THE GENETICS OF ORGANIC ACIDS VARIATION IN CRANBERRY FRUIT

By

STEPHANIE KAY FONG

A dissertation submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Plant Biology

Written under the direction of

Nicholi Vorsa

And approved by

_____

_____

_____

_____

New Brunswick, New Jersey

MAY, 2019

ABSTRACT OF THE DISSERTATION

THE GENETICS OF ORGANIC ACIDS VARIATION IN CRANBERRY FRUIT

By STEPHANIE KAY FONG

Dissertation Director:

Nicholi Vorsa

American cranberry (*Vaccinium macrocarpon* Ait.*)* fruit are renowned for two traits; extreme sourness and their abundance of health-promoting compounds. Extreme sourness in cranberry fruit is due to the high levels of malic, citric, and quinic acids, which together contribute to titratable acidity (TA). Commercially grown cranberries have a TA of 2.3-3.0%, well over two times the amount in most edible fruits. Thus, considerable amounts of 'added-sugar' are necessary for palatability. In addition to citric, malic, and quinic acids, cranberry fruit also have high levels benzoic acid. While benzoic acid does not statistically contribute to TA, but has health promoting benefits.

To ascertain the variation found in organic acids in existing breeding populations, four bi-parental crosses were phenotyped for malic, citric, quinic, and benzoic acids. Generally, the four organic acids displayed transgressive segregation to the parents. Genotyping by sequencing (GBS) was then used to map genetic diversity within the populations. A total of 61 QTLs were identified for the four organic acids.

In addition to the variation in organic acids in breeding populations, there were two accessions with unique genotypes from the germplasm collection with significantly

lower citric acid (≈ 2 mg/g) (*cita*) and malic acid (≈ 2 mg/g) (*mala*). A series of crosses utilizing these accessions revealed that *cita* and *mala* are independently segregating Mendelian loci. A bulked segregant approach with simple sequence repeats (SSRs), then a QTL identification approach with single nucleotide polymorphisms (SNPs) generated through GBS was used to fine map these two traits. Two SSR markers and one SNP marker were identified for the *cita* locus while two SNP markers were identified for the *mala* locus.

The *cita* trait had multiple alleles contributing to differential levels of citric acid concentrations depending on the parent, e.g. Stevens or #35. Both the *cita* and *mala* traits exhibit partial dominance. In two dihybrid crosses with both *cita* and *mala,* an epistatic effect was between these two traits. There was a significant effect of the *cita* alleles on increasing malic acid concentration while the presence of *mala* alleles reduced both citric and malic acid concentrations. This work determined the inheritance and variation of organic acids as well as developed molecular markers linked with low citric and low malic acid traits. These markers will be used for marker assisted selection to accelerate the breeding process of cranberry.

ACKNOWLEDGMENT

First and foremost, I would like to thank Dr. Nicholi Vorsa for taking me on and allowing my dissertation project to take the form it is now. I would like to thank my committee members, Drs. Andrea Gallavotti, James White, and Jack Mardekian for their insights regarding my projects.

I would like to thank Drs. Ann Hirsch, Maskit Maymon, and Nancy Fujishige for mentoring me and inspiring me to pursue a degree in plant breeding.

I would like to thank everyone at the Marucci Research Station who undoubtedly aided me in a plethora of tasks: Sue Vancho, Karen Destefano, Rich Destefano, Ruben Colon, Tony Colon, Kristia Adams, Dr. Tierry Besacon, Dr. Peter Oudaemans. I would especially like to thank Dr. Jennifer Johnson-Cicalese for keeping everyone and everything organized and always talking things out with me. I would like to thank Dr. Jim Polashock for helping with the molecular side of my projects.

I am eternally grateful to Dr. Joe Kawash for the endless supply of advice and most of all, aiding in all the bioinformatics in this dissertation. I would like to thank Dr. Yifei Wang for all the help with the analytical chemistry.

I would like to thank Dr. Josh Honig, Jennifer Vaiciunas, and Chrissy Kubik for their help with the SSR work.

I would like to thank everyone in the Plant Biology Department: Sarah Gilbert, David Brynes, Robert Pyne, Megan Muelbauer, Justin Lombardoni, Henry Qu, Timmy Waller, and Vinnie Averello for their friendship.

Special thanks to Lisa Lamanna, Ariane Vasilatis, and Eve Tang for keeping me sane throughout my graduate studies.

And lastly, thank you to Kyle Yang for taking care of me for all this time.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

Chapter 1: Literature Review

Cranberry (*Vaccinium macrocarpon* Ait.)

Vaccinium macrocarpon, commonly known as American cranberry, is in the Ericaceae family. The genus Vaccinium also includes other economically important crops such as blueberry (section *Cyanococcus*) and lingonberry (section *Vitis-idaea*), along with other locally cultivated crops (Lyrene et al., 2003). The American cranberry, native to North America, is well adapted to sandy, acidic, and highly organic soils. It is a temperate woody evergreen perennial that reproduces both sexually and asexually. Asexually, it proliferates via stolons, which allows for vegetative propagation. In the late summer and early fall, the stolons produce uprights with inflorescence buds, which undergo an overwintering requirement of 800-1000 chilling hours before coming out of dormancy. In the spring, these inflorescence buds then produce 3-7 perfect, protandrous flowers per upright. While the flowers display protandry and promote out-crossing, with pollen shed before the stigma are receptive, they are still self-compatible. After pollination and fruit set in the early summer, round, ovate elliptic, red fruit are produced in the fall. The fruit range from largely 1cm to 4cm in diameter, and contain four vacant locules, allowing the fruit to float in water. (Vorsa and Johnson-Cicalese, 2012)

Commercial cranberry production in the United States occurs in New Jersey, Massachusetts, Maine, Michigan, Wisconsin, Oregon, and Washington, with largest production in Wisconsin. The US cranberry production has a current economic value of $385 million. Internationally, cranberries are cultivated in Canada- British Columbia, Quebec, New Brunswick, Nova Scotia, and Newfoundland, Chile, and Eastern Europe.

Cranberries are traditionally planted in constructed beds with raised borders for holding water for easy harvest and winter flood. In the past, due to vegetative propagation, dormant vines would be “disced-in” or pressed to the soil in early spring. Today however, to maintain genetic homogeneity of the Rutgers’ varieties in each bed, rooted cuttings are planted. This has twofold benefit: first, to maintain the genetic integrity and second, encourage faster establishment and bed establishment of the beds. At harvest, the beds are harvested by flooding the beds with water, and using a tractor equipped with a harrow to remove the fruit from the vines. The fruit then floats and is gathered using a boom. The collected fruit is then sent to a processing station, which evaluates every truckload of fruit for amount of disease and fruit quality traits which include titratable acidity, brix, and total anthocyanins. The price paid per load is determined by a combination of these traits, with more paid out for the optimal amount TAcy and usable fruit. (Vorsa and Johnson-Cicalese, 2012). Fruit is also harvested for the fresh market by a dry method or with a water method.

Currently, fruit rot is one of the most pressing challenges to cranberry production. Thus, a main focus of the Rutgers cranberry breeding program has been for yield and fruit rot resistance. With warming climates, fruit rot is becoming more prevalent and a concern to growers in production areas other than New Jersey (Gallardo et al., 2018). There has been genetic gain in breeding for fruit rot resistance, which started with identifying fruit rot resistant accessions in the germplasm collection, determining heritability, and identifying QTL (Georgi et al., 2013; Johnson-Cicalese et al., 2015; Daverdin et al., 2017).

In recent years, there has been another pressure against the cranberry industry: the need for added sugars to cranberry products. Cranberries have high acidity, from citric, malic, and quinic acids, and are considered to be extremely tart by the general public, thus requiring added sugar for palatability. Due to a 'war on sugar,' added sugar have been the target of health regulations (FDA, 2018). This war on sugar stems from a multitude of studies showing that increased consumption of sugar can be addicting, but also can lead to metabolic syndrome and obesity (Lustig et al., 2012). In addition to being made into a sauce to accompany roast turkey during Thanksgiving, the majority of the cranberry crop (over 95%) is processed into products, primarily sweetened dried cranberries and as a byproduct, cranberry juice cocktail (USDA NASS, 2017).

