that dietary CGA can promote a wide range of pharmacological effects and biological activities in various tissues and organs (Lu et al., 2020). Numerous studies have demonstrated the antioxidant activities of CGA, which inhibit formation of ROS or scavenge them (Kweon et al., 2001). CGA is also negatively correlated with the risk of various harmful conditions, such as oxidative stress (Liang & Kitts, 2015), inflammatory stress (Liang & Kitts, 2015), type 2 diabetes mellitus (Ong et al., 2013; Paynter et al., 2006), cardiovascular disease (Morton et al., 2000; Wan et al., 2013), neurodegenerative disease (Larsson et al., 2011), and cancer (Jiang et al., 2001; Jin et al., 2005; Matsunaga et al., 2002; Mori et al., 1986; Morishita et al., 1997; Nakamura et al., 1997; Tajik et al., 2017).
Antioxidant activity

It has been suggested that consumption of beverages like coffee, tea, and wine is linked to a decreased risk of developing chronic diseases (Liang & Kitts, 2015). This health benefit is attributed to intake of CGAs, the underlying mechanism(s) of which involve mitigating oxidative stress (Liang & Kitts, 2015). Kasai et al. (2000) demonstrated that extracts of various vegetables and fruits share CGA as a common inhibitor of 8-hydroxydeoxyguanosine (8-OH-dG), a key marker of cellular oxidative stress during carcinogenesis. In addition, Tamagnone et al. (1998) showed that inhibition of phenolic acid metabolism in tobacco plants led to premature cell death and abnormal leaf development, which may be attributed to increased oxidative stress and ROS concentration from deficiency of the plants’ natural antioxidants.

Anti-inflammation activity

Oedema is considered one of the main signs of acute inflammation (dos Santos et al., 2006). A study in inflammation and pain rat models found that CGA inhibited paw edema and pain, likely though its inhibitory action in the peripheral synthesis/release of inflammatory mediators involved in these responses (dos Santos et al., 2006). Another study suggested that CGA could be a novel therapeutic agent against inflammatory hyperplasia of the synovium, which acts by enhancing synoviocyte apoptosis in patients with rheumatoid arthritis (Lou et al., 2016). Furthermore, CGA can inhibit interleukin 1 beta (IL-1β)-induced inflammatory response via the nuclear factor kappa B (NF-κB) signaling pathway (Liu et al., 2017), as well as attenuate dextran sulfate sodium-induced colitis in mice by suppressing NF-κB, p-STAT3, p-Akt, p-JNK, and p-ERK expression in the colon (Vukelić et al., 2018).
Anti-diabetic activity

Diabetes mellitus is a chronic metabolic disease caused by an absolute or relative lack of insulin and/or reduced insulin activity, which results in hyperglycemia and impaired metabolism of carbohydrates, proteins, and fats (Kamtchouing et al., 2006). α-glucosidases play important roles in the digestion of carbohydrates, and inhibitors of α-glucosidases like CGA are considered promising candidates for the development of anti-type 2 diabetics (Ma et al., 2008). Sanchez et al. (2017) found that CGA acts as an insulin secretagogue, insulin sensitizer, and lipid-lowering agent, thereby possessing multiple mechanisms of actions in the development of a highly effective therapeutic for the treatment of type 2 diabetes. In addition, Bao et al. (2018) reported that CGA significantly attenuated diabetic renal damage by modulating the Nrf2/HO-1 and NF-κB signaling pathways in streptozotocin-induced diabetic rat kidney. Also, a study in non-alcohol fatty liver disease rats showed that treatment with CGA ameliorated insulin resistance and liver injury by inactivating the c-Jun N-terminal kinase (JNK) pathway, and consequently, suppressing autophagy (Yan et al., 2018).

Cardiovascular disease effects

Collectively, the varying forms of cardiovascular disease are the leading causes of death worldwide (Johnson, 2019). CGA is well known for its many biological properties, including antioxidant and anti-inflammatory activities that can affect the risk factors that contribute to development of cardiovascular diseases (Everson-Rose & Lewis, 2005; Park, 2009; Wang et al., 2016a; Yukawa et al., 2004). Furthermore, consumption of CGA has antihypertensive effects by improving endothelial function and NO bioavailability in the arterial vasculature (Zhao et al., 2012). An in vitro study showed that CGA can mitigate oxidized low-density lipoprotein (LDL)-induced oxidative stress and mitochondrial dysfunction by upregulating SIRT1 and modulating...
the AMPK/PGC-1 signaling pathways (Tsai et al., 2018). Additionally, Fuentes et al. (2014) found that CGA inhibited platelet activation by stimulating the A2A receptor/adenylate cyclase/cAMP/PKA signaling pathway, as well as significantly reduced several platelet inflammatory mediators like soluble P-selectin and IL-1β.

**Neuroprotective effects**

Neurodegenerative diseases have been reported to strongly correlate with inflammation and accumulation of oxidative stress-induced damage (Kim & Lee, 2015). Pre-clinical and clinical studies have suggested that consumption of CGA protects against neurological degeneration and its resulting diseases associated with oxidative stress in the brain (Heitman & Ingram, 2017). A study in rat brain found that CGA exerts neuroprotective effects by inhibiting the activity of two key enzymes, namely acetylcholinesterase and butyrylcholinesterase, as well as suppressing oxidative stress-induced neuronal damage (Oboh et al., 2013). Additionally, Teraoka et al. (2012) demonstrated that CGA exhibited cytoprotective effects against α-synuclein-related toxicity in PC12 cells by suppressing the interaction of oxidized dopamine with α-synuclein. Moreover, CGA also significantly reversed cognitive impairments in scopolamine-induced amnesia mice by inhibiting acetylcholinesterase and malondialdehyde in the hippocampus and frontal cortex (Kwon et al., 2010).

**Anti-cancer activity**

CGA is suggested to reduce the occurrence and progression of carcinogenesis in both *in vivo* and *in vitro* models of cancer (Matsunaga et al., 2002; Mori et al., 1986; Yan et al., 2017). According to Yan et al. (2017), CGA prevented the progression of hepatocellular carcinoma (HCC) cells through multiple pathways. Moreover, CGA inhibited the proliferation of HepC2 cells (human
hepatocellular carcinoma cell lines) *in vitro* and the progression of HepG2 xenografts *in vivo* by inducing inactivation of extracellular signal-regulated kinase 1/2 (ERK1/2) and imbalance of MMP-2 and TIMP-2. Huang et al. (2019) also found that CGA reduced proliferation rate, migration/invasion ability, and mitochondrial ATP production in cancer cells but not in normal cells. In addition, Hou et al. (2017) reported that CGA inhibited viability of human colon cancer cells in a dose-dependent manner by inducing ROS.

### 1.3 Flavonols

#### 1.3.1 Introduction

Flavonols are one of six subclasses of flavonoids, including anthocyanins, flavanols, flavanones, flavonones, and isoflavones (Hollman & Arts, 2000; Kumar & Pandey, 2013). Flavonoids belong to a large group of plant secondary metabolites with variable polyphenolic structures and a broad spectrum of biological activities (Hernández-Rodríguez et al., 2019; Panche et al., 2016). Flavonols are typically present in plants bound to sugars as glycosides (Hollman & Arts, 2000). Free flavonoids without attached sugars, referred to as aglycones, are not usually present in fresh plants, except due to food processing (Hollman & Arts, 2000). The glycosylated forms are believed to possess higher water solubility and mobility than the aglycone forms, making them less likely to interfere with other critical cellular metabolisms when stored in plant cells (Ververidis et al., 2007). Sugar moieties, predominantly bound to the flavonoid nucleus *via* a β-glycosidic bond, can bind to various positions but primarily do so at the 3-position (Hollman & Arts, 2000). Sugars are mono-, di-, tri-, and tetrasaccharides (Hollman & Arts, 2000); however, most flavonoids occur as monomer glycosides (Häkkinen & Auriola, 1998; Ververidis et al., 2007). Mono- and di-substituted B ring flavonols are kaempferol, quercetin, and isorhamnetin,
whereas tri-substituted B ring flavonols are myricetin, laricitrin, and syringetin (Castillo-Muñoz et al., 2007).

Flavonols possess very diverse methylation and hydroxylation patterns, making them perhaps the largest subclass of flavonoids (Panche et al., 2016). They are also the most ubiquitous in foods and widely distributed in nature (El Gharras, 2009; Perez-Vizcaino & Duarte, 2010). They are found in a variety of fruits and vegetables, including onions, kale, tomatoes, apples, grapes, berries, leeks, and broccoli, as well as beverages such as tea and red wine (Hollman & Arts, 2000; Manach et al., 2004; Panche et al., 2016). Flavonols are generally found in relatively low concentrations of about 15-30 mg/kg FW in plant-based foods (El Gharras, 2009). In onions, however, flavonols are present in concentrations up to 1.2 g/kg FW (Manach et al., 2004), and in tea and red wine, they may be present in concentrations up to 45 mg/L (El Gharras, 2009). Among flavonols, quercetin and kaempferol are the main representative molecules (Durazzo et al., 2019; Manach et al., 2005), with quercetin considered the most abundant dietary flavonol (El Gharras, 2009).

1.3.2 Structure

Flavonoids are composed of a phenylbenzopyran chemical structure with a C6-C3-C6 backbone joined to a chroman ring (Figure 1.2A) (Pereira et al., 2009). In most cases, they also have three or more hydroxyl groups linked to the backbone structure (Zhang & Tsao, 2016). Flavonoids are divided into different subclasses based on the carbon of the C ring on which the B ring is attached and the degree of unsaturation and oxidation of the C ring (Panche et al., 2016). These subclasses are further divided into groups depending on the structural features of the C ring (Panche et al., 2016). Flavonols belong to the subclass in which the B ring is linked to the 2-
position of the C ring (Panche et al., 2016). They possess a hydroxyl group in the 3-position of the C ring, which may be glycosylated (Panche et al., 2016), and a double bond at the 2-3 position (Hollman & Arts, 2000). They also have a ketone group (Panche et al., 2016). Flavonols share a 3-hydroxy-2-phenylchromen-4-one common structure, and 90% of plant flavonols have a hydroxyl group in the 5-position and/or 7-position of the A ring (Macheix et al., 1990).

1.3.3 Biosynthesis

Flavonoids are synthesized through a combination of the phenylpropanoid and polyketide pathways (Saito et al., 2013). The phenylpropanoid pathway provides p-coumaroyl-CoA, and the polyketide pathway elongates the C-2 chain by utilizing malonyl-CoA as the condensing unit (Saito et al., 2013). The phenylpropanoid pathway initiates from aromatic amino acids phenylalanine and tyrosine, which are synthesized by the shikimic pathway (Tohge et al., 2017). First, phenylalanine is catalyzed by phenylalanine ammonia lyase (PAL) to form trans-cinnamic acid with the concomitant loss of ammonia (Saito et al., 2013). This trans-cinnamic acid is hydroxylated by cinnamic acid 4-hydroxylase (C4H) at the 4-position to produce p-coumaric acid. To facilitate further conversion, p-coumaric acid requires activation to the corresponding CoA thioester (Saito et al., 2013). 4-coumaric acid:CoA ligase (4CL) catalyzes p-coumaric acid via an ATP-consuming condensation reaction to yield p-coumaroyl-CoA (Tohge et al., 2017).

Entering the central flavonoid biosynthetic pathway, a complex series of reactions including condensation, isomerization, oxidation, and reduction take place to produce flavonoid scaffolds that undergo further modifications via tailoring reactions (Saito et al., 2013). Chalcone synthase (CHS) is the first committed enzyme in the flavonoid pathway (Saito et al., 2013). It catalyzes the formation of a triketide intermediate from p-coumaroyl-CoA and three molecules of malonyl-
CoA (Saito et al., 2013). This intermediate is spontaneously cyclized to yield naringenin chalcone, which is then catalyzed by chalcone isomerase (CHI) to yield (2S)-naringenin (flavanone). Next, flavanone 3-hydroxylase (F3H) catalyzes the oxygenation of (2S)-naringenin at the 3-position to make dihydrokaempferol (dihydroflavonol). Flavonoid 3′-hydroxylase (F3′H), a cytochrome P450 monooxygenase, hydroxylates the 3′-position of the flavonoid B ring (Saito et al., 2013). It accepts either dihydrokaempferol or kaempferol as a substrate and converts them to dihydroquercetin and quercetin, respectively (Saito et al., 2013). Flavonoid 3′,5′-hydroxylase (F3′5′H), also belonging to the cytochrome P450 super family, catalyzes hydroxylation at the 3′ and 5′ positions of the flavonoid B ring (Jeong et al., 2006). It competes with F3′H for dihydrokaempferol as a substrate and converts it to dihydromyricetin (Jeong et al., 2006).

Flavonol synthase (FLS) is the first committed enzyme in the flavonol pathway, catalyzing the first step that branches off from the main trunk pathway towards anthocyanin biosynthesis (Tohge et al., 2017). In the flavonol pathway, dihydroquercetin and dihydromyricetin are catalyzed by FLS, forming a double bond at the 2-3 position, to yield quercetin (Figure 1.2B) and myricetin (Figure 1.2D), respectively (Jeong et al., 2006). Alternatively, flavonol synthesis can also occur via catalyzation of dihydrokaempferol by FLS to produce kaempferol (Figure 1.2C), which is then converted to quercetin by F3′H (Saito et al., 2013).

The great chemical diversity of flavonoids is attributed to the high number of tailoring reactions carried out by glycosyltransferases, methyltransferases, and acyltransferases. Glycosylation of flavonoid aglycones occurs at hydroxyl moieties of the 3-, 5-, and 7-positions and is catalyzed by flavonoid glycosyltransferase (FGT) (Saito et al., 2013; Tohge et al., 2017). Flavonoid
methyltransferase (FMT) is involved in the methylation of flavonols and produces isorhamnetin (Tohge et al., 2017).

Flavonol biosynthesis is a light-dependent process that can be induced by environmental stresses (Czemmel et al., 2009; Dixon & Paiva, 1995). Czemmel et al. (2009) reported that light induced expression of a transcription factor established to be the specific activator of FLS, and of several other promoters of genes involved in flavonol synthesis. Additionally, shaded fruit exhibited reduced expression of FLS and showed less than 10% the flavonol levels contained in light-exposed fruit at time of harvest (Downey et al., 2004). Also, Price et al. (1995) revealed that marked differences in flavonol content exist not only between pieces of fruit on the same tree but also between different sides of the same piece of fruit, depending on sun exposure.