However, cranberries have not always been so augmented to be consumed. Historically, cranberries were used by Native Americans to help preserve and add nutrition to dried meat in the winter. They also consumed cranberries raw or as a sweetened sauce. In addition to nourishment, they also used cranberries as a treatment for wounds and blood poisoning. After the arrival of the Pilgrims to what is now the United States, cranberries were also used to relieve symptoms of scurvy. (Eck, 1990)

Breeding History

Domestication of the American cranberry began in the 1800s with selections from the wild. The first cranberry breeding program was started by the USDA in cooperation with NJAES in 1929 and focused on breeding for resistance to false blossom disease. From this program, a widely grown cultivar, Stevens, was released in 1953. (Vorsa and Johnson-Cicalese, 2012) Since then, Rutgers has released 6 cultivars that are gaining in popularity both in New Jersey and Wisconsin (Vorsa, 2012; 2008a; 2008b; 2007; Vorsa

and Johnson-Cicalese, 2017a; 2017b). The Marucci Center also houses a germplasm collection of American cranberry from various previous programs and collections of wild germplasm from Southern Canada through the Mid-Atlantic and Appalachian regions. This collection also includes germplasm of diploid and polyploid *V. oxycoccus*, commonly known as small cranberry. Diploid *V. oxycoccus* can form interspecific hybrids with *V. macrocarpon*.

Today, breeding for cranberry is primarily conducted by three programs, the USDA-ARS and University of Wisconsin-Madison in Wisconsin, a private breeder in Wisconsin (Valley Corporation), and Rutgers University. Breeding priorities for cranberry are primarily dictated and influenced by industry needs, as the industry (growers and cooperatives) contribute significant amounts of funds for cranberry research. The primary thrust for all three of these programs is yield, TAcy, and disease resistance (Gallardo et al., 2018). The Rutgers' program also focuses on fruit quality, which includes TAcy, proanthocyanidins (PACs), brix, and titratable acidity (TA). From these breeding goals, there have been varieties released from Rutgers with enhanced yield and TAcy such as Mullica Queen, Demoranville, and Scarlet Knight (Vorsa, 2007; 2008; 2012). Fruit with uniform fruit color and firmness are current breeding goals of the Rutgers' program.

Genetics of Cranberry

In addition to traditional breeding, there have been various genetic studies done in cranberry to aid breeding efforts. Genetically, the American cranberry is a diploid (2n=2x=24) with 12 metacentric and submetacentric chromosomes in its genome (Zdepski et al., 2011). Genetic characterization and development of genetic tools in

cranberry began in 1994 with the development of randomly amplified polymorphoic DNA (RAPDs). However, RAPDs had issues as to repeatability, so sequence characterized amplified region (SCAR) markers were generated from these RAPDs. These markers were useful in differentiating different cranberry varieties, but there were not enough for molecular mapping. (Vorsa and Johnson-Cicalese, 2012) Thus, simple sequence repeats (SSRs) were developed and the first genetic map of cranberry was developed (Zhu et al., 2012 and Georgi et al., 2013). As technology progressed, single nucleotide polyphorisms (SNPs) were developed with the advent of reduced sequencing costs and the development of genotyping by sequencing (GBS) (Covarrubias-Pazaran et al., 2016; Schlautman et al., 2017; Daverdin et al., 2017). Currently, there have been five published linkage maps, generated with SSRs, and then SNPs from GBS (Georgi et al., 2013; Schlautman et al., 2015; Covarrubias-Pazaran et al., 2016; Schlautman et al., 2016; Daverdin et al., 2017). There was also study on comparative genetic mapping between diploid blueberry and cranberry (Schlautman et al., 2018). With all these genomic resources developed, an implementation of genomic selection in cranberry using multivariate genomic best linear unbiased predictions (gBLUPs) for the estimation of yield and fruit weight was developed (Covarrubias-Pazaran et al., 2018).

In additional to these genetic studies, the genome has also been sequenced. The plastid genome was first sequenced using Roche 454 shotgun sequencing (Fajado et al., 2013). Then, the mitochondrial genome was also sequenced using the same sequencing platform (Fajado et al., 2014). Finally, a 5^{th} generation inbred Ben Lear individual was used for whole genome sequencing on the Illumina 2x150bp platform with 20x coverage with 229,745 contigs (Polashock et al., 2014). In the same study, the transcriptome was

also sequenced with 5x coverage was used to help predict genes in the genome. In a later study, to identify putative genes associated with flavonoids in cranberry, de novo sequencing of the transcriptome was done (Sun et al., 2015). Currently, using PacBio long read sequencing and Oxford Nanopore long read sequencing, there is a chromosome level genome assembly with 124 contigs with a genome size of 487Mb (Kawash et al., unpublished).

Cranberry Organic Acids

Cranberry fruit are renowned for their tart flavor and contain various organic acids which contribute to both the flavor and high acidity. There are various organic acids that are most commonly found in fleshy fruit, including: malic, citric, quinic, oxalic, and tartaric acids. These acids are primarily linked with several metabolic pathways: the Krebs cycle, the shikimate pathway, glycolysis, and the glyoxylate pathway in plants. Overall, these pathways play a role in general plant metabolism as well as energy production and consumption. Additionally, fleshy fruits may accumulate these organic acids for mechanisms of seed dispersal e.g. animal dispersal. In most fruit species, fruit is high in acid early in development to prevent herbivory. When the seeds are mature, the organic acids get metabolized and/or the soluble sugar content increases, to entice animals to consume and disperse the seeds. (Walker and Famiani, 2018)

In cranberry fruit, the primary organic acids are citric, malic, and quinic acids, with benzoic acid in smaller quantities (Wang et al., 2017). Citric and malic acids are the primary contributors to titratable acidity (TA), a measure of sourness. Quinic acid, a major component contributing to bitterness in coffee, also imparts a bitter taste to cranberry fruit (McCamey et al., 1990). Benzoic acid does not contribute to taste but act

as a preservative and is likely why cranberries have a relatively longer shelf life (Marwan and Nagel, 1986). Quinic and benzoic acids can also have human health benefits. Quinic acid has anti-inflammatory effects while benzoic acid has anti-microbial effects (Sheng et al., 2005; Marwan and Nagel, 1986).

Objectives

The focus of this thesis research is threefold, 1) to identify quantitative variation of organic acids and fruit quality traits in breeding populations, 2) determine the genetics of low acid traits found in two germplasm accessions, and 3) to identify markers associated to qualitative low citric and low malic acid traits and explore the interaction between these two traits. Although there have been studies quantifying the amounts of organic acids in different cranberry cultivars, there has not been any studies towards identifying the genetic basis for variation of these acids. Additionally, until now, there has not been identification and characterization of low citric and low malic acid traits in cranberry. The contributions of this research will aid with breeding for lower acidity cranberries that require reduced 'added-sugar' and have enhanced health benefits to sustain and provide more cultivars for the cranberry industry. The scientific results of this research may also be applied in breeding another economically important crop, blueberries, for taste and human health.

In the following chapters, a study on the characterization and genetic control of the levels of organic acids in cranberry fruit is described and discussed. Chapter 2, entitled Variation and QTL Mapping of Fruit Chemistry Traits in Fruit Rot Resistant Breeding Populations of American Cranberry, focuses on identifying QTL and quantitative variation for organic acids. Chapter 3, entitled Genetics and Molecular

Mapping of a Low Citric Acid Trait in Cranberry Fruit and Its Relationship to Titratable Acidity, describes the identification of a low citric acid trait in the fruit and identifies the mode of inheritance and genetic markers linked with this trait. Chapter 4, entitled Development of Molecular Markers for Low Malic Acid and Its Effect on Citric Acid in Cranberry, describes the identification of markers for a low malic acid trait in the fruit and describes the relationship between malic and citric acids. Chapter 5, Interaction of Low Malic Acid and Low Citric Acid Alleles Affecting Titratable Acidity in Cranberry, describes the effect of low malic and low citric acid alleles on the fruit organic acid content and titratable acidity. The findings described and discussed in this dissertation contributes to the current knowledge on cranberry breeding, especially for fruit acidity. Overall, the research presented in this dissertation provides knowledge on the genetic behavior of organic acids in cranberry fruit. This is the first time the genetics and variation of organic acids has been explored in cranberry. This research contributes to the current breeding program to incorporate marker assisted selection of fruit organic acids, to improve the healthiness and palatability of cranberry fruit.