Furthermore, flavonols have been shown to accumulate in outer and aerial tissues due to their light-stimulated biosynthesis (Seeram, 2006). They are present exclusively in the outer epidermis and in some layers of seed coat and stored in the inner, thick-walled layers of the hypodermis (Fontes et al., 2011; McDonald et al., 1998). Therefore, the thickness of berry epidermis influences their concentration (McDonald et al., 1998). Zifkin et al. (2012) reported that flavonols were detected in the exocarp, mesocarp, and placental tissues of young stage 1 blueberry fruit but only minimally in the mesocarp of larger stage 3 fruit. Additionally, flavonols were observed in the placentae and exocarp through developing stage 6 fruit, in the placentae and cuticle through fully mature stage 8 fruit, and in the developing seed. Winkel-Shirley (2002) suggested that the localization of flavonols is related to its role as a protectant against UV radiation, extreme temperatures, and free radicals.
1.3.4 Biological functions of flavonols in plants

Flavonoids are secondary metabolites that display a variety of functional roles in higher plants in response to a wide array of environmental stimuli (Dixon & Paiva, 1995; Taylor & Grotewold, 2005; Winkel-Shirley, 2002). There have been more than 6,500 flavonoid compounds identified in various plants and greater than 8,000 structures reported (Forkmann & Martens, 2001; Harborne & Williams, 2000; Pietta, 2000). Despite this fact, flavonols (one of three major subclasses of dietary flavonoids) are almost exclusively responsible for the responses of plants to stressful agents and act solely as developmental regulators and/or signaling molecules. They show the greatest potential to keep stress-induced changes in reactive oxygen species (ROS) homeostasis under control and to regulate the development of individual organs and the whole plant, even at very low concentrations (Pollastri & Tattini, 2011).

Antioxidant activity

Watkins et al. (2017) reported that flavonols facilitated stomatal aperture opening to modulate leaf gas exchange by dampening the abscisic acid (ABA)-dependent ROS bursts in guard cells that drive stomatal closure. Additionally, antioxidant activities of extracts of roots, twigs, and leaves of barberry showed that total phenol and flavonol contents are well correlated with antioxidant activity in the studied plant organs (Zovko Končić et al., 2010).

Anti-microbial, anti-fungal, anti-viral activity

Jia et al. (2010) found that quercetin induced resistance to the virulent strain *Pseudomonas syringae* pv. Tomato DC3000 (Pst) in *Arabidopsis thaliana* via $\text{H}_2\text{O}_2$ burst and involvement of both salicylic acid and NPR1 protein. However, many pathogens can still cause disease in flavonol-rich plant tissues. Chen et al. (2019) revealed that a quercetin dioxygenase gene in
*Sclerotinia sclerotiorum*, a necrotrophic fungal pathogen, catalyzes cleavage of the flavonol carbon skeleton. In the absence of the gene, flavonols were found to exhibit toxicity towards the pathogen.

*Protection against herbivores and pests*

Flavonols play an important role in plants’ interactions with other organisms, acting as repellants, antifeedants, or even toxins (Haribal & Feeny, 2003; Mallikarjuna et al., 2004; Mierziak et al., 2014; Nuessly et al., 2007). Goławska et al. (2014) determined that higher concentrations of the flavanone naringenin and flavonol quercetin increased development time, pre-reproductive period, and mortality in pea aphids, as well as decreased fecundity and their intrinsic rate of natural increase. In addition, Sosa et al. (2004) showed that flavonoids in the exudate of *Cistus ladanifer*, of which 3,7-di-O-methylkaempferol is predominant, elicited an avoidance reaction in herbivores by impairment of mouth skeletal muscle relaxation.

*Protection against abiotic stressors*

Experiments with wild-type and overexpression mutants demonstrated that flavonoids can act as mitigators of oxidative and drought stress in *Arabidopsis* via over accumulation of flavonol glycosides and anthocyanins (Nakabayashi et al., 2014). Furthermore, evaluation of *Nigella sativa* callus cultures indicated that total flavonol and anthocyanin contents were highly associated with salinity stress-induced antioxidant activity, which may suggest that secondary metabolites enhance plant defense mechanisms against oxidative damage induced by salinity stress (Golkar et al., 2020).
**Signaling molecule effects**

According to Brunetti et al. (2019), a flavonol-phytohormone relationship might be at the core of land plant evolution. Moreover, the ABA signaling pathway is thought to have promoted biosynthesis of flavonols in early bryophytes, which offered further protection against novel environmental stressors encountered by early land plants (Brunetti et al., 2019). In turn, according to Watkins et al. (2014), flavonols regulate the ABA signaling pathway by suppressing its second messengers. Also, flavonols can act as modulators of auxin transport during nodulation (Zhang et al., 2009). In particular, Zhang et al. (2009) noted accumulation of kaempferol led to localized inhibition of auxin transport, which resulted in initiation of nodule primordia, and ultimately, a functional nitrogen-fixing nodule.

### 1.3.5 Health benefits in humans

Flavonols are among the most consumed dietary flavonoid subclasses in the U.S. diet (Chun et al., 2007). Moreover, their health-promoting effects are suggested to be stronger than that of other subclasses (Barreca et al., 2021). Intake of dietary flavonols, which play a key role in reducing oxidative damage (Shahidi & Ambigaipalan, 2015), is associated with decreased risk of various harmful conditions (Hertog et al., 1993; Hollman & Katan, 1998; Keli et al., 1996; Knekt et al., 1997). Many studies have investigated flavonols as potential pharmaceuticals and novel therapeutics (Eitah et al., 2019; Wang et al., 2020). Results have shown that high dietary intake of flavonols is negatively correlated with the risk of non-communicable chronic diseases, such as oxidative disease (Caddeo et al., 2019; Wang et al., 2016b), chronic inflammation (Piovezana Bossolani et al., 2019), type 2 diabetes mellitus (Eitah et al., 2019), cardiovascular disease (Hügel et al., 2016; Serban et al., 2016; Shabrova et al., 2011), neurodegenerative disease (Holland et al., 2020; Zhang et al., 2020c), and cancer (Chahar et al., 2011).
Antioxidant activity

Dietary flavonols play a key role in human health by reducing oxidative damage, which may help reduce the severity of chronic diseases (Barreca et al., 2021; Crozier et al., 2009). The free radical scavenging activity of flavonoids was shown to occur via electron donation from free hydroxyl groups on the nucleus (Pannala et al., 1997). Therefore, flavonols, which possess three hydroxyl groups, could be the strongest antioxidants among flavonoid compounds (Pannala et al., 1997). Singh et al. (2011) reported that curry leaf flavonols prevented cupric ion-induced oxidation of LDLs. In addition, Li et al. (2010) showed that icariin induced apoptosis in human hepatoma cells via a ROS/c-Jun N-terminal kinase (JNK)-dependent mitochondrial pathway. Furthermore, according to Lee et al. (2010), kaempferol reduced 2-deoxy-D-ribose reducing sugar-mediated pancreatic β-cell damage by interfering with ROS metabolism, thereby protecting the cells from lipid peroxidation.

Anti-inflammation activity

Acute inflammation is a normal protective response of living tissue to injury or infection and is characterized by pain, fever, redness, swelling, and increased blood flow to the tissue (Rankin, 2004). Chronic inflammation, on the other hand, is associated with harmful diseases that collectively represent the most significant cause of mortality worldwide (Furman et al., 2019). A study in rat model of inflammation and pain demonstrated that kaempferol glycosides reduced inflammation by blocking synthesis of prostaglandins important to the genesis of oedema and pain responses (De Melo et al., 2009). Furthermore, Crespo et al. (2008) determined that kaempferol and quercetin modulate pro-inflammatory genes in a concentration-dependent
manner via interference with the activation of NF-κB and dAP-1. Likewise, isorhamnetin inhibited pro-inflammatory gene expression by inactivating NF-κB (Yang et al., 2013).

**Anti-diabetic activity**

A diet rich in polyphenols might help in the prevention of type 2 diabetes (Rienks et al., 2018). Moreover, epidemiological data revealed that habitual flavonol intake is associated with lower incidence of type 2 diabetes (Jacques et al., 2013). Zhang and Liu (2011) found that the anti-diabetic effects of kaempferol are exerted through the protection of β-cells against glucotoxicity. A study in non-obese diabetic mice demonstrated that epicatechin prevents the onset of type 1 diabetes by modulating immune function, thereby preserving islet mass (Fu et al., 2013).

**Cardiovascular disease effects**

It is well known that high intake of dietary flavonoids is beneficial against cardiovascular diseases (Manach et al., 2005; Medina-Remón et al., 2017; Tresserra-Rimbau et al., 2014; Yochum et al., 1999; Zern & Fernandez, 2005). Quercetin is a prominent dietary flavonol with strong antioxidant and anti-inflammatory properties that are associated with the prevention and therapy of cardiovascular diseases (Boots et al., 2008; Mirsafaei et al., 2020; Russo et al., 2012). A trial demonstrated that quercetin lowers blood pressure in stage 1 hypertensive individuals (Edwards et al., 2007). In addition, a study focused on the anti-atherosclerotic property of flavonols indicated that quercetin 3-(6-malonylglucoside) can attenuate atherosclerotic lesion development by enhancing LDL resistance to oxidative modification (Enkhmaa et al., 2005). Furthermore, an in vitro and in vivo study showed that dietary quercetin reduced inflammatory cardiovascular risk factors and alleviated atherosclerosis in a humanized
transgenic model of cardiovascular disease (Kleemann et al., 2011). Also, quercetin was found to significantly decrease the activities of inflammatory enzymes and mediators during atherosclerosis progression and regression (Bhaskar et al., 2013).

Neuroprotective effects

Neurodegenerative diseases are a large, heterogeneous group of neurological disorders that affect distinct subsets of neurons in specific anatomical locations (Maher, 2019). Several studies have investigated flavonols as preventive and therapeutic treatments of age-related neurodegenerative diseases (Bombardi Duarte et al., 2018; Hou et al., 2010; Sandhir & Mehrotra, 2013). The neuroprotective benefits of flavonols may be attributed to their antioxidant, anti-proliferative, and anti-inflammatory properties (Bombardi Duarte et al., 2018). Moreno et al. (2017) showed that quercetin with higher oral absorption and bioavailability improved the cognition and memory impairments characteristic of Alzheimer’s disease. In addition, Suganya and Sumathi (2017) demonstrated that rutin significantly improved the biochemical, behavioral, and histological alterations induced by 3-nitropropionic acid (3-NPA) administration in rat models of Huntington’s disease (HD). Similarly, another study in 3-NPA-induced neurodegeneration found that kaempferol exerted effective protection against striatal degeneration and largely attenuated the neurological deficits in HD rats (Lagoa et al., 2009).

Anti-cancer activity

Cancer is largely environmentally determined, and dietary polyphenols are closely related with the risk for several types of cancer (Block et al., 1992; Ramos, 2008). Epidemiological data has shown that flavonol consumption decreases the incidence of cancer (Lea, 2015). Additionally, flavonols exert their anti-cancer effects, both in vitro and in vivo, via several genes, enzymes,
and signaling pathways involved in different stages of carcinogenesis (Chahar et al., 2011; Khan et al., 2020). Chen and Chen (2013) revealed that dietary kaempferol significantly inhibited cancer cell growth and angiogenesis and induced cancer cell apoptosis, while preserving normal cell viability, by modifying cellular signal transduction pathways. Furthermore, Lang et al. (2020) found that chrysosplenol D induced ERK1/2-mediated apoptosis in triple negative human breast cancer cells, possibly by activating autophagy. Moreover, evidence from rat mammary cancer models indicated that quercetin inhibited both the incidence and number of palpable tumors (Verma et al., 1988).

1.4 Bioavailability

In the last decade, there has been an increase in research interest in phenolic compounds, likely due to a growing awareness of their antioxidant power and potential health benefits (Cosme et al., 2020). The ability of phenolic compounds to exert biological activities primarily depends on their bioavailability, which broadly defined, refers to the fraction of a nutrient in a food that is absorbed and utilized (Cosme et al., 2020; Wood, 2005). Therefore, further studies on the absorption and bioavailability of phenolic compounds are necessary to evaluate their protective health effects.

1.4.1 Chlorogenic acids

The absorption and bioavailability of CGAs are still controversial due to disparities in evidence between different studies. It was initially believed that CGA might be poorly absorbed through the small intestine barrier (Farah & de Paula Lima, 2019). According to Azuma et al. (2000), oral administration of CGA in rats resulted in detection of only traces of metabolites in plasma, which suggests poor absorption from the digestive tract. In humans, analysis of plasma collected
after administration of coffee, a common beverage with high levels of CGA, did not detect CGA (Nardini et al., 2002).

Subsequent studies, however, have demonstrated that CGA is absorbed in the stomach and/or small intestine. After ingestion of CGA supplement, traces of it were detected excreted in urine (Olthof et al., 2001). The authors concluded that a maximum of 33% of the ingested CGA was absorbed from the small intestine, which implied that some CGA from food will enter the blood circulation. Furthermore, according to Lafay et al. (2006), CGA is absorbed as its intact form in the stomach and as its hydrolyzed forms in the small intestine. The remaining intact CGA reaches the large intestine and is then hydrolyzed and further metabolized into other aromatic acids by colonic microbiota and gastric esterase prior to reabsorption (Olthof et al., 2001; Stalmach et al., 2010; Tomas-Barberan et al., 2014; Williamson & Stalmach, 2012). Notably, at least six major CGA compounds were detected in human plasma after acute coffee consumption, although the compounds were differentially absorbed and/or metabolized, with large inter-individual variation (Monteiro et al., 2007).

Currently, it is thought that the inconsistent bioavailability of CGA may arise from differences in dosage, duration, model, organ, cell types, inter-individual variation, limitation of detection assay, and/or degradation of sample (Lu et al., 2020). The transformation of CGA by human gut microbiota might also contribute to potential inter-individual differences in bioavailability (Tomas-Barberan et al., 2014). Furthermore, different isomers differ in their susceptibility to intestinal chlorogenic acid esterase, with 5-CQA hydrolyzed more readily than 3-CQA, and 4-CQA being particularly resistant to hydrolysis (Guy et al., 2009).
1.4.2 Flavonols

Similarly, it was earlier thought that absorption of flavonols from the diet was negligible because of their presence in plants bound to sugars as β-glycosides (Hollman & Arts, 2000). Only aglycones were considered absorbable, while glycosides were considered non-absorbable. Indeed, studies with germ-free rats seemed to indicate that enzymes capable of splitting the β-glycosidic bonds are not secreted into the gut or present on the intestinal wall (Griffiths & Barrow, 1972). Bacteria in the colon could hydrolyze glycosides but would degrade the liberated aglycones (Griffiths & Smith, 1972). Therefore, only marginal absorption of flavonoid glycosides was expected (Kühnau, 1976).

Later studies demonstrated the occurrence of flavonol absorption in humans, although the question as to which form is effectively absorbed remained (Aherne & O’Brien, 2002; Wang et al., 2016b). According to Hollman et al. (1995), significant amounts of quercetin in both aglycone and glycosidic forms were absorbed in healthy ileostomy volunteers, chosen to circumvent the problem of bacterial degradation. Moreover, absorption of quercetin glucoside (52%) was far greater than that of the aglycone (24%). Consistently, quercetin from onions has been reported to be absorbed more efficiently than that from apples or rutin (Hollman et al., 1996, 1997). Onions contain mostly quercetin glucosides, whereas apples contain a mixture of galactosides and xylosides, indicating that the attached sugar moiety affects the rate of absorption and is an important determinant in absorption and metabolism of flavonols (Hollman et al., 1999; Manach et al., 2005).

Notably, a study on the effect of the sugar moiety on absorption of dietary flavonoid glycosides suggested that quercetin glucoside is absorbed from the small intestine, whereas the rutinoside
might be absorbed from the colon (Hollman et al., 1999). Absorption occurs less readily in the colon than in the small intestine due to a smaller exchange area and lower density of transport systems (Manach et al., 2004). Consequently, the rutinoside of quercetin is absorbed much less efficiently than the glucoside or aglycone (Hollman et al., 1999). On the other hand, it was revealed that the position of the sugar moiety does not affect the time course of quercetin concentration in plasma, indicating that the nature, and not the position, of the sugar moiety plays a role in absorption (Olthof et al., 2000).

1.5 Objectives

The available scientific evidence regarding the beneficial health effects of blueberries mainly concerns tetraploid highbush blueberries. However, there are numerous blueberry species at the diploid level that offer diverse germplasm (Wang et al., 2019), of which many have a limited body of analyses describing their genetics. Identification of quantitative trait loci (QTLs) for multiple segregating traits in a diploid blueberry population would add greater value to our insight into the genetic architecture of blueberry. Of particular interest is the evergreen subtropical *V. darrowii* species that appears to have fruit chemistry quite distinct from that of the highbush species (Baloga et al., 1995; Wang et al., 2019). This species was crossed to a diploid northern highbush (*V. corymbosum*) to give F1s, and then F1 x F1 crosses were subsequently made to produce pseudo F2 populations. Breeding at the diploid level should offer for simpler genetic analysis.

Therefore, the objectives of this thesis are to evaluate the variation for CGAs and flavonol glycosides in diploid blueberry genotypes and to phenotypically score the distribution of these
compounds in a segregating diploid F2 population for the intended purpose of future QTL mapping.

In the next chapter, our findings on the content and composition of CGAs, flavonol glycosides, and organic acids among diploid *V. corymbosum*, *V. darrowii*, and their F1s are given. Chapter 3 focuses on the phenotypic evaluation of CGAs and flavonol glycosides in a segregating F2 population derived from the six (four grandparents and two F1s) blueberry genotypes described in Chapter 2. The findings discussed in this thesis will contribute to our knowledge on blueberries’ genetics of fruit chemical composition, especially as it relates to its nutritional and health beneficial compounds.
1.6 References


Jeong, S. T., Goto-Yamamoto, N., Hashizume, K., & Esaka, M. (2006). Expression of the flavonoid 3′-hydroxylase and flavonoid 3′,5′-hydroxylase genes and flavonoid composition in grape...


Dietary Flavonol Quercetin. Cancer Research, 48(20), 5754 LP – 5758. http://cancerres.aacrjournals.org/content/48/20/5754.abstract


Serum Lipid Levels in Humans. Biochemistry (Moscow), 69(1), 70–74. https://doi.org/10.1023/B:BIRY.0000016354.05438.0f


Zovko Končić, M., Kremer, D., Karlović, K., & Kosalec, I. (2010). Evaluation of antioxidant activities and phenolic content of Berberis vulgaris L. and Berberis croatica Horvat. Food and
Chemical Toxicology, 48(8), 2176–2180.
https://doi.org/10.1016/j.fct.2010.05.025
1.7 Figures and tables

Figure 1.1 Chemical structure of 5-caffeoylquinic acid and 3-caffeoylquinic acid. A 5-caffeoylquinic acid isomer (chlorogenic acid) and B 3-caffeoylquinic acid (neochlorogenic acid).

Figure 1.2 Basic skeleton structures of flavonoid compound and flavonol aglycones. A flavonoid, B quercetin, C kaempferol, and D myricetin.
CHAPTER 2

Variation of phenolic compounds and organic acids in two diploid blueberry species, Vaccinium corymbosum and V. darrowii, and their F1 hybrids, using high-performance liquid chromatography

2.1 Introduction

Phenolic compounds are associated with the prevention of various chronic diseases such as cardiovascular diseases, type 2 diabetes, and cancer (Hollman & Katan, 1998; Hügel et al., 2016; Li et al., 2016; Yukawa et al., 2004). Diets high in fruits and vegetables (Block et al., 1992; Joshipura et al., 1999; Key, 2011; Willett, 1994), or replete in polyphenolic compounds like chlorogenic acids (CGAs) and flavonols (Hertog et al., 1993; Keli et al., 1996; Knekt et al., 1997), have been suggested to increase longevity and decrease incidence of these diseases.

Chlorogenic acids and flavonols offer strong antioxidant and anti-inflammatory properties that may reduce biomarkers and risk of chronic diseases by mitigating oxidative damage (Chen & Ho, 1997; Hotta et al., 2002; Liu et al., 2017; Moreira et al., 2013; Niggeweg et al., 2004; Piovezana Bossolani et al., 2019; Shahidi & Ambigaipalan, 2015).

Blueberries have become increasingly recognized due to their rich nutritional and bioactive compounds (Bujor et al., 2019). They contain high amounts of CGA, a major phenolic acid in ripe blueberry (Kang et al., 2015) and well-known nutritional antioxidant in plant-based foods (Kundu & Vadassery, 2019), which constitutes 10-16% of the total acids in blueberry fruit (Kalt & McDonald, 1996; Markakis et al., 1963). In addition, blueberries are one of the most concentrated sources of dietary flavonols in the human diet (Vrhovsek et al., 2012), although
flavonol content varies significantly among genotypes (Cho et al., 2005). Quercetin is considered a prominent dietary flavonol due to its high bioaccessibility (≈ 80%) (Gapski et al., 2019) and is present in blueberries up to 1.7-2.9 mg/100 g FW (fresh weight) (Bilyk & Sapers, 1986; Häkkinen et al., 1999). Furthermore, ripe blueberries have relatively high acid contents of 1-2%, of which citric acid is predominant (Ehlenfeldt et al., 1994; Song & Hancock, 2011). The composition of organic acids is influenced by multiple factors such as variety, fruit maturity, and environmental conditions (Zhang et al., 2020a). Organic acids also contribute to blueberry’s characteristic tart taste (Padmanabhan et al., 2016).

Compared with the work done on tetraploid highbush blueberry, there has been relatively little research conducted on the fruit chemistry of diploid blueberry species of the *Cyanococcus* subsection and what they have to offer for breeding and genetics of the crop. There are numerous blueberry species at the diploid level that can offer diverse germplasm (Wang et al., 2019). Of particular interest is the evergreen *V. darrowii* species native to Florida. This species is a source of many desirable traits such as low chilling requirement, resistance to mummy berry, late flowering and ripening, and various fruit quality traits that include powder blue color, firmness, and complex flavor (Baloga et al., 1995; Hancock et al., 2008; Ranger et al., 2007). Moreover, *V. darrowii* appears to have fruit chemistry that is unique from that of the highbush species (Baloga et al., 1995; Wang et al., 2019). Should sufficient variation of phenolic compounds exist between *V. corymbosum* and *V. darrowii*, it would be useful to construct a segregating F2 population. Identification of quantitative trait loci (QTLs) in such a diploid F2 population would offer greater insight into the genetics of fruit chemistry in blueberry.
In the present study, we investigated the content and composition of blueberry phenolic compounds in the parent species and their F1 hybrids. We particularly focused on the antioxidant compounds, including CGAs and flavonol glycosides. Also of interest, the fruit organic acid content was observed in the species and hybrids. HPLC methods were optimized to identify individual blueberry phenolics based on Wang et al. (2019) and provide accurate quantitative data. Four phenolic acids, ten flavonol glycosides, and three organic acids were analyzed. The objectives were to: (1) determine the content and composition of compounds in two clones of diploid *V. corymbosum*, two clones of diploid *V. darrowii*, and two F1 hybrids; (2) assess whether the compound levels are significantly different across species; and (3) investigate possible sources of variance, e.g., branch, plant clone, or harvest date, that may affect compound levels.

### 2.2 Materials and methods

#### 2.2.1 Plant materials

Four species (two *V. corymbosum* and two *V. darrowii* genotypes) and two F1 hybrids were used in this study. OPB-15 and OPB-8 are wild diploid *V. corymbosum* plants that were collected from a native population in Burlington County, NJ. NJ88-12-41 and NJ88-14-03 are wild diploid *V. darrowii* plants that were collected from native populations in Liberty County, Florida and along Route 98S in the Saint Joseph Bay area of Florida, respectively. In 2005, crosses were made between NJ88-14-03 and OPB-15 and between OPB-8 and NJ88-12-41. The resulting F1 hybrid plants were BNJ05-218-9 and BNJ05-237-8 (Figure 2.1). All plants were maintained in potted culture in a greenhouse through seasonal plant growth and winter dormancy (greenhouse maintained in a “cold” state, allowing for winter chilling at minimum 0-4 °C). Bumblebees (Koppert Biological Systems, MI, USA) were brought into the greenhouse in late spring/early
summer during flowering for open pollination for fruit set. Fully ripe, i.e., blue, blueberry fruit samples were harvested from each plant at 7-14-day intervals over the fruiting period, initiating at first harvest in April through last harvest in August, at the P.E. Marucci Center for Blueberry and Cranberry Research and Extension in 2019 to 2021. Samples were separately placed in polyethylene bags and kept chilled until weight measurements were taken. After weighing, samples were stored at -80 °C until analysis.

2.2.2 Chemicals and reagents
All solvents, including water, acetonitrile, methanol, and acetone, were purchased from EMD Millipore (Billercia, MA, USA) and were of HPLC grade. Acetic acid was purchased from Avantor Performance Materials (Center Valley, PA, USA), formic acid was purchased from Mallinckrodt baker (Phillipsburg, NJ, USA), and phosphoric acid was purchased from Amresco (Solon, OH, USA). Commercial standards of flavonol glycosides (quercetin-3-galactoside, quercetin-3-glucoside, quercetin-3-rutinoside, quercetin-3-rhamnoside, myricetin-3-galactoside, and myricetin-3-rhamnoside) were obtained from Indofine Chemical Company (Somerville, NJ, USA), and chlorogenic acid standard was obtained from Sigma-Aldrich, Inc. (St. Louis, MO, USA). Organic acid standards, including citric acid, malic acid, and shikimic acid, were obtained from Sigma-Aldrich, Inc. (St. Louis, MO, USA).

2.2.3 Extraction of blueberry phenolic compounds and organic acids
For chlorogenic acid and flavonol glycoside quantification, depending on sample availability, 2-8 g of fruits were weighed, average berry weight (AW) was recorded, and samples were ground with a Precellos Evolution homogenizer (Bertin Corp., Rockville, MD, USA) using 2.8 mm ceramic beads at 7200 rpm for 1.5 min. 80% aqueous acetone with 0.1% acetic acid (1:4 sample to
solvent w/v) was added to suspend ground fruit, and samples were extracted overnight at 4 °C. Liquid extracts were then centrifuged at 13,300 rpm for 2 min, and 1 mL aliquots of clear supernatant were taken. Aliquots were dried with a SpeedVac vacuum concentrator (Savant® SPD2010-220, Thermo Scientific, Waltham, MA, USA) under no heat and redissolved in 500 μL of 100% methanol by sonication for 15 min. Samples were then centrifuged at 11,000 rpm for 5 min, and clear supernatants were analyzed with high-performance liquid chromatography (HPLC).

For organic acid quantification, if remaining fruit sample was available, 2-5 g of fruits were weighed, AW was recorded, and samples were ground with a Precelllys Evolution homogenizer (Bertin Corp., Rockville, MD, USA) using 2.8 mm ceramic beads at 7200 rpm for 1.5 min. Distilled water (1:10 sample to solvent w/v) was added to suspend ground fruit, and then 1 mL aliquots of clear supernatant were heated in a 90 °C water bath for 10 min. Samples were then centrifuged at 11,000 rpm for 5 min, and clear supernatants were analyzed with HPLC.

For epidermis versus flesh quantification, depending on sample availability, 2-5 g of fruits were peeled using tweezers. The epidermal and flesh contents were weighed, AW was recorded, and samples were ground with a Precelllys Evolution homogenizer (Bertin Corp., Rockville, MD, USA) using 2.8 mm ceramic beads at 7200 rpm for 1.5 min. 80% aqueous acetone with 0.1% acetic acid (1:4 sample to solvent w/v) was added to suspend ground fruit tissues, and samples were extracted overnight at 4 °C. Liquid extracts were then centrifuged at 13,300 rpm for 2 min, and 1 mL aliquots of clear supernatant were taken. Aliquots were dried with a SpeedVac vacuum concentrator (Savant® SPD2010-220, Thermo Scientific, Waltham, MA, USA) under no heat and
reddissolved in 500 μL of 100% methanol by sonication for 15 min. Samples were then centrifuged at 11,000 rpm for 5 min, and clear supernatants were analyzed with HPLC.