References

Covarrubias-Pazaran, G., Diaz-Garcia, L., Schlautman, B., Deutsch, J., Salazar, W., Hernandez-Ochoa, M., Grygleski, E., Steffan, S., Iorizzo, M., Polashock, J., Vorsa, N., Zalapa, J., 2016. Exploiting genotyping by sequencing to characterize the genomic structure of the American cranberry through high-density linkage mapping. BMC Genomics 17, 451. https://doi.org/10.1186/s12864-016-2802-3

Daverdin, G., Johnson-Cicalese, J., Zalapa, J., Vorsa, N., Polashock, J., 2017. Identification and mapping of fruit rot resistance QTL in American cranberry using GBS. Mol Breeding 37, 38. https://doi.org/10.1007/s11032-017-0639-3

Eck, P., 1990. The American cranberry. Rutgers University Press, New Brunswick.

Fajardo, D., Schlautman, B., Steffan, S., Polashock, J., Vorsa, N., Zalapa, J., 2014. The American cranberry mitochondrial genome reveals the presence of selenocysteine (tRNA-Sec and SECIS) insertion machinery in land plants. Gene 536, 336–343. https://doi.org/10.1016/j.gene.2013.11.104

Fajardo, D., Senalik, D., Ames, M., Zhu, H., Steffan, S.A., Harbut, R., Polashock, J., Vorsa, N., Gillespie, E., Kron, K., Zalapa, J.E., 2013. Complete plastid genome sequence of Vaccinium macrocarpon: structure, gene content, and rearrangements revealed by next generation sequencing. Tree Genetics & Genomes 9, 489–498. https://doi.org/10.1007/s11295-012-0573-9

Gallardo, R.K., Klingthong, P., Zhang, Q., Polashock, J., Atucha, A., Zalapa, J., Rodriguez-Saona, C., Vorsa, N., Iorizzo, M., 2018. Breeding Trait Priorities of the Cranberry Industry in the United States and Canada. HortScience 53, 1467–1474. https://doi.org/10.21273/HORTSCI13219-18

Georgi, L., Johnson-Cicalese, J., Honig, J., Das, S.P., Rajah, V.D., Bhattacharya, D., Bassil, N., Rowland, L.J., Polashock, J., Vorsa, N., 2013. The first genetic map of the American cranberry: exploration of synteny conservation and quantitative trait loci. Theor Appl Genet 126, 673–692. https://doi.org/10.1007/s00122-012-2010-8

Johnson-Cicalese, J., Polashock, J.J., Honig, J.A., Vaiciunas, J., Ward, D.L., Vorsa, N., 2015. Heritability of Fruit Rot Resistance in American Cranberry. Journal of the American Society for Horticultural Science 140, 233–242. https://doi.org/10.21273/JASHS.140.3.233

Lustig, R.H., Schmidt, L.A., Brindis, C.D., 2012. Public health: The toxic truth about sugar. Nature 482, 27–29. https://doi.org/10.1038/482027a

Lyrene, P.M., Vorsa, N. and Ballington, J.R., 2003. Polyploidy and sexual polyploidization in the genus Vaccinium. *Euphytica*, *133*(1), pp.27-36.

Marwan, A.G., Nagel, C.W., 1986. Characterization of Cranberry Benzoates and Their Antimicrobial Properties. Journal of Food Science 51, 1069–1070. https://doi.org/10.1111/j.1365-2621.1986.tb11236.x

McCamey, D.A., Thrope, T.M., and McCarthy, J.P. 1990. Coffee Bitterness. p. 169-182. In: R.L. Rouseff (ed.). Bitterness in Foods and Beverages. Elsevier Science Publishers B.V. Amsterdam, The Netherlands.

U.S. Food and Drug Administration. 2019. Labeling & Nutrition - Changes to the Nutrition Facts Label [WWW Document]. URL https://www.fda.gov/Food/GuidanceRegulation/GuidanceDocumentsRegulatoryInformation/LabelingNutrition/ucm385663.htm (accessed 2.6.19).

Polashock, J., Zelzion, E., Fajardo, D., Zalapa, J., Georgi, L., Bhattacharya, D., Vorsa, N., 2014. The American cranberry: first insights into the whole genome of a species adapted to bog habitat. BMC Plant Biology 14, 165. https://doi.org/10.1186/1471-2229-14-165

Schlautman, B., Covarrubias-Pazaran, G., Diaz-Garcia, L., Iorizzo, M., Polashock, J., Grygleski, E., Vorsa, N., Zalapa, J., 2017. Construction of a High-Density American Cranberry (Vaccinium macrocarpon Ait.) Composite Map Using Genotyping-by-Sequencing for Multi-pedigree Linkage Mapping. G3 (Bethesda) 7, 1177–1189. https://doi.org/10.1534/g3.116.037556

Schlautman, B., Covarrubias-Pazaran, G., Diaz-Garcia, L.A., Johnson-Cicalese, J., Iorrizo, M., Rodriguez-Bonilla, L., Bougie, Tierney, Bougie, Tiffany, Wiesman, E., Steffan, S., Polashock, J., Vorsa, N., Zalapa, J., 2015. Development of a high-density cranberry SSR linkage map for comparative genetic analysis and trait detection. Mol Breeding 35, 177. https://doi.org/10.1007/s11032-015-0367-5

Schlautman, B., Diaz-Garcia, L., Covarrubias-Pazaran, G., Schlautman, N., Vorsa, N., Polashock, J., Ogden, E.L., Brown, A., Lin, Y.-C., Bassil, N., Buck, E.J., Wiedow, C., McCallum, S., Graham, J., Iorizzo, M., Rowland, L.J., Zalapa, J., 2018. Comparative genetic mapping reveals synteny and collinearity between the American cranberry and diploid blueberry genomes. Mol Breeding 38, 9. https://doi.org/10.1007/s11032-017-0765-y

Sheng, Y., Åkesson, C., Holmgren, K., Bryngelsson, C., Giamapa, V., Pero, R.W., 2005. An active ingredient of Cat's Claw water extracts: Identification and efficacy of quinic acid. Journal of Ethnopharmacology 96, 577–584. https://doi.org/10.1016/j.jep.2004.10.002

Sun, H., Liu, Y., Gai, Y., Geng, J., Chen, L., Liu, H., Kang, L., Tian, Y., Li, Y., 2015. De novo sequencing and analysis of the cranberry fruit transcriptome to identify

putative genes involved in flavonoid biosynthesis, transport and regulation. BMC Genomics 16. https://doi.org/10.1186/s12864-015-1842-4

United States Department of Agriculture- National Agriculture Statistics Service. 2017. *Wisconsin Ag News- Cranberries* [Data file] Retrieved from https://www.nass.usda.gov/Statistics_by_State/Wisconsin/Publications/Crops/2018/WI-Cranberries-Annual-06-18.pdf

Vorsa, N., 2012. Cranberry plant named 'CNJ95-20-20.' USPP22541P3.

Vorsa, N., 2008a. Cranberry variety named 'NJS98-35.' USPP18911P2.

Vorsa, N., 2008b. Cranberry plant named 'CNJ97-105-4.' USPP19434P3.

Vorsa, N., 2007. Cranberry variety named 'NJS98-23.' USPP18252P3.

Vorsa, N., Johnson-Cicalese, J., 2012. American Cranberry, in: Badenes, M.L., Byrne, D.H. (Eds.), Fruit Breeding. Springer US, Boston, MA, pp. 191–223. https://doi.org/10.1007/978-1-4419-0763-9_6

Vorsa, N., Johnson-Cicalese, J., 2017a. Cranberry variety named 'CNJ99-52-15.' USPP27709P3.

Vorsa, N., Johnson-Cicalese, J., 2017b. Cranberry variety named 'CNJ99-9-96.' USPP27657P3.