### 2.2.4 HPLC apparatus and conditions

Three HPLC systems were used for phenolic compound and organic acid identification and quantification:

Chlorogenic acids and flavonol glycosides were analyzed in a Waters Alliance® LC system composed of a Waters e2695 Separations Module and Waters 2998 PDA Detector (Waters Corp., Milford, MA, USA). A Gemini® 150 x 4.6 mm 5 μm C18 110 Å LC column (Phenomenex, Torrance, CA, USA) was used for separation, and both compounds were detected at 366 nm (Figure 2.2). The injection volume was 10 μL.

For identification of chlorogenic acids and flavonol glycosides, the samples were also analyzed with an identical method using the Waters ACQUITY® UPLC I-Class system coupled with a Waters Vion Ion Mobility Quadrupole Time of Flight (IMS QTof) mass spectrometer (MS) (Waters Corp., Milford, MA, USA) previously described in Wang et al. (2018). Compounds were identified by LC-MS-MS based on accurate masses, retention times, and UV absorbance at 305 to 390 nm. The same column, solvent system, and elution gradient as described in this study were used with the system for compound identification. In addition, a 1:3 splitter was used to direct one-fourth of the flow (0.25 mL/min) into the MS.

Organic acids were analyzed in a Dionex® HPLC system with AS50 Autosampler, AS50 Thermal Compartment, PDA-100 Detector, and GP-40 Gradient Pump (Thermo Scientific, Waltham, MA,
USA). A Gemini® 250 x 4.6 mm 5 μm C18 110 Å LC column (Phenomenex, Torrance, CA, USA) was used for separation, and organic acids were detected at 210 to 230 nm. The injection volume was 20 μL.

All solvent systems and elution gradients are summarized in Table 2.2.

2.2.5 Compound identification with MS spectrometry of samples
IM HRMS data was acquired in high-definition MSE mode, with the following parameters: ion source, ESI negative ion; analyzer type, sensitivity; source temperature, 100 °C; desolvation temperature, 400 °C; cone gas flow, 50 L/h; desolvation gas flow, 850 L/h; capillary voltage, 2.50 kV; low collision energy, 6.0 eV; high collision energy, 15.0-45.0 eV; mass range, 50-2000 m/z; scan rate, 0.25 s. Leucine encephalin (50 pg/mL, 10 μL/min) was used for lock mass correction at 0.25 min intervals. MS and ion mobility data were acquired and processed in UNIFI (Waters Corp., Milford, MA, USA).

2.2.6 Chromatograms were used for characterization and quantification of phenolic compounds and organic acids
Chlorogenic acid, flavonol glycoside, and organic acid characterization was carried out by comparing their LC retention times, UV spectra and/or MS/MS data with standards (Table 2.3).

For quantification, chromatograms were viewed at absorbance 366 nm (phenolic compounds) and 210 nm and 214 nm (organic acids) using the photodiode array (PDA) detectors. Compounds were quantified as equivalents of their available standard, chlorogenic acid for phenolic acid analysis, or quercetin-3-galactoside for flavonol glycoside analysis (Table 2.4). The concentration
of each compound is expressed in milligrams of its equivalent external standard per gram of FW sample.

2.2.7 Statistical analysis

Statistical analyses were performed using R Studio version 4.1.1 (The R Foundation for Statistical Computing) and Microsoft® Excel® for Microsoft 365 MSO (New York, NY, USA). The lme4, lmerTest, and emmeans packages in R were used to fit a linear regression model to the data, determine whether a significant difference exists between genotypes with Satterthwaite's method, and run post-hoc analyses on differences between pairs of genotypes using Tukey multiple comparison tests with Kenward-Roger's degree of freedom method. The corrplot package in R was used to generate a Pearson’s correlation matrix among AW and phenolic concentrations. Kruskal-Wallis H test was used to evaluate harvest year and localization effects on total phenolic content. Excel was used to generate tables using mean and standard deviation values derived from the data.

2.3 Results

2.3.1 Qualitative and quantitative analysis of phenolics

Figure 2.2 shows a selected HPLC chromatogram of blueberry phenolic compounds. Table 2.3 shows the retention times and MS spectra of the 14 studied blueberry phenolic compounds, abbreviated as listed in Table 2.1. Figure 2.3 shows significant differences between average concentrations, by weight, for all phenolic compounds in fruit among two genotypes of V. corymbosum, two genotypes of V. darrowii, and two interspecific hybrids (BNJ05-218-9 and BNJ05-237-8). Differences were also observed between V. corymbosum, V. darrowii, and F1 hybrids in terms of major phenolic acids and flavonol glycosides (Table 2.5). In V. corymbosum
blueberries, ACQA1 was the major phenolic acid, whereas 3-CQA predominated in *V. darrowii*, and ACQA1 and 3-CQA were intermediate in F1 hybrid blueberries. In addition, S3GLU and S3GAL were the major flavonol glycosides in both *V. corymbosum* and *V. darrowii* blueberries, while Q3RHA predominated in F1 hybrid blueberries. Also, Q3RHA was considerably higher in *V. darrowii* (0.02 mg/g FW) compared to less than 0.01 mg/g FW in *V. corymbosum*.

Table 2.5 summarizes the mean and standard deviation values of all phenolic compounds among the genotypes. For *V. corymbosum* and *V. darrowii*, phenolic acid content varied from 0.01±0.02 mg/g FW to 0.17±0.07 mg/g FW and 0.04±0.001 mg/g FW to 1.12±0.20 mg/g FW, respectively. Flavonol glycoside content varied from <0.01±0.005 mg/g FW to 0.04±0.01 mg/g FW and <0.01±0.0004 mg/g FW to 0.04±0.01 mg/g FW, respectively. For F1 hybrids, phenolic acid content ranged from 0.06±0.006 mg/g FW to 0.27±0.06 mg/g FW and flavonol glycoside content ranged from <0.01±0.001 mg/g FW to 0.03±0.006 mg/g FW.

Additionally, differences were found between relative proportions of phenolic classes (Figure 2.5A). In *V. corymbosum*, phenolic acids were the predominant phenolic constituents (65-75% proportion of total phenolics). Phenolic acids also predominated in *V. darrowii* (90% proportion of total phenolics). In *V. darrowii* and F1 hybrids, quercetin was the major aglycone, and galactoside and glucoside were the major glycosides (Figure 2.5B, 2.5C). In *V. corymbosum*, syringetin was the major aglycone and rhamnoside was the major glycoside. In terms of composition, both *V. darrowii* and F1 hybrids contained higher amounts of quercetin, followed by syringetin and myricetin (Figure 2.5B). *V. corymbosum* had higher amounts of syringetin, followed by quercetin and myricetin.
2.3.2 Effect of harvest year on total phenolic content

The total phenolic content did not differ significantly between samples collected during two different harvest years \( (p = 0.25) \) (Table 2.7A). However, the total phenolic content in harvest year 1 was lower than the total phenolic content in harvest year 2 \( (0.66 \text{ mg/g FW average vs. 0.81 mg/g FW average}) \).

2.3.3 Distribution of total phenolic content between epidermis and flesh

The total phenolic content did significantly differ between samples extracted from the epidermis and from the flesh of blueberry fruit \( (p = 0.02) \) (Table 2.7B, Figure 2.6). Variation was observed among totals for different phenolic classes. For instance, the accumulation of total flavonol glycosides in the epidermis was greater than that of total phenolic acid accumulation.

2.3.4 Correlations among phenolics in grandparents and F1s

All correlations, either negative or positive, were classified as high if \( r > 0.7 \), moderate if \( 0.4 > r > 0.7 \), or weak if \( r < 0.4 \). According to these criteria, many phenolic compounds showed strong positive correlations \( (3\text{-CQA and CA, 3\text{-CQA and Q3GLC, ACQA1 and ACQA2, M3GAL and Q3GLU, CA and Q3GAL, CA and QCGAL, Q4GAL and S3GAL, Q3GAL and QCGAL, Q3GLU and S3GLU, Q3GLU and S3GLC, and S3GLU and S3GLC}) \ (p < 0.001) \) (Table 2.6). Other compounds correlated negatively \( (AW and 3\text{-CQA, AW and CA, AW and Q3GAL, AW and S3GAL, AW and QCGAL, 3\text{-CQA and ACQA2, 3\text{-CQA and M3GAL, 3\text{-CQA and S3GLC, ACQA1 and CA, ACQA1 and Q3GAL, ACQA1 and QCGAL, ACQA2 and CA, ACQA2 and Q3RUT, ACQA2 and Q3GAL, ACQA2 and QCGAL, M3GAL and CA, M3RHA and S3GAL, Q3RUT and Q3RHA, and Q3GLU and Q3RHA}) \ (p < 0.001). In addition, statistically significant weak correlations were observed between AW and several phenolic compounds \( \text{(ACQA2, M3GAL, M3RHA, Q3RHA, and S3GLC}) \ (p \leq 0.001).} \)
2.3.5 Qualitative and quantitative analysis of organic acids

Genotypes were evaluated for three major organic acids previously identified by Wang et al. (2019). Organic acid content and composition data indicated that there were significant differences for all organic acids among genotypes (Figure 2.7). Our results agreed with the previous report. *V. corymbosum* contained the highest levels of citric acid and possessed the least amount of quinic acid. In contrast, *V. darrowii* contained the highest levels of quinic acid and possessed the least amount of citric acid. The F1 hybrids exhibited intermediate levels of citric acid and quinic acid. Furthermore, shikimic acid was detected in all three groups, with higher levels evident in *V. darrowii* blueberries.

2.4 Discussion

In recent years, research interest in the high phenolic content of blueberry has rapidly increased due to its putative human health benefits. Most studies have focused on the three commercial species of blueberries; however, diploids are thought to be the progenitors of cultivated blueberry species and may offer valuable unique fruit chemistry traits for breeding (Bruederle and Vorsa 1994; Camp 1945). In the present study, we evaluated the CGA and flavonol glycoside profiles of two *V. corymbosum* and two *V. darrowii* wild diploid plants, as well as two F1 hybrids. We demonstrated that significant variation for these compounds does exist among the different blueberry species. In addition, the study found that the variance in phenolic content due to branch, plant clone, and harvest date was negligible. Also, the results indicated that species accounted for a large proportion of variability in organic acids, suggesting that organic acids might be useful as markers in breeding.
2.4.1 Qualitative and quantitative analysis of phenolics

A total of four phenolic acids and ten flavonol glycosides were identified and quantified in two genotypes each of diploid species V. corymbosum and V. darrowii and two F1 hybrids (Table 2.5). The phenolic acids were mainly derivatives of caffeic acid, including 3-caffeoylquinic acid, two acylated caffeoylquinic acid isomers, and caffeoylarbutin. The flavonol glycosides consisted of the three aglycones quercetin, myricetin, and syringetin (Figure 2.5B). Our results were generally consistent with the previous report of Wang et al. (2019), but we were not able to conclusively identify laricitrin, kaempferol, or isorhamnetin in our extracts.

The different blueberry genotypes varied in both phenolic concentration and composition (Figure 2.3, Table 2.5). Except for V. corymbosum, all other genotypes showed quercetin as their most abundant flavonol class, which agreed with the result of previous studies (Figure 2.5B) (Vrhovsek et al., 2012; Wang et al., 2019). V. corymbosum exhibited syringetin as its most abundant flavonol class (Figure 2.5B). Regarding this high abundance, the comparatively low amount of myricetin identified may have affected the overall composition outlook. The data set included only two myricetin glycosides at low to moderate amounts (Table 2.5). Moreover, the high amount of syringetin detected indicates strong activity of F3’5’OH and methyltransferase enzymes.

V. corymbosum blueberries exhibited a higher amount of myricetin than V. darrowii blueberries (Figure 2.5B). Furthermore, rhamnoside showed high levels of variation among the genotypes (Figure 2.5C), which suggests that flavonol rhamnosyltransferases (key enzymes for biosynthesis of rhamnosides) may be differentially expressed or more efficient in different blueberry species (Frydman et al., 2004; Saito et al., 2013). As the sugar moiety is an important determining factor
in flavonol bioavailability, information on their conjugation may be helpful in investigating flavonol nutrition. Galactoside (ranging from 27.47% to 51.33%) and glucoside (ranging from 16.96% to 32.62%) were the most abundant conjugate sugars in both *V. corymbosum* and *V. darrowii* (Figure 2.5C). In F1 hybrids, the most abundant conjugate sugars were rhamnoside (ranging from 33.08% to 33.16%) and galactoside (ranging from 27.32% to 27.93%). These findings suggest that sugar moiety differs in species.

Figure 2.3 shows that all phenolic compounds studied exhibited some significant levels of variability among the six blueberry genotypes. This observation suggests that, for flavonols, both the biosynthesis of the three aglycones and conjugation of their sugar moieties are regulated differently in species. Some of these compounds also appeared able to differentiate between genotypes. For example, 3-CQA, which exhibited high variability with an average concentration ranging from 0.09-1.12 mg/g FW (Table 2.5), could be used to discriminate between *V. corymbosum* and *V. darrowii* genotypes but not clones. Similarly, M3GAL, ranging from <0.01-0.01 mg/g FW, could also be used to differentiate between *V. corymbosum* and *V. darrowii* genotypes.

All examined blueberry samples showed high amounts of 3-CQA in agreement with data previously reported by Grace et al. (2019) (Table 2.5). Powerful antioxidant properties have been associated with CGAs, which are major soluble phenolics that accumulate to substantial levels in a wide variety of plants (Dastmalchi et al., 2011; Niggeweg et al., 2004). Relatively high concentrations of CGA have also been reported in coffee and correlated with potential health benefits (Moeenfard et al., 2014; Rodriguez-Mateos et al., 2012). Q3RHA also stood out among the studied flavonol glycosides. This compound has been reported at significant amounts in *V.*
darrowii blueberries but no quantity observed in *V. corymbosum* blueberries (Wang et al., 2019). In our study, the presence of Q3RHA was detected in *V. corymbosum* blueberries, in amounts less than 10% of the total flavonol glycoside content (Figure 2.3, Table 2.5). This inconsistent result indicates that repeat experiments should be carried out.