Walker, R.P., Famiani, F., 2018. Organic Acids in Fruits, in: Horticultural Reviews. John Wiley & Sons, Ltd, pp. 371–430. https://doi.org/10.1002/9781119431077.ch8

Wang, Y., Johnson-Cicalese, J., Singh, A.P., Vorsa, N., 2017. Characterization and quantification of flavonoids and organic acids over fruit development in American cranberry (Vaccinium macrocarpon) cultivars using HPLC and APCI-MS/MS. Plant Science 262, 91–102. https://doi.org/10.1016/j.plantsci.2017.06.004

Zdepski, A., Debnath, S.C., Howell, A., Polashock, J., Oudemans, P., Vorsa, N., Michael, T.P., 2016. Cranberry, in: Genetics, Genomics and Breeding of Berries. CRC Press, pp. 41–63.

Zhu, H., Senalik, D., McCown, B.H., Zeldin, E.L., Speers, J., Hyman, J., Bassil, N., Hummer, K., Simon, P.W., Zalapa, J.E., 2012. Mining and validation of pyrosequenced simple sequence repeats (SSRs) from American cranberry (Vaccinium macrocarpon Ait.). Theor Appl Genet 124, 87–96. https://doi.org/10.1007/s00122-011-1689

Chapter 2: Variation and QTL Mapping of Fruit Chemistry Traits in Fruit Rot Resistant Breeding Populations of American Cranberry

Introduction

The American cranberry, *Vaccinium macrocarpon* Ait. ($2n = 2x = 24$), is a woody perennial and a temperate fruit crop species native to eastern North America. Cranberry has been considered a "super fruit" for many years, having various health benefits such as urinary tract health (Howell, 2007), anticancer bioactivities (Pappas and Schaich, 2009), anti-atherogenic activities (Shabrova et al., 2011), and prevention of metabolic syndrome (Kowalsk and Olejenik, 2016). Major compounds of cranberries contributing to the health effects and antioxidant activities are the flavonoid classes: proanthocyanidins, flavonols, and anthocyanins (Kowalsk and Olejenik, 2016).

Cranberries also present uniquely high levels of acid and are renowned for their sourness, affecting their palatability as fresh fruit. There is an optimal sugar to acid ratio for palatable fruit product formulation (Bates et al., 2001). Consequently, considerable amounts of sugar are required to balance the high acidity in cranberry products, such as sweetened-dried cranberries and cranberry juice cocktails (Cunningham et al., 2004). Products labeled with 'added-sugar' have been stigmatized in recent years due to a 'war on sugar' to regulate and reduce the amounts of sugar consumed in the United States (Lustig et al., 2012). Although some acidity, e.g., TA≈0.5%, is required for flavor expression, reducing excessive acidity should enable a reduction in the amount of sugar necessary for palatability, and help maintain cranberry's super fruit status (Bates et al., 2001).

The cranberry industry is worth about 266 million dollars in the United States, with the majority of fruit production in New Jersey, Massachusetts, Wisconsin, Washington,

and Oregon. Canada is the second largest producer of cranberries, followed by Chile and smaller productions in Eastern Europe (Vorsa and Johnson-Cicalese, 2012). Cranberry fruit are commonly consumed as sweetened dried cranberries, juice, sauces, and processed into nutraceutical products. Fruit quality is important in the production of cranberry products, especially sweetened dried cranberries, which currently makes up 90% of the market (Ocean Spray, personal communication).

Fruit quality parameters include the physical characteristics size, shape, firmness, and resistance to field and storage 'rot' diseases. For processed fruit, three fruit chemistry criteria are also typically measured: total anthocyanins (Tacy) (mg/100g FW), Brix (percent soluble solids), and titratable acidity (TA) (in citric acid equivalents). TAcy is a particularly important trait to the cranberry industry as a quantitative measurement of fruit 'redness' and an indicator of ripeness and product appeal (Vorsa and Johnson-Cicalese, 2012). For juice products, a minimum TAcy level is required and cultivars have been developed with increased TAcy, for example; Scarlet Knight®, Crimson Queen®, Demoranville®, and HyRed® (Vorsa, 2007; 2008; 2012; McCown, 2003). However, desired Tacy levels for sweetened-dried cranberry are within a narrow window (30-50mg/100g FW), and generally lower than desired for juice concentrate. Brix, a measurement of soluble solids, is generally a measure of sugars in most fruit species, but because of the high acid content in cranberry, organic acids also contribute to the brix value (Cunningham et al., 2004). Cranberry products, e.g., juice, are formulated based on percent brix of the expressed juice (Bates et al., 2001). TA is a measurement of sourness, and is largely the result of the sum of citric and malic acids, and to a lesser degree quinic acid.

Benzoic acid is present at much lower concentrations, 0.1%, and is a weaker acid than the other three acids, contributing little to variation in TA.

High quinic acid (pKa 3.46) levels are found in just kiwi, plum, and cranberry fruit, and is sometimes used to distinguish cranberry juice from other similar berry fruit juices (Marsh et al., 2009; Walker and Famiano, 2018; Ehling and Cole, 2011). Quinic acid has been shown to work as an anti-inflammatory compound (Sheng et al., 2005) and is also attributed to bitterness in coffee (McCamey et al., 1990), potentially contributing to cranberry's unique taste. Benzoic acid is also quite unique to cranberry fruit. With the exception of cloudberry, *Rubus chamaemorus,* benozic acid it is not usually found in fruit of other crop species (Baardseth and Russwurm, 1978). Quinic and benzoic acids are thought to have high anti-fungal activity in vitro against the fruit rot complex found in cranberry, while malic acid had less inhibitory activity (Tadych et al., 2015). Benzoic acid specifically has been found to have an anti-microbial effect (Marwan and Nagel, 1986). Citric acid is the strongest acid (pKa 3.13, 4.76, 6.40) contributing to TA. It is commonly found in other fruit, most notably citrus, and contributes to the sour taste of cranberry. Malic acid (pKa 3.4, 5.11) is the second strongest acid that affects TA. It is most notably found in apples, and contributes a tart taste to cranberry fruit (Rubico and McDaniel, 1992). Overall, the balance of malic, citric, and quinic acids along with brix contribute to the taste, while quinic and benzoic acids contribute to the health-promoting properties of cranberries.

A major and growing challenge for cranberry production in the Northeastern United States (Massachusetts and New Jersey) is fruit rot, which is the degradation of the cranberry fruit caused by a fungal complex (Gallardo et al., 2018). Thus, current breeding efforts are focused on developing fruit rot resistant varieties through introgression of three fruit rot

resistant germplasm accessions into elite horticultural backgrounds (Johnson-Cicalese et al., 2016). Multiple QTL studies have been performed to identify QTL for fruit rot resistance, yield, TAcy, brix, and TA (Daverdin et al., 2017; Georgi et al., 2012; Diaz-Garcia et al., 2018). Two of the fruit rot resistant germplasm accessions, Budd's Blues and US89-3, had notably high quinic and benzoic acid levels. Utilizing these previously SNP-genotyped populations, this study identified QTL for specific organic acids and flavonoids. Tacy, brix, TA, and proanthocyanidin QTL have previously been identified in cranberry (Georgi et al., 2012; Diaz-Garcia et al., 2018), but this is the first study to identify QTLs for organic acids. As a woody perennial with a long generation time (about three years from seed to fruit), utilizing QTL and marker assisted selection has the potential to accelerate cranberry breeding efforts.

The primary objective of this study was to evaluate fruit quality traits, particularly organic acids, to determine if there is a genetic component to quantitative variation for fruit organic acids in cranberry. Additionally, the data collected in this study were used to gain preliminary insight on the relationship, or lack thereof, of fruit quality traits with fruit rot resistance. Levels of TAcy, brix, TA, and the four organic acids (citric, malic, quinic, and benzoic acids) were quantified in four biparental breeding populations represented by three way half-sib families, and a pair of two way half-sib families. The GBS SNP data of Daverdin et al. (2017) were employed to assess QTL associated with these traits.

Materials and Methods

Plant Materials and Fruit Collection. Three diverse fruit rot resistance germplasm accessions were used as parents for crosses, Budd's Blues (BB), Cumberland (CU), and US89-3, along with the elite cultivar Crimson Queen (CQ) (Daverdin et al., 2017). BB and

US89-3 are considered resistant to fruit rot with low yield while CU is considered moderately resistant to fruit rot with higher yield. CQ is susceptible to fruit rot with high yield (Johnson-Cicalese et al., 2015). Progeny from four crosses, BBxCQ, BBxCU, BBxUS89-3, and CUxUS89-3 were planted in field plots in 2009. Each population was divided in two blocks in a randomized complete block design with one plot per progeny. Each parent and standard (Susceptible: Stevens, Mullica Queen, CQ, and Resistant: CU, US89-3, US88-1, and BB) was replicated three times.