### 2.4.2 Effect of harvest year on total phenolic content

A difference in total phenolic content was observed between harvest years (*p* > 0.05) (Figure 2.4). Blueberries harvested later in the season exhibit increased total phenol, anthocyanin, and antioxidant levels than those fruit harvested earlier in the same season (Castrejón et al., 2008; Mallik & Hamilton, 2017; Prior et al., 1998). During harvest year 2, we began harvesting of fruit later in the season compared to the previous year (harvest year 1). In addition, a greater quantity of fruit was available for extractions in 2020. A higher proportion of harvest year 2 (2020) fruit than harvest year 1 (2019) fruit in the data set may have contributed to the observed difference in total phenolic content. Furthermore, a sufficiently higher degree of UV radiation registered during the growth period of harvest year 2 compared to harvest year 1 may also have affected the phenolic content (Pertuzatti et al., 2021; Surjadinata et al., 2017).

### 2.4.3 Distribution of total phenolic content between epidermis and flesh

We observed that the distribution of total phenolic content was significantly greater in the epidermis of blueberry fruit compared to its flesh (Figure 2.6, Table 2.7B). It has been reported that phenolic compounds are concentrated in the epidermal tissues of blueberries, meaning that small berries contain proportionally more phenolic compounds than large berries, when based on per weight of whole berries (Connor et al., 2002b; Howard et al., 2003; Kim et al.,
More specifically, smaller berries with a larger surface area to volume ratio tend to contain higher quantities of these compounds compared to larger berries (Prior et al., 1998).

It has been shown that fruit volume and fruit weight are highly correlated, and that fruit weight can be used as a proxy for estimating fruit size (Mengist et al., 2020). In our study, AW was included in the analyses as a covariate. Differences in significance results were observed between analyses run with the covariate and without the covariate for 3-CQA, ACQ1, ACQA2, M3RHA, Q3RUT, S3GAL, and QCGAL. Among these compounds, 3-CQA, Q3RUT, S3GAL, and QCGAL exhibited moderately strong, negative correlations with AW (Table 2.6). The results support that at least some phenolic compounds are specifically located in the epidermis of blueberries. Therefore, it is important to use a covariate to adjust for fruit size when analyzing phenolic content with whole fruit.

2.4.4 Correlations among phenolics in grandparents and F1s

The relationships between the studied phenolic compounds are summarized in Table 2.6. Overall, moderate to strong correlations were observed between quercetin flavonols ($p < 0.001$), while weak correlations were observed between myricetin and syringetin flavonols. Strong correlations were also evident between glucosides but not rhamnosides or galactosides ($p < 0.001$). Phenolic acids also exhibited moderate to strong correlations, especially between the two acetyl-caffeoylquinic acid isomers ($p < 0.001$). Only 3-CQA, ACQA1, CA, Q3GAL, S3GAL, and QCGAL exhibited moderate to strong correlations to the AW covariate, despite findings that phenolic compounds are differentially localized in the fruit epidermis. This result suggests that our AW variable may not be the most effective covariate to use when adjusting for fruit size. Further studies may benefit from utilizing a standardized method of assigning fruit to different
size categories, and subsequently, deriving fruit weight calculations from those recorded measurements.

On the other hand, this lack of relationship between fruit size and phenolic content correlates to a previous study by Kalt et al. (2001). The authors hypothesized that the lack of any relationship may reflect differences in the evolutionary histories of the different blueberry species. Furthermore, it has been observed that blueberry varieties have different numbers of layers with various-sized pigment-containing cells (Allan-Wojtas et al., 2001). This data suggests that the distribution of compounds between the epidermal and subepidermal layers in our fruit might have also been a factor.

In the flavonol biosynthesis pathway, syringetin is produced from myricetin via two sequential methylations carried out by \(O\)-methyltransferases (Cacace et al., 2003). Yet M3RHA showed a moderate, negative correlation with S3GAL and weak correlations with S3GLU and S3GLC. M3GAL showed a weak correlation with S3GAL and moderate, positive correlations with S3GLU and S3GLC. These results support a divergence in the metabolic pathway of synthesizing syringetin from myricetin, which agrees with evidence in the literature (Cacace et al., 2003). Furthermore, the significant correlations found among multiple compounds suggests the influence of an ancestral conserved genetic control on these traits.

2.4.5 Qualitative and quantitative analysis of organic acids

Citric, quinic, and shikimic acids were previously identified as major organic acids in blueberry by Wang et al. (2019). The organic acid compositions observed among the genotypes in this study were consistent with previous results; most notably, the citric versus quinic acid difference in
the two species. Figure 2.7 shows that citric acid was significantly low in *V. darrowii*, whereas quinic acid was significantly low in *V. corymbosum*. Because citric acid and quinic acid can be used to discriminate between different species, these compounds may prove important in the study of organic acid biosynthesis regulation.

### 2.5 Conclusions

The current study evaluated the CGA, flavonol glycoside, and organic acid profiles of two diploid *V. corymbosum*, two diploid *V. darrowii*, and two F1 hybrid blueberries. Four phenolic acids, ten flavonol glycosides, and three organic acids were identified and quantified. Results revealed that different blueberry species exhibited significant variation in the content and composition of some of the studied compounds. It was also found that phenolic compound distribution was much greater in the epidermis of fruit than in the flesh of fruit. This result was consistent with evidence indicating that smaller berries contain proportionally greater phenolic content than larger berries, when based on per weight of whole berries, due to a larger surface area to volume ratio (Connor et al., 2002b; Howard et al., 2003; Kim et al., 2013; Prior et al., 1998).

It is worth noting that to express fruit flavor, acid content (e.g., citric acid and malic acid) in most fruit crops must be at levels that give a titratable acidity (TA) of about 0.5% (Fong et al., 2020). Despite both tetraploid (Ehlenfeldt et al., 1994) and diploid (Baloga et al., 1995) *V. corymbosum* having a TA of about 0.5%, only the tetraploid species was domesticated in having citric acid.

Moreover, it must be mentioned that most, if not all, of the bioactive compounds found in fruits and vegetables impart bitterness, acridity, or astringency (Drewnowski & Gomez-Carneros, 2000). Consequently, responding to taste-driven consumer demand, the food industry removes
phenolic compounds from plant foods. Thus, while studies suggest that enhancing phytonutrient content through breeding or genetic improvement is a potent option to lower the risk of various diseases, care must be taken not to produce a variety that is sensorially objectionable to the consumer.

The differential phenolic profiles seen in this study may provide valuable information in blueberry genetics for fruit chemistry and phytochemical improvement. According to our results, *V. darrowii* could offer an opportunity for the improvement of antioxidant activity in domesticated highbush *V. corymbosum*. *V. darrowii* was found to contain high levels of 3-CQA, a phenolic acid that, together with anthocyanins, likely accounts for a large proportion of the antioxidant activity in blueberries. Notably, 3-CQA does not impart bitterness (Nagel et al., 1987), thereby allowing it to be enhanced to improve human health without having a negative impact on taste.
2.6 References


2.7 Figures and tables

Figure 2.1 Pedigree chart of blueberry plants.

Figure 2.2 HPLC chromatogram of blueberry phenolic compounds. All peaks were detected by a photodiode array detector at 366 nm wavelength.
Figure 2.3 Bar chart of linear mixed effects model estimated average concentrations of phenolic compounds for six wild diploid blueberry genotypes (mg/g FW). A 3-caffeoylquinic acid, B acetyl-caffeoylquinic acid isomer 1, C acetyl-caffeoylquinic acid isomer 2, D myricetin-3-galactoside, E caffeoylarbutin, F myricetin-3-rhamnoside, G quercetin-3-rutinoside, H quercetin-3-galactoside, I quercetin-3-glucoside, J quercetin-3-rhamnoside, K syringetin-3-galactoside, L syringetin-3-glucoside, M syringetin-3-glucuronide, and N quercetin-3-(6"'-caffeyl)-galactoside. Different letters above bars indicate significant difference ($p < 0.05$) among genotypes.
Figure 2.4 Effect of harvest year on concentrations of total phenolic compounds (mg/g FW).
Figure 2.5 Average composition of different compound groups. A phenolic compound groups, B flavonol groups based on aglycone structure, and C flavonol groups based on glycosylation type.

PHEN = phenolic acids, FLAV = flavonol glycosides, TOTAL = total phenolics, QUER = quercetin, MYRI = myricetin, SYRIN = syringetin, GAL = galactoside, RHA = rhamnose, RUT = rutinoside, GLU = glucoside, and GLC = glucuronide.
Figure 2.6 Effect of localization in epidermis or flesh of blueberry fruit on concentrations of total phenolic compounds (mg/g FW). PHEN = phenolic acids, FLAV = flavonol glycosides, and TOTAL = total phenolics.

Figure 2.7 Average composition of different organic acids in genotypes.
Table 2.1 List of abbreviations for all studied compounds.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Compound name</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-CQA</td>
<td>3-caffeoylquinic acid</td>
</tr>
<tr>
<td>ACQA1</td>
<td>Acetyl-caffeoylquinic acid isomer 1</td>
</tr>
<tr>
<td>ACQA2</td>
<td>Acetyl-caffeoylquinic acid isomer 2</td>
</tr>
<tr>
<td>M3GAL</td>
<td>Myricetin-3-galactoside</td>
</tr>
<tr>
<td>CA</td>
<td>Caffeoylarbutin</td>
</tr>
<tr>
<td>M3RHA</td>
<td>Myricetin-3-rhamnoside</td>
</tr>
<tr>
<td>Q3RUT</td>
<td>Quercetin-3-rutinoside</td>
</tr>
<tr>
<td>Q3GAL</td>
<td>Quercetin-3-galactoside</td>
</tr>
<tr>
<td>Q3GLU</td>
<td>Quercetin-3-glucoside</td>
</tr>
<tr>
<td>Q3RHA</td>
<td>Quercetin-3-rhamnoside</td>
</tr>
<tr>
<td>S3GAL</td>
<td>Syringetin-3-galactoside</td>
</tr>
<tr>
<td>S3GLU</td>
<td>Syringetin-3-glucoside</td>
</tr>
<tr>
<td>S3GLC</td>
<td>Syringetin-3-glucuronide</td>
</tr>
<tr>
<td>QCGAL</td>
<td>Quercetin-3-(6''-caffeyl)-galactoside</td>
</tr>
</tbody>
</table>

Table 2.2 Summary of chemical analysis methods for chlorogenic acids, flavonol glycosides, and organic acids.

<table>
<thead>
<tr>
<th>Phenolic compounds</th>
<th>Organic acids</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time (min)</td>
<td>Flow (ml/min)</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>30</td>
<td>1</td>
</tr>
<tr>
<td>35</td>
<td>1</td>
</tr>
<tr>
<td>38</td>
<td>1</td>
</tr>
<tr>
<td>40</td>
<td>1</td>
</tr>
<tr>
<td>43</td>
<td>1</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
</tr>
<tr>
<td>80</td>
<td>0</td>
</tr>
</tbody>
</table>

Solvent A: 0.1% formic acid in water
Solvent B: 0.5% phosphoric acid in water
Table 2.3 Liquid chromatography retention time, [M-H]-, and fragmentation time of blueberry phenolic compounds.

<table>
<thead>
<tr>
<th>Peak Number</th>
<th>Retention time (min)</th>
<th>[M-H] (m/z)</th>
<th>Main fragment ions in ESI-MS-MS (m/z)</th>
<th>Formula</th>
<th>ppm</th>
<th>Tentative peak identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.657</td>
<td>353.0873</td>
<td>191.0556(100)</td>
<td>C16H18O9</td>
<td>-1.3</td>
<td>Chlorogenic acid (3-Caffeoylquinic acid)</td>
</tr>
<tr>
<td>2</td>
<td>7.146</td>
<td>355.1029</td>
<td>160.0159(100), 170.0395(58)</td>
<td>C16H20O9</td>
<td>-1.6</td>
<td>Ferulic acid hexoside</td>
</tr>
<tr>
<td>4</td>
<td>7.822</td>
<td>495.1141</td>
<td>300.0264(57), 315.0502(100)</td>
<td>C22H24O13</td>
<td>-0.6</td>
<td>Methyl-epicatecholiglucoside</td>
</tr>
<tr>
<td>5</td>
<td>10.01</td>
<td>395.0980</td>
<td>191.0556(100), 233.0662(75)</td>
<td>C18H20O10</td>
<td>-1.0</td>
<td>Acetyl-caffeoylquinic acid</td>
</tr>
<tr>
<td>6</td>
<td>11.437</td>
<td>395.0979</td>
<td>233.0662(100)</td>
<td>C18H20O10</td>
<td>-1.2</td>
<td>Acetyl-caffeoylquinic acid</td>
</tr>
<tr>
<td>7</td>
<td>12.083</td>
<td>479.0828</td>
<td>316.0218(100)</td>
<td>C21H20O13</td>
<td>-0.6</td>
<td>Myricetin-3-galactoside</td>
</tr>
<tr>
<td>8</td>
<td>15.005</td>
<td>433.1199</td>
<td>161.0236(100), 323.0763 (9)</td>
<td>C21H22O10</td>
<td>-0.3</td>
<td>Caffeoylarbutin</td>
</tr>
<tr>
<td>9</td>
<td>17.575</td>
<td>609.1455</td>
<td>301.0270(100)</td>
<td>C27H30O16</td>
<td>-1.0</td>
<td>Quercetin-3-rutinoside</td>
</tr>
<tr>
<td>10</td>
<td>18.456</td>
<td>463.0880</td>
<td>271.0241(41), 300.0269(100)</td>
<td>C21H22O12</td>
<td>-0.4</td>
<td>Quercetin-3-galactoside</td>
</tr>
<tr>
<td>11a</td>
<td>19.692</td>
<td>463.0880</td>
<td>271.0241(51), 300.0279(100)</td>
<td>C21H22O12</td>
<td>-0.4</td>
<td>Quercetin-3-glucoside</td>
</tr>
<tr>
<td>12a</td>
<td>27.458</td>
<td>447.0929</td>
<td>301.0345(100)</td>
<td>C21H22O11</td>
<td>-0.9</td>
<td>Quercetin-3-rhamnoside</td>
</tr>
<tr>
<td>13</td>
<td>28.761</td>
<td>507.1142</td>
<td>344.0534(100)</td>
<td>C23H24O13</td>
<td>-0.4</td>
<td>Syringetin-3-galactoside</td>
</tr>
<tr>
<td>14</td>
<td>30.532</td>
<td>507.1141</td>
<td>344.0536(100)</td>
<td>C23H24O13</td>
<td>-0.4</td>
<td>Syringetin-3-glucoside</td>
</tr>
<tr>
<td>15</td>
<td>34.777</td>
<td>475.1244</td>
<td>NA</td>
<td>C23H24O11</td>
<td>-0.4</td>
<td>Unknown (non-flavonol)</td>
</tr>
<tr>
<td>16</td>
<td>35.894</td>
<td>521.0937</td>
<td>345.0613(100)</td>
<td>C23H22O14</td>
<td>0.0</td>
<td>Syringetin-3-glucuronide</td>
</tr>
<tr>
<td>17</td>
<td>36.442</td>
<td>625.1196</td>
<td>301.0271(100), 301.0343(69), 463.0885(100)</td>
<td>C30H26O15</td>
<td>-0.5</td>
<td>Quercetin-3-(6''-cafeoyl)-galactoside</td>
</tr>
</tbody>
</table>

Table 2.4 Calibration curves used in the quantification of phenolic compounds and organic acids.