A 10-gram fruit sample per plot was collected for organic acid extraction from 2014 to 2016. In 2014, fruit from the parents and standards were collected every 10 days throughout the growing season (July 4th through October 8th), for a total of 9 harvest dates to quantify the quinic, benzoic, citric, and malic acid levels in developing fruit, and determine the variation throughout the season. The last harvest date was October 8th, which is considered to be the peak in New Jersey's commercial cranberry harvest. For BB x CQ progeny, fruit were collected once at the start of the season (July) and once at the end of the season (October) in 2014 and 2015. For BB x CU, fruit were collected in October 2014, July 2015, October 2015, and August 2016. For BB x US89-3 and CU x US89-3, fruit were collected just in October 2016. For Tacy, TA, and brix analysis, additional fruit samples, 100g/plot, for all populations were collected in September and October of 2014, 2015, or 2016. Parents and standards were also harvested on each of the above dates. All fruit was stored in a -4°C freezer until processed.

Fruit Quality Traits Evaluation. Organic acid quantification was performed as in Wang et al. (2017). Organic acids were extracted using a 10:1 ratio of water (mL) to grams of fruit, ground in a standard laboratory blender until smooth. For the parents and standard

varieties, single berries from each replicated plot were used to make the extract. For the four populations, three berries of approximately equal size and color were homogenized together for the organic acid extraction. Quinic, citric, malic, and benzoic acids were quantified using a Dionex High-Performance Liquid Chromatography (HPLC) system with a PDA detector and a Waters Atlantis 250 x 4.6mm dC18 5um LC column. TAcy, brix, and TA analyses were performed as described by Vorsa and Johnson-Cicalese (2012). TAcy was measured using a spectrophotometric method at 515nm, brix was measured using an ATAGO PR-32 digital refractometer, and TA was measured by titrating to an endpoint of pH 8.2 using 0.1N NaOH.

Fruit Rot Resistance. Fruit rot resistance previously quantified and published in Daverdin et al. (2017) for all the individuals used in the current study.

Statistical Analysis. SAS v9.4 was used for statistical analysis. PROC UNIVARIATE was used to determine if the data was normally distributed. PROC GLM was used to determine year-to-year variation, year was set as a fixed effect. PROC CORR with the standard parameters was used to determine correlation between traits in each population.

QTL analysis. For QTL analysis, individual population SNP genetic maps from Daverdin et al. (2017) were analyzed with WinQTL Cartographer 2.5. Composite interval mapping was used with the following parameters: zmapqtl model 6 and hypothesis test 10. QTL analysis was performed individually for each harvest date and each trait. In all the populations, QTL were detected if the LOD score was greater than 3.

Results

Phenotype Data: Parents and Standards. Overall, quinic acid levels decreased during fruit development while benzoic and malic acids increased. The decrease in quinic acid as related to harvest date is largely linear, with R^2 of 58% for linear fit, with only 8% additional variation accounted for with a quadratic fit. Citric acid levels had a large increase between July 14th and August 20th, and a 57% linear fit. After August 20th, citric acid remained fairly level, with only 2% linear fit. Malic acid was largely linear, 45% linear fit, with an additional 9% variation accounted for by a quadratic fit of 54%, with malic acid levels increasing with later harvest dates. Benzoic acid is not detectable in some cultivars until July 30th and continues to increase until harvest. Overall, benzoic acid had a 49% linear fit and 54% quadratic fit, increasing with later harvest dates.

There were significant differences among varieties for each acid, on the multiple sampling dates in 2014, which indicate differential temporal variation during fruit development and ripening (Figure 1). Overall, the July 30th date had the greatest variability among the varieties in organic acid levels, for example, there was a range of 19.6 to 37.1 mg/g quinic acid. Over all the dates, Budd's Blues had the highest levels of benzoic acid while US89-3 had the highest levels of quinic acid (Figure 1). For malic acid, there was no significant variation between varieties on the final harvest date (October 8th).

However, in commercial cranberry production, fruit is harvested from late September to early October. Since there was no significant variation for variety by date for the last two dates, the mean of the last two harvest dates was used to test the variability among the parents (BB, CQ, US89-3, and CU). For quinic, malic, and benzoic acids, there was significant variation between the fruit rot resistant varieties (BB, CU, and US89-3) and

CQ (Figure 2). Thus, parental variability suggested that segregation among progeny might be expected and these populations were suitable for QTL analysis. In addition, there was also significant variation for TAcy and TA as well (Figure 1 and Table 1).

Phenotype Data: Progeny. For all crosses, the organic acid data was normally distributed on all harvest dates except for benzoic acid on the July 30th sampling dates, due to very low levels early in the season (similar to the parents in Figure 1d). There was less year-to-year variation on the October harvest dates compared to the July harvest dates (fruit development and sizing period), which tended to have significant differences year-to-year for all four organic acids (data not shown). For example, the BB x CQ July population mean was significantly different for citric and malic acids between the two harvest years ($p < 0.005$), and the BB x CU July population mean was significantly different for all four acids between the harvest years. In the BB x CQ and BB x CU populations, there was no significant year-to-year variation for TAcy, but there was significant variation between years for brix and TA. There were also significant differences between the September and October harvest dates for TAcy, brix, and TA (data not shown). Thus, all dates were used for QTL analysis.

The progeny displayed transgressive segregation for organic acid and fruit chemistry levels between the two parents for each cross, and had a wide range in values (Table 1). In some cases, there was a four-fold difference between the minimum and maximum values in the progeny.

The correlation matrices for both BB x CQ and BB x CU during the October harvest dates show significant correlations between malic and citric acids ($r= 0.25$ to 0.48, with $p < 0.05$) (Figure 3a). For the same two populations, there was a significant negative

correlation between TAcy and berry weight (r= -0.20 to -0.26, with $p < 0.05$) (Figure 3b). Malic and citric acid were correlated in the BB x US89-3 population (r= 0.49, with $p < 0.005$) (Figure 3c). In the BB x CU population in 2014 and 2015, there was a significant positive correlation of malic acid with TA (r= 0.35 and 0.22, with $p < 0.0005$ and $p < 0.05$, respectively), and citric acid with TA (r= 0.37 and 0.28, with $p < 0.0005$ and $p < 0.005$, respectively) (Figure 3b). However, in the other populations, there was no or non-significant correlation of malic and citric acids to TA. For both the BB x US89-3 and CU x US89-3 populations, there was a strong negative correlation between berry weight and quinic acid concentration (r= -0.50 and -0.60, with $p < 0.005$ and $p < 0.0005$, respectively) (Figures 2c and 2d).

QTL analysis. In the BB x CQ population a total of 53 QTLs were detected (Table 2A). For the organic acids, nineteen QTLs were detected on the Budd's Blues parental map and 18 QTLs on the Crimson Queen parental map. For Tacy, brix, and TA, nine and seven QTLs were detected on the BB and the CQ maps, respectively. In the BB x CU population, a total of 36 QTLs were detected (Table 2B). For organic acids, 12 QTLs and 11 QTLs were detected on the BB and CU maps respectively. For Tacy, brix, and TA, 5 QTLs were detected on BB, and 8 QTLs were detected on CU. In the BB x US89-3 population, fewer QTLs were detected, a total of 9 QTLs were detected for organic acids and 13 for TA, brix, and TAcy (Table 2C). In the CU x US89-3 population, a total of 11 QTLs were detected for organic acids and 9 for TA and brix (Table 2D).