<table>
<thead>
<tr>
<th>Standard compound</th>
<th>Calibration range (mg/mL)</th>
<th>Calibration curve</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chlorogenic acid</td>
<td>1.46E-05 - 5.71</td>
<td>$y = 2E+06x - 47023$</td>
<td>0.9997</td>
</tr>
<tr>
<td>Myricetin-3-galactoside</td>
<td>1.83E-06 - 0.71</td>
<td>$y = 2E+07x - 59267$</td>
<td>0.9994</td>
</tr>
<tr>
<td>Myricetin-3-rhamnoside</td>
<td>1.83E-06 - 0.71</td>
<td>$y = 1E+07x - 45617$</td>
<td>0.9990</td>
</tr>
<tr>
<td>Quercetin-3-rutinoside</td>
<td>1.83E-06 - 0.71</td>
<td>$y = 2E+07x - 42499$</td>
<td>0.9995</td>
</tr>
<tr>
<td>Quercetin-3-galactoside</td>
<td>1.83E-06 - 0.71</td>
<td>$y = 2E+07x - 56857$</td>
<td>0.9995</td>
</tr>
<tr>
<td>Quercetin-3-glucoside</td>
<td>1.83E-06 - 0.71</td>
<td>$y = 2E+07x - 52842$</td>
<td>0.9995</td>
</tr>
<tr>
<td>Quercetin-3-rhamnoside</td>
<td>1.83E-06 - 0.71</td>
<td>$y = 2E+07x - 53432$</td>
<td>0.9993</td>
</tr>
<tr>
<td>Quinic acid</td>
<td>6.51E-03 - 6.67</td>
<td>$y = 16.537x - 0.3447$</td>
<td>0.9989</td>
</tr>
<tr>
<td>Citric acid</td>
<td>6.51E-03 - 6.67</td>
<td>$y = 34.389x - 0.1498$</td>
<td>0.9999</td>
</tr>
<tr>
<td>Shikimic acid</td>
<td>6.51E-05 - 0.07</td>
<td>$y = 1760.5x - 0.2854$</td>
<td>0.9996</td>
</tr>
</tbody>
</table>
Table 2.5 Mean and standard deviation values of all phenolic compounds among genotypes (mg/g FW). 3-CQA = 3-caffeoylquinic acid, ACQA1 = acetyl-caffeoylquinic acid isomer 1, ACQA2 = acetyl-caffeoylquinic acid isomer 2, M3GAL = myricetin-3-galactoside, CA = caffeoylarbutin, M3RHA = myricetin-3-rhamnoside, Q3RUT = quercetin-3-rutinoside, Q3GAL = quercetin-3-galactoside, Q3GLU = quercetin-3-glucoside, Q3RHA = quercetin-3-rhamnoside, S3GAL = syringetin-3-galactoside, S3GLU = syringetin-3-glucoside, S3GLC = syringetin-3-glucuronide, and QCGAL = quercetin-3-(6''-caffeyl)-galactoside.

<table>
<thead>
<tr>
<th>V. corymbosum</th>
<th>F1 hybrids</th>
<th>V. darrowii</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean ± SD</td>
<td>Mean ± SD</td>
<td>Mean ± SD</td>
</tr>
<tr>
<td>3-CQA</td>
<td>0.09 ± 0.031</td>
<td>0.13 ± 0.055</td>
</tr>
<tr>
<td>ACQA1</td>
<td>0.11 ± 0.038</td>
<td>0.17 ± 0.070</td>
</tr>
<tr>
<td>ACQA2</td>
<td>0.07 ± 0.017</td>
<td>0.10 ± 0.037</td>
</tr>
<tr>
<td>M3GAL</td>
<td>0.01 ± 0.005</td>
<td>0.01 ± 0.003</td>
</tr>
<tr>
<td>CA</td>
<td>0.01 ± 0.018</td>
<td>0.02 ± 0.020</td>
</tr>
<tr>
<td>M3RHA</td>
<td>&lt;0.01 ± 0.001</td>
<td>&lt;0.01 ± 0.005</td>
</tr>
<tr>
<td>Q3RUT</td>
<td>0.02 ± 0.011</td>
<td>0.01 ± 0.003</td>
</tr>
<tr>
<td>Q3GAL</td>
<td>&lt;0.01 ± 0.001</td>
<td>0.01 ± 0.001</td>
</tr>
<tr>
<td>Q3GLU</td>
<td>0.02 ± 0.005</td>
<td>0.02 ± 0.004</td>
</tr>
<tr>
<td>Q3RHA</td>
<td>&lt;0.01 ± 0.003</td>
<td>&lt;0.01 ± &lt;0.001</td>
</tr>
<tr>
<td>S3GAL</td>
<td>0.02 ± 0.008</td>
<td>0.04 ± 0.012</td>
</tr>
<tr>
<td>S3GLC</td>
<td>0.03 ± 0.011</td>
<td>0.01 ± 0.004</td>
</tr>
<tr>
<td>S3GLC</td>
<td>0.02 ± 0.005</td>
<td>0.01 ± 0.002</td>
</tr>
<tr>
<td>QCGAL</td>
<td>&lt;0.01 ± &lt;0.001</td>
<td>&lt;0.01 ± &lt;0.001</td>
</tr>
</tbody>
</table>
Table 2.6 Pearson’s correlation matrix with coefficients (r) and p-values of phenolic compounds and average fruit weight in six wild diploid blueberry genotypes (N = 208). High if $r > 0.7$ (blue), moderate if $0.4 > r > 0.7$ (red), or weak if $r < 0.4$ (black). AVG.WT = average fruit weight, 3-CQA = 3-caffeoylquinic acid, ACQA1 = acetyl-caffeoylquinic acid isomer 1, ACQA2 = acetyl-caffeoylquinic acid isomer 2, M3GAL = myricetin-3-galactoside, CA = caffeoylarbutin, M3RHA = myricetin-3-rhamnoside, Q3RUT = quercetin-3-rutinoside, Q3GAL = quercetin-3-galactoside, Q3GLU = quercetin-3-glucoside, Q3RHA = quercetin-3-rhamnoside, S3GAL = syringetin-3-galactoside, S3GLU = syringetin-3-glucoside, S3GLC = syringetin-3-glucuronide, and QCGAL = quercetin-3-(6''-caffeyl)-galactoside.

| Phenolic Compounds | AVG.WT | 3-CQA | ACQA1 | ACQA2 | M3GAL | CA | M3RHA | Q3RUT | Q3GAL | Q3GLU | Q3RHA | S3GAL | S3GLU | S3GLC | QCGAL |
|--------------------|--------|-------|-------|-------|-------|----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| Syringetin-3-Galactoside |       |       |       |       |       |    |       |       |       |       |       |       |       |       |       |
| Quercetin-3-Glucoside |       |       |       |       |       |    |       |       |       |       |       |       |       |       |       |
| Quercetin-3-Rhamnoside |       |       |       |       |       |    |       |       |       |       |       |       |       |       |       |
| Myricetin-3-Galactoside |       |       |       |       |       |    |       |       |       |       |       |       |       |       |       |
| Caffeoylquinic Acid |       |       |       |       |       |    |       |       |       |       |       |       |       |       |       |
| Acetyl-Caffeoylquinic Acid |       |       |       |       |       |    |       |       |       |       |       |       |       |       |       |


<table>
<thead>
<tr>
<th></th>
<th>AVG.WT</th>
<th>3-CQA</th>
<th>ACQA1</th>
<th>ACQA2</th>
<th>M3GAL</th>
<th>CA</th>
<th>M3RHA</th>
<th>Q3RUT</th>
<th>Q3GAL</th>
<th>Q3GLU</th>
<th>Q3RHA</th>
<th>S3GAL</th>
<th>S3GLU</th>
<th>S3GLC</th>
<th>QCGAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVG.WT</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-CQA</td>
<td>-0.64</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACQA1</td>
<td>0.42</td>
<td>-0.60</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACQA2</td>
<td>0.38</td>
<td>-0.56</td>
<td>0.97</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M3GAL</td>
<td>0.30</td>
<td>-0.43</td>
<td>0.31</td>
<td>0.26</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td>-0.52</td>
<td>0.79</td>
<td>-0.46</td>
<td>-0.44</td>
<td>-0.46</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M3RHA</td>
<td>0.38</td>
<td>-0.30</td>
<td>0.51</td>
<td>0.45</td>
<td>0.18</td>
<td>-0.15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q3RUT</td>
<td>-0.09</td>
<td>0.31</td>
<td>-0.39</td>
<td>-0.41</td>
<td>0.41</td>
<td>0.03</td>
<td>-0.36</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q3GAL</td>
<td>-0.47</td>
<td>0.60</td>
<td>-0.42</td>
<td>-0.42</td>
<td>-0.29</td>
<td>0.89</td>
<td>-0.25</td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q3GLU</td>
<td>0.12</td>
<td>-0.34</td>
<td>-0.08</td>
<td>-0.12</td>
<td>0.73</td>
<td>-0.39</td>
<td>-0.13</td>
<td>0.54</td>
<td>-0.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q3RHA</td>
<td>-0.22</td>
<td>0.20</td>
<td>0.13</td>
<td>0.19</td>
<td>-0.27</td>
<td>0.48</td>
<td>0.33</td>
<td>-0.44</td>
<td>0.39</td>
<td>-0.49</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3GAL</td>
<td>-0.54</td>
<td>0.51</td>
<td>-0.31</td>
<td>-0.30</td>
<td>-0.05</td>
<td>0.61</td>
<td>-0.44</td>
<td>0.18</td>
<td>0.79</td>
<td>0</td>
<td>0.08</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3GLU</td>
<td>0.04</td>
<td>-0.22</td>
<td>-0.13</td>
<td>-0.18</td>
<td>0.44</td>
<td>0.07</td>
<td>0.03</td>
<td>0.25</td>
<td>0.10</td>
<td>0.75</td>
<td>-0.17</td>
<td>-0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3GLC</td>
<td>0.29</td>
<td>-0.43</td>
<td>-0.01</td>
<td>-0.06</td>
<td>0.44</td>
<td>-0.33</td>
<td>0.03</td>
<td>0.27</td>
<td>-0.13</td>
<td>0.78</td>
<td>-0.40</td>
<td>-0.19</td>
<td>0.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>QCGAL</td>
<td>-0.58</td>
<td>0.82</td>
<td>-0.58</td>
<td>-0.56</td>
<td>-0.43</td>
<td>0.96</td>
<td>-0.29</td>
<td>0.15</td>
<td>0.90</td>
<td>-0.28</td>
<td>0.39</td>
<td>0.69</td>
<td>-0.04</td>
<td>-0.29</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 2.7 Kruskal-Wallis rank sum test results of total phenolic contents. A harvest year and B localization in epidermis or flesh of blueberry fruit.

<table>
<thead>
<tr>
<th>A</th>
<th>Kruskal-Wallis chi-squared</th>
<th>df</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>harvest.year</td>
<td>1.32</td>
<td>1</td>
<td>0.25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B</th>
<th>Kruskal-Wallis chi-squared</th>
<th>df</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>total phenolic acids</td>
<td>5.33</td>
<td>1</td>
<td>0.02</td>
</tr>
<tr>
<td>total flavonol glycosides</td>
<td>5.33</td>
<td>1</td>
<td>0.02</td>
</tr>
<tr>
<td>total phenolics</td>
<td>5.33</td>
<td>1</td>
<td>0.02</td>
</tr>
</tbody>
</table>
CHAPTER 3

Phenotypic evaluation of a segregating (Vaccinium corymbosum x V. darrowii) diploid F2 blueberry population for phenolic compounds

3.1 Introduction

Blueberry is a woody perennial shrub that belongs to the genus *Vaccinium* (Syn. Health) in the family Ericaceae (Vander Kloet, 1988). Most commercial production of blueberry comes from species in the subsection *Cyanococcus*, which are native to and occur naturally in eastern and northcentral North America (Hancock & Draper, 1989; Song & Hancock, 2011). Blueberries are the most popular berry crop in *Vaccinium* (Song & Hancock, 2011), and in 2020, had an annual U.S. production reaching 324,000 tons (Kramer et al., 2020). Over the past decade, annual growth for cultivated blueberry production in the fresh market has averaged 5%, climbing from 246.4 million pounds in 2010 to 373 million pounds in 2019 (Kramer et al., 2020). This growth was likely because of an increasing general awareness of blueberry’s high antioxidant capacity and its many health benefits (Connor et al., 2002b, 2002a; Ehlenfeldt & Prior, 2001; Prior et al., 1998; Seeram, 2006). Regular, moderate intake of blueberries has also been shown to reduce risk of cardiovascular disease (Vendrame et al., 2013), obesity (Jiao et al., 2019), type 2 diabetes (Stote et al., 2017), and neurodegenerative diseases (Kalt et al., 2020).

Blueberries are comprised of approximately 150 to more than 400 species worldwide (Finn et al., 1990; Vander Kloet, 1988). However, almost all commercial blueberries are harvested from one of only several species: tetraploid highbush *V. corymbosum* L., hexaploid rabbiteye *V. ashei* Reade, and tetraploid lowbush *V. angustifolium* and *V. myrtilloides* (Retamales & Hancock, 2012; Song & Hancock, 2011). Among them, tetraploid highbush blueberry, *V. corymbosum* L., is
considered the most important commercial crop in the *Vaccinium* genus (Hancock, 1995; Song & Hancock, 2011). Highbush blueberry can be classified into northern and southern types depending on their cold hardiness and chilling requirements (Song & Hancock, 2011). Southern types were developed by incorporating low-chilling diploid *V. darrowii* genes into the northern tetraploid *V. corymbosum* background (Hancock et al., 2008; Rowland et al., 2014).

Although blueberry breeding has made great progress since its domestication at the end of the 19th century (Song & Hancock, 2011), the process of developing new cultivars can still take more than ten years (Hancock et al., 2008). Therefore, the development of germplasm resources could help in breeding new blueberry cultivars, and in future marker-assisted selections, to meet rising consumer demand. For this reason, an interspecific F2 diploid population suitable for mapping multiple plant and fruit quality traits was developed in 2016. This population was generated by first crossing *V. corymbosum* and *V. darrowii* to develop F1 hybrids. Then, F1 x F1 crosses (diploid blueberry is obligatory outcrossing) were made to create a biparental pseudo F2 population that showed a wide distribution of traits. A true F2 population would be difficult, if not impossible, to create due to self-sterility in diploid blueberry.

In the present study, we phenotyped the F2 population for two major classes of blueberry phenolic compounds – chlorogenic acids (CGAs) and flavonol glycosides – and for average fruit weight (AW). The objectives of this study, herein, were to score the phenotypes and describe their distribution in the segregating population mentioned above. HPLC methods were optimized to identify individual blueberry phenolics based on Wang et al. (2019) and provide accurate quantitative data. Four phenolic acids and ten flavonol glycosides were analyzed.
Possible modes of genetic inheritance of the constituents have been proposed using frequency
distribution data and chi-square tests.

3.2 Materials and methods

3.2.1 Plant materials

Four species (two V. corymbosum and two V. darrowii genotypes) and two F1 hybrids were used
in this study. OPB-15 and OPB-8 are wild diploid V. corymbosum plants that were collected from
a native population in Burlington County, NJ. NJ88-12-41 and NJ88-14-03 are wild diploid V.
darrowii plants that were collected from native populations in Liberty County, Florida and along
Route 98S in the Saint Joseph Bay area of Florida, respectively. In 2005, crosses were made
between NJ88-14-03 and OPB-15 and between OPB-8 and NJ88-12-41. The resulting F1 hybrid
plants were BNJ05-218-9 and BNJ05-237-8 (Figure 2.1). In 2016, crosses were made between
the BNJ05-218-9 and BNJ05-237-8 hybrid plants, creating the F2 population used in this study
with an inbreeding coefficient of F = 0 (Figure 1.1). All plants were maintained in potted culture
in a greenhouse through seasonal plant growth and winter dormancy (greenhouse maintained
in a “cold” state, allowing for winter chilling at minimum 0-4 °C). Bumblebees (Koppert
Biological Systems, MI, USA) were brought into the greenhouse in late spring/early summer
during flowering for open pollination for fruit set. Fully ripe, i.e., blue, blueberry fruit samples
were harvested from each plant at 7-14-day intervals over the fruiting period, initiating at first
harvest in April through last harvest in August, at the P.E. Marucci Center for Blueberry and
Cranberry Research and Extension in 2019 to 2021. Samples were separately placed in
polyethylene bags and kept chilled until weight measurements were taken. After weighing,
samples were stored at -80 °C until analysis.
3.2.2 Chemicals and reagents

All solvents, including water, acetonitrile, methanol, and acetone, were purchased from EMD Millipore (Billercia, MA, USA) and were of HPLC grade. Acetic acid was purchased from Avantor Performance Materials (Center Valley, PA, USA) and formic acid was purchased from Mallinckrodt Baker (Phillipsburg, NJ, USA). Commercial standards of flavonol glycosides (quercetin-3-galactoside, quercetin-3-glucoside, quercetin-3-rutinoside, quercetin-3-rhamnoside, myricetin-3-galactoside, and myricetin-3-rhamnoside) were obtained from Indofine Chemical Company (Somerville, NJ, USA), and chlorogenic acid standard was obtained from Sigma-Aldrich, Inc. (St. Louis, MO, USA).

3.2.3 Extraction of blueberry phenolic compounds

For chlorogenic acid and flavonol glycoside quantification, depending on sample availability, 2-8 g of fruits were weighed, average berry weight (AW) was recorded, and samples were ground with a Precellys Evolution homogenizer (Bertin Corp., Rockville, MD, USA) using 2.8 mm ceramic beads at 7200 rpm for 1.5 min. 80% aqueous acetone with 0.1% acetic acid (1:4 sample to solvent w/v) was added to suspend ground fruit, and samples were extracted overnight at 4 °C. Liquid extracts were then centrifuged at 13,300 rpm for 2 min, and 1 mL aliquots of clear supernatant were taken. Aliquots were dried with a SpeedVac vacuum concentrator (Savant®, SPD2010-220, Thermo Scientific, Waltham, MA, USA) under no heat and redissolved in 500 μL of 100% methanol by sonication for 15 min. Samples were then centrifuged at 11,000 rpm for 5 min, and clear supernatants were analyzed with high-performance liquid chromatography (HPLC).
3.2.4 HPLC apparatus and conditions

Two HPLC systems were used for chlorogenic acid and flavonol glycoside identification and quantification:

Chlorogenic acids and flavonol glycosides were analyzed in a Waters Alliance® LC system composed of a Waters e2695 Separations Module and Waters 2998 PDA Detector (Waters Corp., Milford, MA, USA). A Gemini® 150 x 4.6 mm 5 μm C18 110 Å LC column (Phenomenex, Torrance, CA, USA) was used for separation, and both compounds were detected at 366 nm (Figure 2.2). The injection volume was 10 μL.

For identification of chlorogenic acids and flavonol glycosides, the samples were also analyzed with an identical method using the Waters ACQUITY® UPLC I-Class system coupled with a Waters Vion Ion Mobility Quadrupole Time of Flight (IMS QTof) mass spectrometer (MS) (Waters Corp., Milford, MA, USA) previously described in Wang et al. (2018). Compounds were identified by LC-MS-MS based on accurate masses, retention times, and UV absorbance at 305 to 390 nm. The same column, solvent system, and elution gradient as described in this study were used with the system for compound identification. In addition, a 1:3 splitter was used to direct one-fourth of the flow (0.25 mL/min) into the MS.

All solvent systems and elution gradients are summarized in Table 2.2

3.2.5 Compound identification with MS spectrometry of samples

IM HRMS data was acquired in high-definition MSE mode, with the following parameters: ion source, ESI negative ion; analyzer type, sensitivity; source temperature, 100 °C; desolvation
temperature, 400 °C; cone gas flow, 50 L/h; desolvation gas flow, 850 L/h; capillary voltage, 2.50 kV; low collision energy, 6.0 eV; high collision energy, 15.0-45.0 eV; mass range, 50-2000 m/z; scan rate, 0.25 s. Leucine encephalin (50 pg/mL, 10 μL/min) was used for lock mass correction at 0.25 min intervals. MS and ion mobility data were acquired and processed in UNIFI (Waters Corp., Milford, MA, USA).

3.2.6 Chromatograms were used for characterization and quantification of phenolic compounds

Chlorogenic acid and flavonol glycoside characterization was carried out by comparing their LC retention times, UV spectra and/or MS/MS data with standards (Table 2.3).

For quantification, chromatograms were viewed at absorbance 366 nm using the photodiode array (PDA) detector. Compounds were quantified as equivalents of their available standard, chlorogenic acid for phenolic acid analysis, or quercetin-3-galactoside for flavonol glycoside analysis (Table 2.4). The concentration of each compound is expressed in milligrams of its equivalent external standard per gram of FW sample.

3.2.7 Statistical analysis

Statistical analyses were performed using R Studio version 4.1.1 (The R Foundation for Statistical Computing) and Microsoft® Excel® for Microsoft 365 MSO (New York, NY, USA). Excel was used to generate frequency distribution histograms and tables using F2 data. Chi-square goodness of fit test was used to evaluate segregation of each phenolic compound. The factoextra package in R was used to apply principal component analysis to the F2 data to highlight similarities and differences in phenolic compounds between individuals in the F2 population.
3.3 Results

3.3.1 Segregation of phenolic traits in the F2 population

The phenolic traits used to score individuals in the F2 population were identified from previous data on average content of phenolic compounds among the four wild diploid species and two F1 hybrids (Table 3.2). After characterizing the parent species phenotypes, the F2 progeny were counted for each trait and summed according to pre-established bins in frequency distribution tables (Table 3.1).

Figure 3.1 shows distribution data for all the phenolic compound traits as frequency distribution histograms. A largely continuous distribution for each compound value was observed for the F2 population. Most phenolic compounds segregated in an approximately normal distribution, whereas ACQA1 and S3GAL segregated in a bimodal distribution (Figure 3.1B, K). In addition, the distributions for ACQA1, ACQA2, Q3RUT, Q3GLU, Q3RHA, S3GAL, and QCGAL did not span that of the species range as might be expected in the F2 population (Figure 3.1B, C, G, I, J, K, N). Data for some compounds also included outlier data, which skewed distribution curves towards the right and created long tails on the histograms. 3-CQA showed the broadest range of concentration (0.05-1.26 mg/g FW) and M3GAL showed the narrowest range (<0.01-0.01 mg/g FW).

Phenolic traits were scored in F2 progeny based on parent species phenotypes (Table 3.1, 3.2). Specifically, these phenotypes constituted concentration values greater than or less than a certain value (e.g., 3-CQA: *V. corymbosum* trait if < 0.13 mg/g FW and *V. darrowii* trait if > 0.13 mg/g FW). Expected counts of F2 progeny were calculated by multiplying the total number of F2
individuals by the expected Mendelian phenotypic ratio (3:1). The segregation of each trait was tested with a chi-square ($X^2$) test for goodness of fit, with the formula $X^2 = (O-E)^2/E$, where $O$ represents the observed progeny and $E$ represents the expected progeny. The null hypothesis was that there is no significant difference between the observed counts and the expected counts of Mendelian distribution with a single locus model. Results were considered significant ($p < 0.05$) if the $X^2$ value was greater than the critical value of 3.84 ($df = 1$). The chi-square test results reported non-significance for the compounds, indicating a good fit. Table 3.1 shows the chi-square test results for all phenolic compounds.

3.3.2 Principal component analysis (PCA)

Principal component analysis (PCA) was carried out to classify the F2 samples based on their relative composition of CGAs and flavonol glycosides. Figure 3.2A shows the percentage of variance explained by the first two dimensions. The first dimension, i.e., variance component, explained 20.19% of variance and the second dimension explained 16.35%. Together, the first two dimensions described 36.54% of the total variance and presented eigenvalues of 2.83 and 2.29 (Table 3.3A), respectively.

The first principal component (PC1) (Figure 3.2B) shows that most compounds have negative coefficients that somewhat separated the F2 individuals, largely on composition of compounds. Dim1 exhibits moderate to large negative associations with ACQA1, ACQA2, M3GAL, Q3RUT, Q3GLU, Q3RHA, S3GLU, and S3GLC (Table 3.3B). Dim2, the second principal component (PC2), exhibits moderate to large positive associations with ACQA1, ACQA2 and S3GAL and negative associations with CA, Q3RHA, S3GLC, and QCGAL. Figure 3.2B visually shows the first two dimensions. Some potentially interesting F2 individuals were also observed in Figure 3.2D: fruit
samples from plants 62 and 287 located in the upper-left quadrant influenced by compounds with high positive loadings on PC2, and fruit samples from plants 239, 237, and 64 located in the lower-right quadrant influenced by compounds with high positive loadings on PC1.

3.4 Discussion

Blueberry consumption is seeing rapid growth, having an annual per capita consumption that has grown over 510% since the early 2000s (USDA, 2021). Development of new genetic and genomic tools for breeding new cultivars is becoming increasingly important for keeping up with this high demand (Rowland et al., 2012). Mapping populations for identifying quantitative trait loci (QTLs) associated with fruit quality traits are one such essential resource. In the present study, we phenotyped a diploid pseudo F2 population containing 288 individuals, and based on this data, determined possible models of gene control for 14 phenolic compounds. The results demonstrated that these compound levels were segregating with a wide distribution, and based on distribution of parent species phenotypes, suggested that they might be controlled by a single dominant gene. Additionally, the data indicated that this population could be useful for identifying QTLs. Breeding with V. darrowii germplasm could offer to genetically enhance 3-CQA, Q3GAL, Q3RHA, and S3GAL levels, thereby increasing the potential benefits for human health.

3.4.1 Segregation of phenolic traits in the F2 population

A continuous distribution for compound concentration was shown in Figure 3.1, which may be attributed to most compounds being quantitative in nature, likely involving multiple loci. Additionally, ACQA1 and S3GAL segregated in a bimodal distribution, suggesting the presence of major loci. Furthermore, transgressive segregants were observed for ACQA1, ACQA2, M3RHA, Q3RUT, Q3GLU, Q3RHA, S3GAL, and QCGAL (Figure 3.1B, C, F, G, I, J, K N). For ACQA1, ACQA2,
Q3GLU, and Q3RHA, F2 individuals showed concentration values higher than that of the parent V. corymbosum, V. darrowii, and F1 hybrids. On the other hand, for M3RHA, Q3RUT, Q3GLU, S3GAL, and QCGAL, F2 individuals showed concentration values lower than that of the parent species and hybrids. The presence of transgressive segregants suggests V. corymbosum and V. darrowii each have alleles in various proportions for loci governing these compounds.

Figure 3.1 also shows that M3GAL possesses little innate variation (0.006 mg/g FW min to 0.01 mg/g FW max), suggesting that M3GAL might not be a valuable target for breeding efforts. In addition, it was observed that the inclusion of outlier data skewed distribution curves to the right. Outlier data that were inconsistent within replications were removed from the data set. The remaining outlier data was assumed to be, at least partially, due to natural variation in concentration of phenolic compounds.

Chi-square goodness of fit tests revealed that there were no significant differences (p > 0.05) between observed and expected counts of F2 individuals for all phenolic compounds tested (Table 3.1). In other words, the F2 population segregated in a 3:1 phenotypic ratio following a single dominant gene model of control.