Six QTLs were identified that were consistent from year-to-year for the BB x CQ population (Table 2A, blue highlight). On chromosome 9: marker 36, a citric acid QTL was identified in Crimson Queen for both July harvest dates. It has a LOD score of 6.7 and

5.4 for July 2014 and 2015. The proportion of phenotypic variance accounted for by this QTL was 23.3% and 16.9%. A TA QTL on chromosome 11: markers 9 and 10 in Budd's Blues was found for both September harvest dates (September 2014 and 2015) with a LOD score of 4.4 and 4.7, respectively. The proportion of phenotypic variance accounted for by this QTL was 14.9% and 17.4%. For Tacy, there were two QTL on chromosome 7: marker 59, 76, and 74 in Crimson Queen for the September harvest dates, with a LOD score of 7.66 and 5.95, and 4.87 and 4.64, for 2014 and 2015. The proportion of phenotypic variance accounted for by this QTL was 24.97% and 18.55%, and 16.21% and 14.82%, respectively. In addition, there were two malic acid QTL that were consistent during two different harvest dates from year-to-year.

Discussion

Fruit organic acid levels in the parents and standard varieties over the 9 sampling dates (fruit development and ripening) followed similar trends as previously reported in Wang et al. (2017). Quinic acid decreased from fruit set on, while benzoic acid began to increase after early fruit development. Malic acid increased from fruit set on. Overall, the general trend for citric acid was to decrease in Wang et al. (2017), while in the present study, citric acid increased until the fourth sampling date, around mid-season, and decreased slowly for the rest of the season. In the present study, the first sampling date was July 14th, 5 days earlier than the first harvest date in Wang et al. (2017). The earlier sampling date in the present study allowed observation of citric acid concentration earlier in fruit development. During early fruit development, citric acid increases due to increased respiration (Etienne et al., 2013). Then, as fruit ripens, citrate is consumed through gluconogensis (Etienne et al., 2013). Also, similar to Wang et. al. (2017), varieties

exhibited differential temporal variation across fruit development and maturity, with varieties exhibiting a significant cultivar by harvest date interaction. In this study, we found that overall, there is considerable year-to-year variation in organic acids, as found in other fruit QTL studies in grape and peach (Etienne et al., 2002; Chen et al., 2015). This is likely because there are environmental effects influencing organic acid production and metabolism in cranberry fruit. Transgressive segregation was observed for most of the traits measured, with progeny exhibiting both higher and lower values than their parents. Multiple QTL were identified, with variation across fruit development and some QTL that were consistent across years. The differential responses of genotypes across fruit development for certain acids, e.g. citric, likely reflect variable genetic components responding differently to environmental cues. Transgressive segregation is observed because the parents in this study are likely heterozygous (Vorsa, unpublished). Cultivars derived from hybridization of genetically diverse parents, such as Crimson Queen, would be expected to be relatively highly heterozygous, whereas, wild selections, Cumberland, US89-3 and Budd's Blues may have some level of inbreeding. Segregation data indicates the wild selection parents used in this study do have some level of heterozygosity. However, even mild inbreeding may have fixed various chromosomal regions. In addition, differential parental recombination rates might be expected to affect segregation as well. Daverdin et al. (2017) found that parental genetic map lengths varied: CU had genetic map lengths of >2,200cM, whereas, BB genetic map lengths were <1,200cM. Lengths of genetic maps would be expected to be reduced with increasing homozygosity.

Phenotypic correlations between TAcy and berry weight were negatively correlated in the progeny of the crosses. This is expected as anthocyanins are primarily found in the

epidermis of cranberry (Vorsa and Welker, 1985). Larger fruit, measured by average fruit weight, would have less surface area of epidermis to flesh, yielding lower TAcy values. Benzoic acid is also primarily found in the epicuticular wax of cranberry fruit, explaining its negative correlation with fruit weight (Coroteau and Fagerson, 1971). Malic acid was positively correlated with TA both years in BB x CU and BB x US89-3. Citric acid was also positively correlated with TA in BB x CU. This is expected because TA is directly affected by malic and citric acid (Vorsa et al., unpublished data). Although there was no correlation between organic acids and fruit rot resistance in the segregating populations, there were significant differences in organic acid levels between the parents, specifically between the fruit rot resistant germplasm and the commercial cultivars (data not shown).

In this analysis, a total of 134 QTLs were detected across 4 populations and 7 traits. A few QTLs and the phenotypic variation were found to be consistent across the two years, suggesting that there is a genetic component influencing organic acids, brix, and TA, but with major environmental influence. The detection of multiple QTLs for organic acids across different chromosomes indicate that these traits are also likely controlled by multiple genes. There were 6 QTLs that were consistent year-to-year. Despite a consistent QTL year-to-year for TA found in Budd's Blues in the BBxCQ population, this QTL was not observed in the other Budd's Blues populations (BBxCU and BBxUS89-3). Reduced recombination in CU, along with the smaller BBxCU and BBxUS89-3 population sizes, would result in lower segregation power, and result in the lack of QTL detection (Daverdin et al., 2017). Polygenic control of organic acids and TA would explain the transgressive segregation observed. It is interesting to note that there were citric and malic acid QTL identified that were not associated with the TA QTL found. TA reflects a sum of all acids.

So, variation in TA may be muted by levels of individual acids fluctuating relative to each other. The correlation of TA with individual acids, quinic, malic, and citric acids, were all less than r=0.4, which suggests that no one acid contributes to TA completely.

There were some consistencies in the TAcy and TA QTL of this study in comparison with Georgi et al. (2013). Georgi et al. (2013) reported two QTL associated with TAcy (Vm7: scf142e and an unplaced marker near Vm8, SCAR0910); and three QTL for TA (Vm3: scf2s, Vm12: vccj1d, and Vm11: vcc1a). For TAcy, the scf142e marker is not placed on the current genetic map, however, the SNP markers found at Vm7 is 30 cM from the TAcy QTL that encloses the scf142e region. For TA, the vcc1a marker is on the same chromosome as our SNP TA QTL on Budd's Blues Vm11. However, the vcc1a marker is not placed on the current genetic map and the closest marker is more than 50 cM away. A recent QTL study by Diaz-Garcia et al. (2018) reported a QTL on VM7 (LG3 in Schlautman et al., 2015) for TAcy and a QTL for TA on VM11 (LG11 in Schlautman et al., 2015) that match with the QTL identified in the current study. Interestingly, Diaz-Garcia identified a major QTL, with segregation between red and yellow fruit, on VM11. This indicates that there is a major gene controlling TAcy, with other quantitative controls of TAcy located in other linkage groups.

In the BBxUS89-3 cross, our SNP TAcy QTL on Vm8, marker 11, on the Budd's Blues map, is 1 cM from the Daverdin et al (2017) QTL for yield. This might be explained by the fact that TAcy is negatively correlated with fruit size, and fruit size is an important contributor to yield. In the BBxCU cross, on the CU map, the malic acid QTL on Vm1, marker 87, corresponds to a QTL for fruit rot rating (Daverdin et al., 2017). Another SNP QTL for malic acid is on Vm, marker 41, 1 on the Budd's Blue's (BBxCQ) map, and is 1

cM from the QTL for fruit rot rating identified in Daverdin et al. (2017). The fruit rot rating QTL coincides with malic acid QTLs in two different fruit rot resistant genotypes indicating that there could be a direct or indirect (linkage drag) relationship between fruit rot resistance and malic acid that should be further explored. Tadych et al. (2015) reported malic acid to inhibit ROS production by fruit rotting cranberry pathogens. While there are some QTL found in the present study that coincide with QTL identified for fruit rot resistance and yield, there remains many more that are not. Thus, there is not necessarily a detrimental selection against good fruit quality while selecting for fruit rot resistance. Further breeding efforts can use these unlinked QTL and potentially combine fruit rot resistance, yield, and phytochemical properties into a single variety.

One of the weaknesses of QTL mapping is the verification of QTL in different populations. In this study, we found four QTLs that were consistent year-to-year (citric acid, malic acid, TA, and TAcy QTLs), as well as ones that were found in prior QTL studies (TA and TAcy). This study is the first QTL study of individual organic acids in cranberry fruit. From the stable QTLs we have identified, molecular markers can be designed for marker assisted selection and used in future breeding efforts to enhance not only disease resistance and yield, but also fruit quality and phytochemistry. These QTLs will help us select fruit with improved phytochemistry, for both human health and palatability, and further studies will help evaluate the validity and use of other QTL detected.