3.4.2 Principal component analysis (PCA)

PCA was used to explain the pattern of correlations between different phenolic compounds in the F2 population and to distinguish genotypes especially rich in phenolics. Table 3.3B showed that ACQA1, ACQA2, M3GAL, Q3RUT, Q3GLU, Q3RHA, S3GLU, and S3GLC had moderate to large negative associations with Dim1, indicating that Dim1 might reflect relatively greater flavonol content. ACQA1, ACQA2 and S3GAL had moderate to large positive associations with Dim2,
whereas CA, Q3RHA, S3GLC, and QCGAL had negative associations with Dim2. These observations may reflect relatively greater CGA content at positive Dim2 scores and increasingly non-CGA phenolic acid and flavonol content at negative Dim2 scores. Furthermore, the biplot also showed strong associations between phenolic compounds and F2 individuals present in the upper and lower left quadrants (Figure 3.2D). F2 individuals closely associated with S3GAL were in the upper-right quadrant, and F2 individuals associated with 3-CQA, as well as CA, were in the lower-right quadrant.

The variable plot (Figure 3.3C) indicates that 3-CQA is more strongly correlated to quercetin than to myricetin. This result suggests that some F2 individuals, such as BNJ16-4-159 and BNJ16-4-6, may be of use in future breeding efforts to select both strong antioxidant activity and greater bioavailability of health-related compounds.

3.5 Conclusions

This present study phenotyped a genetically segregating F2 population and determined possible gene models of control for 14 phenolic compounds (four phenolic acids and ten flavonol glycosides). 288 F2 individuals were evaluated. Results revealed that the different phenolic compounds may likely be controlled by a single dominant gene or few loci. This phenotypic data will be used with genotyping-by-sequencing (GBS) data to map QTLs. Although commercial blueberry is tetraploid, genetic analysis at the diploid level offers a simpler model in understanding the genetics for phenolics biosynthesis and fruit composition. These findings might prove useful in the next steps of developing a genetic map and/or identifying loci involved with fruit chemistry.
3.6 References


3.7 Figures and tables

Figure 3.1 Frequency distribution histograms of phenolic compounds in the F2 population (mg/g FW). The vertical lines mark the average concentration of grandparent and F1 genotypes: OPB-8 (dark blue), OPB-15 (blue), NJ88-14-03 (dark green), NJ88-12-41 (green), BNJ05-218-9 (orange) and BNJ05-237-8 (yellow). A 3-caffeoylquinic acid, B acetyl-caffeoylquinic acid isomer 1, C acetyl-caffeoylquinic acid isomer 2, D myricetin-3-galactoside, E caffeoylarbutin, F myricetin-3-rhamnoside, G quercetin-3-rutinoside, H quercetin-3-galactoside, I quercetin-3-glucoside, J quercetin-3-rhamnoside, K syringetin-3-galactoside, L syringetin-3-glucoside, M syringetin-3-glucuronide, and N quercetin-3-(6''-caffeyl)-galactoside.
Figure 3.2 Principal component analysis of phenolic compounds in the F2 population. A scree plot, B individual plot, C variable plot, and D biplot. CQA = 3-caffeoylquinic acid, ACQA1 = acetyl-caffeoylquinic acid isomer 1, ACQA2 = acetyl-caffeoylquinic acid isomer 2, M3GAL = myricetin-3-galactoside, CA = caffeoylarbutin, M3RHA = myricetin-3-rhamnoside, Q3RUT = quercetin-3-rutinoside, Q3GAL = quercetin-3-galactoside, Q3GLU = quercetin-3-glucoside, Q3RHA = quercetin-3-rhamnoside, S3GAL = syringetin-3-galactoside, S3GLU = syringetin-3-glucoside, S3GLC = syringetin-3-glucuronide, and QCGAL = quercetin-3-(6''-caffeyl)-galactoside.
Table 3.1 Observed and expected segregation ratios and chi-square analysis for the F2 population segregating for fruit phenolic compound concentration. 3-CQA = 3-caffeoylquinic acid, ACQA1 = acetyl-caffeoylquinic acid isomer 1, ACQA2 = acetyl-caffeoylquinic acid isomer 2, M3GAL = myricetin-3-galactoside, CA = caffeoylarbutin, M3RHA = myricetin-3-rhamnoside, Q3RUT = quercetin-3-rutinoside, Q3GAL = quercetin-3-galactoside, Q3GLU = quercetin-3-glucoside, Q3RHA = quercetin-3-rhamnoside, S3GAL = syringetin-3-galactoside, S3GLU = syringetin-3-glucoside, S3GLC = syringetin-3-glucuronide, and QCGAL = quercetin-3-(6''-caffeyl)-galactoside.

<table>
<thead>
<tr>
<th>Observed phenotype</th>
<th>Expected phenotype</th>
<th>N</th>
<th>Test ratio</th>
<th>$X^2$</th>
<th>DF</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-CQA</td>
<td>68 220</td>
<td>72 216</td>
<td>288</td>
<td>3:1</td>
<td>0.30</td>
<td>1</td>
</tr>
<tr>
<td>ACQA1</td>
<td>215 73</td>
<td>216 72</td>
<td>288</td>
<td>3:1</td>
<td>0.02</td>
<td>1</td>
</tr>
<tr>
<td>ACQA2</td>
<td>217 71</td>
<td>216 72</td>
<td>288</td>
<td>3:1</td>
<td>0.02</td>
<td>1</td>
</tr>
<tr>
<td>M3GAL</td>
<td>226 62</td>
<td>216 72</td>
<td>288</td>
<td>3:1</td>
<td>1.85</td>
<td>1</td>
</tr>
<tr>
<td>CA</td>
<td>69 219</td>
<td>72 216</td>
<td>288</td>
<td>3:1</td>
<td>1.07</td>
<td>1</td>
</tr>
<tr>
<td>M3RHA</td>
<td>213 75</td>
<td>216 72</td>
<td>288</td>
<td>3:1</td>
<td>1.07</td>
<td>1</td>
</tr>
<tr>
<td>Q3RUT</td>
<td>221 67</td>
<td>216 72</td>
<td>288</td>
<td>3:1</td>
<td>0.46</td>
<td>1</td>
</tr>
<tr>
<td>Q3GAL</td>
<td>221 67</td>
<td>216 72</td>
<td>288</td>
<td>3:1</td>
<td>0.46</td>
<td>1</td>
</tr>
<tr>
<td>Q3GLU</td>
<td>83 205</td>
<td>72 216</td>
<td>288</td>
<td>3:1</td>
<td>2.24</td>
<td>1</td>
</tr>
<tr>
<td>Q3RHA</td>
<td>75 213</td>
<td>72 216</td>
<td>288</td>
<td>3:1</td>
<td>0.17</td>
<td>1</td>
</tr>
<tr>
<td>S3GAL</td>
<td>73 215</td>
<td>72 216</td>
<td>288</td>
<td>3:1</td>
<td>0.02</td>
<td>1</td>
</tr>
<tr>
<td>S3GLU</td>
<td>58 230</td>
<td>72 216</td>
<td>288</td>
<td>3:1</td>
<td>3.63</td>
<td>1</td>
</tr>
<tr>
<td>S3GLC</td>
<td>64 224</td>
<td>72 216</td>
<td>288</td>
<td>3:1</td>
<td>1.19</td>
<td>1</td>
</tr>
<tr>
<td>QCGAL</td>
<td>224 64</td>
<td>216 72</td>
<td>288</td>
<td>3:1</td>
<td>1.19</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 3.2 Phenotypic characteristics of the F2 parent species (V. corymbosum and V. darrowii) for the fruit phenolic compound traits to be analyzed (mg/g FW). 3-CQA = 3-caffeoylquinic acid, ACQA1 = acetyl-caffeoylquinic acid isomer 1, ACQA2 = acetyl-caffeoylquinic acid isomer 2, M3GAL = myricetin-3-galactoside, CA = caffeoylarbutin, M3RHA = myricetin-3-rhamnoside, Q3RUT = quercetin-3-rutinoside, Q3GAL = quercetin-3-galactoside, Q3GLU = quercetin-3-glucoside, Q3RHA = quercetin-3-rhamnoside, S3GAL = syringetin-3-galactoside, S3GLU = syringetin-3-glucoside, S3GLC = syringetin-3-glucuronide, and QCGAL = quercetin-3-(6''-caffeyl)-galactoside.

<table>
<thead>
<tr>
<th></th>
<th>V. corymbosum phenotype (mg/g)</th>
<th>V. darrowii phenotype (mg/g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-CQA</td>
<td>&lt; 0.130</td>
<td>&gt; 0.130</td>
</tr>
<tr>
<td>ACQA1</td>
<td>&gt; 0.050</td>
<td>&lt; 0.050</td>
</tr>
<tr>
<td>ACQA2</td>
<td>&gt; 0.060</td>
<td>&lt; 0.060</td>
</tr>
<tr>
<td>M3GAL</td>
<td>&gt; 0.007</td>
<td>&lt; 0.007</td>
</tr>
<tr>
<td>CA</td>
<td>&lt; 0.040</td>
<td>&gt; 0.040</td>
</tr>
<tr>
<td>M3RHA</td>
<td>&gt; 0.009</td>
<td>&lt; 0.009</td>
</tr>
<tr>
<td>Q3RUT</td>
<td>&lt; 0.015</td>
<td>&gt; 0.015</td>
</tr>
<tr>
<td>Q3GAL</td>
<td>&lt; 0.014</td>
<td>&gt; 0.014</td>
</tr>
<tr>
<td>Q3GLU</td>
<td>&gt; 0.015</td>
<td>&lt; 0.015</td>
</tr>
<tr>
<td>Q3RHA</td>
<td>&lt; 0.008</td>
<td>&gt; 0.008</td>
</tr>
<tr>
<td>S3GAL</td>
<td>&gt; 0.030</td>
<td>&lt; 0.030</td>
</tr>
<tr>
<td>S3GLU</td>
<td>&gt; 0.019</td>
<td>&lt; 0.019</td>
</tr>
<tr>
<td>S3GLC</td>
<td>&gt; 0.012</td>
<td>&lt; 0.012</td>
</tr>
<tr>
<td>QCGAL</td>
<td>&lt; 0.008</td>
<td>&gt; 0.008</td>
</tr>
</tbody>
</table>
Table 3.3 Eigenanalysis of the correlation matrix for phenolic compounds found in the F2 population. **A** eigenvalues and **B** loadings. 3-CQA = 3-caffeoylquinic acid, ACQA1 = acetyl-caffeoylquinic acid isomer 1, ACQA2 = acetyl-caffeoylquinic acid isomer 2, M3GAL = myricetin-3-galactoside, CA = caffeolarbutin, M3RHA = myricetin-3-rhamnoside, Q3RUT = quercetin-3-rutinoside, Q3GAL = quercetin-3-galactoside, Q3GLU = quercetin-3-glucoside, Q3RHA = quercetin-3-rhamnoside, S3GAL = syringetin-3-galactoside, S3GLU = syringetin-3-glucoside, S3GLC = syringetin-3-glucuronide, and QCGAL = quercetin-3-(6''-caffeyl)-galactoside.

### A

<table>
<thead>
<tr>
<th></th>
<th>Eigenvalue</th>
<th>Variance (%)</th>
<th>Cumulative (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dim.1</td>
<td>2.83</td>
<td>20.19</td>
<td>20.19</td>
</tr>
<tr>
<td>Dim.2</td>
<td>2.29</td>
<td>16.35</td>
<td>36.54</td>
</tr>
</tbody>
</table>

### B

<table>
<thead>
<tr>
<th></th>
<th>Dim.1</th>
<th>Dim.2</th>
<th>coord</th>
<th>cos2</th>
<th>contrib</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-CQA</td>
<td>0.07</td>
<td>-0.32</td>
<td>0.11</td>
<td>0.11</td>
<td>2.09</td>
</tr>
<tr>
<td>ACQA1</td>
<td>-0.40</td>
<td>0.61</td>
<td>0.53</td>
<td>0.53</td>
<td>10.34</td>
</tr>
<tr>
<td>ACQA2</td>
<td>-0.45</td>
<td>0.56</td>
<td>0.52</td>
<td>0.52</td>
<td>10.14</td>
</tr>
<tr>
<td>M3GAL</td>
<td>-0.48</td>
<td>0.18</td>
<td>0.26</td>
<td>0.26</td>
<td>5.11</td>
</tr>
<tr>
<td>CA</td>
<td>0.04</td>
<td>-0.46</td>
<td>0.21</td>
<td>0.21</td>
<td>4.10</td>
</tr>
<tr>
<td>M3RHA</td>
<td>-0.11</td>
<td>0.26</td>
<td>0.08</td>
<td>0.08</td>
<td>1.49</td>
</tr>
<tr>
<td>Q3RUT</td>
<td>-0.56</td>
<td>-0.13</td>
<td>0.33</td>
<td>0.33</td>
<td>6.51</td>
</tr>
<tr>
<td>Q3GAL</td>
<td>-0.35</td>
<td>0.40</td>
<td>0.28</td>
<td>0.28</td>
<td>5.51</td>
</tr>
<tr>
<td>Q3GLU</td>
<td>-0.78</td>
<td>0.12</td>
<td>0.62</td>
<td>0.62</td>
<td>12.18</td>
</tr>
<tr>
<td>Q3RHA</td>
<td>-0.55</td>
<td>-0.48</td>
<td>0.54</td>
<td>0.54</td>
<td>10.46</td>
</tr>
<tr>
<td>S3GAL</td>
<td>0.25</td>
<td>0.57</td>
<td>0.38</td>
<td>0.38</td>
<td>7.50</td>
</tr>
<tr>
<td>S3GLU</td>
<td>-0.63</td>
<td>-0.12</td>
<td>0.41</td>
<td>0.41</td>
<td>7.99</td>
</tr>
<tr>
<td>S3GLC</td>
<td>-0.63</td>
<td>-0.42</td>
<td>0.57</td>
<td>0.57</td>
<td>11.19</td>
</tr>
<tr>
<td>QCGAL</td>
<td>-0.14</td>
<td>-0.51</td>
<td>0.28</td>
<td>0.28</td>
<td>5.38</td>
</tr>
</tbody>
</table>