References

Baardseth, P., Russwurm, H., 1978. Content of some organic acids in cloudberry (Rubus chamaemorus L.). Food Chemistry 3, 43–46. https://doi.org/10.1016/0308-8146(78)90046-8

Bates, R.P., Morris, J.R. and Crandall, P.G., 2001. *Principles and practices of small-and medium-scale fruit juice processing* (No. 146). Food & Agriculture Org.

Chen, J., Wang, N., Fang, L.-C., Liang, Z.-C., Li, S.-H., Wu, B.-H., 2015. Construction of a high-density genetic map and QTLs mapping for sugars and acids in grape berries. BMC Plant Biology 15, 28. https://doi.org/10.1186/s12870-015-0428-2

Croteau, R., Fagerson, I.S., 1971. Volatile substances derived from the cuticle wax of cranberry. Phytochemistry 10, 3247–3249. https://doi.org/10.1016/S0031-9422(00)97380-1

Cunningham, D.G., Vannozzi, S.A., Turk, R., Roderick, R., O'Shea, E., Brilliant, K., 2003. Cranberry Phytochemicals and Their Health Benefits, in: Nutraceutical Beverages, ACS Symposium Series. American Chemical Society, pp. 35–51. https://doi.org/10.1021/bk-2004-0871.ch004

Daverdin, G., Johnson-Cicalese, J., Zalapa, J., Vorsa, N., Polashock, J., 2017. Identification and mapping of fruit rot resistance QTL in American cranberry using GBS. Mol Breeding 37, 38. https://doi.org/10.1007/s11032-017-0639-3

Diaz-Garcia, L., Schlautman, B., Covarrubias-Pazaran, G., Maule, A., Johnson-Cicalese, J., Grygleski, E., Vorsa, N., Zalapa, J., 2018. Massive phenotyping of multiple cranberry populations reveals novel QTLs for fruit anthocyanin content and other important chemical traits. Mol Genet Genomics 293, 1379–1392. https://doi.org/10.1007/s00438-018-1464-z

Ehling, S., Cole, S., 2011. Analysis of Organic Acids in Fruit Juices by Liquid Chromatography−Mass Spectrometry: An Enhanced Tool for Authenticity Testing. J. Agric. Food Chem. 59, 2229–2234. https://doi.org/10.1021/jf104527e

Gallardo, R.K., Klingthong, P., Zhang, Q., Polashock, J., Atucha, A., Zalapa, J., Rodriguez-Saona, C., Vorsa, N., Iorizzo, M., 2018. Breeding Trait Priorities of

the Cranberry Industry in the United States and Canada. HortScience 53, 1467–1474. https://doi.org/10.21273/HORTSCI13219-18

Georgi, L., Johnson-Cicalese, J., Honig, J., Das, S.P., Rajah, V.D., Bhattacharya, D., Bassil, N., Rowland, L.J., Polashock, J., Vorsa, N., 2013. The first genetic map of the American cranberry: exploration of synteny conservation and quantitative trait loci. Theor Appl Genet 126, 673–692. https://doi.org/10.1007/s00122-012-2010-8

Howell, A.B., 2007. Bioactive compounds in cranberries and their role in prevention of urinary tract infections. Molecular Nutrition & Food Research 51, 732–737. https://doi.org/10.1002/mnfr.200700038

Johnson-Cicalese, J., Polashock, J.J., Honig, J.A., Vaiciunas, J., Ward, D.L., Vorsa, N., 2015. Heritability of Fruit Rot Resistance in American Cranberry. Journal of the American Society for Horticultural Science 140, 233–242. https://doi.org/10.21273/JASHS.140.3.233

Kowalska, K., Olejnik, A., 2016. Beneficial effects of cranberry in the prevention of obesity and related complications: Metabolic syndrome and diabetes – A review. Journal of Functional Foods 20, 171–181. https://doi.org/10.1016/j.jff.2015.11.001

Lustig, R.H., Schmidt, L.A., Brindis, C.D., 2012. Public health: The toxic truth about sugar. Nature 482, 27–29. https://doi.org/10.1038/482027a

Marsh, K.B., Boldingh, H.L., Shilton, R.S., Laing, W.A., 2009. Changes in quinic acid metabolism during fruit development in three kiwifruit species. Functional Plant Biology 36, 463. https://doi.org/10.1071/FP08240

Marwan, A.G., Nagel, C.W., 1986. Characterization of Cranberry Benzoates and Their Antimicrobial Properties. Journal of Food Science 51, 1069–1070. https://doi.org/10.1111/j.1365-2621.1986.tb11236.x

McCamey, D.A., Thrope, T.M., and McCarthy, J.P. 1990. Coffee Bitterness. p. 169-182. In: R.L. Rouseff (ed.). Bitterness in Foods and Beverages. Elsevier Science Publishers B.V. Amsterdam, The Netherlands.

McCown, B.H. and Zeldin, E.L., 2003. 'HyRed', an Early, High Fruit Color Cranberry Hybrid. *HortScience*, *38*(2), pp.304-305.

Pappas, E., Schaich, K.M., 2009. Phytochemicals of Cranberries and Cranberry Products: Characterization, Potential Health Effects, and Processing Stability. Critical Reviews in Food Science and Nutrition 49, 741–781. https://doi.org/10.1080/10408390802145377

Polashock, J., Zelzion, E., Fajardo, D., Zalapa, J., Georgi, L., Bhattacharya, D., Vorsa, N., 2014. The American cranberry: first insights into the whole genome of a species adapted to bog habitat. BMC Plant Biology 14, 165. https://doi.org/10.1186/1471-2229-14-165

Rubico, S.M., McDaniel, M.R., 1992. Sensory evaluation of acids by free-choice profiling. Chemical Senses 17, 273–289. https://doi.org/10.1093/chemse/17.3.273

Shabrova, E.V., Tarnopolsky, O., Singh, A.P., Plutzky, J., Vorsa, N., Quadro, L., 2011. Insights into the Molecular Mechanisms of the Anti-Atherogenic Actions of Flavonoids in Normal and Obese Mice. PLoS One 6. https://doi.org/10.1371/journal.pone.0024634

Sheng, Y., Åkesson, C., Holmgren, K., Bryngelsson, C., Giamapa, V., Pero, R.W., 2005. An active ingredient of Cat's Claw water extracts: Identification and efficacy of quinic acid. Journal of Ethnopharmacology 96, 577–584. https://doi.org/10.1016/j.jep.2004.10.002

Tadych, M., Vorsa, N., Wang, Y., Bergen, M.S., Johnson-Cicalese, J., Polashock, J.J., White, J.F.J., 2015. Interactions between cranberries and fungi: the proposed function of organic acids in virulence suppression of fruit rot fungi. Front. Microbiol. 6. https://doi.org/10.3389/fmicb.2015.00835

Vorsa, N., 2012. Cranberry plant named 'CNJ95-20-20.' USPP22541P3.

Vorsa, N., 2008. Cranberry variety named 'NJS98-35.' USPP18911P2.

Vorsa, N., 2007. Cranberry variety named 'NJS98-23.' USPP18252P3.

Vorsa, N., Johnson-Cicalese, J., 2012. American Cranberry, in: Badenes, M.L., Byrne, D.H. (Eds.), Fruit Breeding. Springer US, Boston, MA, pp. 191–223. https://doi.org/10.1007/978-1-4419-0763-9_6

Vorsa, N., Welker, W.V., 1985. Relationship between fruit size and extractable anthocyanin content in cranberry. HortScience 20:402-403.

Walker, R.P., Famiani, F., 2018. Organic Acids in Fruits, in: Horticultural Reviews. John Wiley & Sons, Ltd, pp. 371–430. https://doi.org/10.1002/9781119431077.ch8

Wang, Y., Johnson-Cicalese, J., Singh, A.P., Vorsa, N., 2017. Characterization and quantification of flavonoids and organic acids over fruit development in American cranberry (Vaccinium macrocarpon) cultivars using HPLC and APCI-MS/MS. Plant Science 262, 91–102. https://doi.org/10.1016/j.plantsci.2017.06.004

Zdepski, A., Debnath, S.C., Howell, A., Polashock, J., Oudemans, P., Vorsa, N., Michael, T.P., 2016. Cranberry, in: Genetics, Genomics and Breeding of Berries. CRC Press, pp. 41–63.

Table 2.1. TAcy (mg/100g fresh fruit), brix (% soluble solids), %TA, organic acid levels (mg/g fruit), and fruit weight in parents and progeny of four breeding populations. BB x CQ and BB x CU data from October 2015, and BB x US89-3 and CU x US89-3 data from October 2016. Significant differences between parents indicated by *($p < 0.05$).

Cross	Trait	Parent 1	Parent 2	Progeny Mean	Std Dev	Min	Max
BB x CQ	*Tacy**	23.50	31.00	30.4	9.5	10	75
October 2015	*Brix*	9.79	8.91	8.25	0.95	6.2	10.8
n=90	*TA*	2.30	2.25	2.15	0.2	1.76	2.84
	*Quinic**	13.06	13.36	14.6	2.78	7.02	22.83
	*Citric**	7.34	6.57	7.12	1.85	3.24	12.82
	Malic	4.54	5.68	4.65	0.9	2.12	6.64
	*Benzoic**	0.13	0.04	0.079	0.034	0.02	0.23
	*Fruit Weight**	1.50	2.20	1.66	0.27	1	2.3
BB x CU	*Tacy*	23.50	14.00	18.25	6.91	5	45
October 2015	*Brix*	9.79	9.35	8.68	0.88	5.72	10.56
n=98	*TA**	2.30	2.19	2.15	0.26	1.52	2.82
	Quinic	13.06	13.44	16.37	2.68	9.68	24.31
	Citric	7.34	7.60	7.74	1.37	4.5	12.02
	Malic	4.54	3.35	4.02	0.58	2.32	5.6
	*Benzoic**	0.13	0.10	0.11	0.038	0.04	0.25
	Fruit Weight	1.50	1.70	1.33	0.19	0.69	1.78
BB x US89-3	*Tacy**	36.00	46.50	31.12	9.44	19	53
October 2016	*Brix**	9.02	8.58	8.99	0.67	7.48	10.12
n=43	*TA*	2.48	2.82	2.45	0.18	2.02	2.8
	Quinic	21.74	25.88	23.51	3.16	14.91	28.91
	Citric	9.74	10.95	11.28	2.34	7.74	17.24
	*Malic**	4.41	5.79	4.62	0.86	3.22	7.02
	*Benzoic**	0.06	0.07	0.047	0.017	0.02	0.08
	*Fruit Weight**	1.50	0.80	1	0.23	0.62	1.72
CU x US89-3	*Tacy**	14.00	46.50	20.56	7.57	8	45
October 2016	*Brix**	8.91	8.58	8.18	0.61	7.04	9.9
n=39	*TA**	2.23	2.82	2.37	0.19	1.91	2.69
	*Quinic**	19.52	25.88	21.59	3.59	13.69	31.14
	*Citric**	11.36	10.95	11.12	2.05	6.77	15.72
	*Malic**	4.87	5.79	4.53	0.79	2.77	6.18
	*Benzoic**	0.09	0.07	0.031	0.012	0.01	0.06
	*Fruit Weight**	1.70	0.80	1.3	0.46	0.7	2.63

Table 2.2. QTL identified in four breeding populations. QTL were listed if LOD scores were > 3. Marker is the marker number on the linkage group, as reported in Daverdin et al. (2017). R^2 is the % of phenotypic variance explained by that QTL, TR^2 is the "proportion of variance explained by a QTL conditioned on the cofactors in composite interval mapping". Blue highlights are QTL that are consistent year-to-year. **A)** BBxCQ, **B)** BBxCU, **C)** BBxUS89-3, **D)** CUxUS89-3. (Fig. 3C and 3D QTL are all from October 2016 samples. Note: The US89-3 parental map contain a translocation event between linkage groups 5 and 6, noted by T5/6-1 and T5/6-2. The CU parental map in CUxUS89-3 has considerably more linkage groups than the other parental maps; synteny analysis was able to resolve the 23 linkage groups into the expected 12 (Daverdin et al., 2017). These are denoted by "_1" or "_2" after the linkage group number.

A) Map	**Budd's Blues**									**Crimson Queen**							
Traits	Year	LG	Marker	Position (cM)	LOD	Additive	R^2	TR^2		Year	LG	Marker	Position (cM)	LOD	Additive	R^2	TR^2
Quinic	July 2014	6	25	25.7	3.86	-2.27	0.13	0.44		July 2014	6	62	33.1	4.66	-2.77	0.19	0.45
		9	63	71.4	3.62	-2.25	0.11	0.44		Oct 2014	3	105	83.2	4.12	-2.48	0.16	0.32
	Oct 2015	6	22	19.7	4.32	2.28	0.15	0.38		Oct 2015	3	102	82.4	4.11	3.78	0.18	0.37
		10	71	56	3.19	1.94	0.11	0.37									
Citric	July 2014	10	92	70.7	7.34	0.13	0.26	0.44		July 2014	5	121	81.9	5.27	0.10	0.18	0.44
	Oct 2014	11	10	6.5	3.87	-1.46	0.15	0.37			7	9	2.3	3.46	0.05	0.11	0.44
	July 2015	11	20	12.4	5.01	-0.05	0.17	0.42			9	36	13.2	6.67	0.08	0.23	0.44
		11	21	18.7	4.04	0.04	0.14	0.39		Oct 2014	11	28	18.3	5.06	-1.95	0.18	0.38
	Oct 2015	2	42	45.4	3.92	0.25	0.13	0.40		July 2015	9	36	13.2	5.36	0.05	0.17	0.47
		6	52	51.1	3.99	0.28	0.14	0.38			9	104	56.4	5.38	-0.05	0.17	0.47
		6	71	59.6	4.74	0.31	0.17	0.40		Oct 2015	12	6	4.3	2.93	0.23	0.11	0.31
		9	38	38.6	3.17	-0.24	0.12	0.35									
Malic	July 2014	2	35	34.2	3.63	0.05	0.14	0.39		Oct 2014	6	41	20.9	4.29	0.95	0.15	0.41
	Oct 2014	1	41	35.8	4.45	-0.94	0.17	0.35			6	79	50.4	4.15	-0.93	0.14	0.41
		1	60	45.4	6.31	1.13	0.23	0.42		July 2015	6	41	20.9	4.38	0.05	0.15	0.40
		1	106	77.6	3.21	0.89	0.09	0.50		Oct 2015	10	51	24.3	5.67	0.85	0.21	0.42
	July 2015	1	41	35.8	4.87	-0.05	0.17	0.41			11	60	33	3.11	0.91	0.11	0.39
		1	60	45.4	6.24	0.06	0.21	0.46									
		2	37	39.1	4.72	0.05	0.16	0.46									
	Oct 2015	1	122	101.9	4.62	-0.85	0.17	0.38									
Benzoic	July 2015	1	98	66.6	4.71	0.01	0.17	0.38		July 2014	7	100	54.6	2.96	0.00	0.05	0.72
											9	160	84.9	6.58	0.00	0.27	0.40
											9	176	99.2	4.84	0.00	0.21	0.40
										July 2015	7	78	42.5	4.50	-0.02	0.15	0.42
										Oct 2015	5	68	46.4	6.28	-0.09	0.22	0.42
TA	Sept 2014	11	10	4.5	4.38	-0.25	0.15	0.41		Oct 2015	12	10	5.5	4.65	0.22	0.17	0.35
	Oct 2014	2	51	53	3.45	0.02	0.11	0.43			12	43	27.3	5.89	-0.29	0.24	0.32
		5	41	32.2	4.20	0.03	0.14	0.43									
		11	15	9.8	3.77	0.03	0.14	0.43									
	Sept 2015	11	9	3.9	4.73	-0.18	0.17	0.37									
	Oct 2015	11	4	1.1	3.43	-0.14	0.12	0.39									
		12	82	66.5	3.53	0.22	0.12	0.38									
TAcy	Sept 2015	3	75	54.7	4.44	-0.14	0.16	0.39		Sept 2014	7	59	30	7.66	-0.17	0.25	0.51
		9	89	106.4	6.79	-0.36	0.26	0.40			7	76	39.7	5.95	-0.15	0.20	0.46
										Sept 2015	7	59	30	4.87	-0.14	0.16	0.46
											7	74	38.1	4.64	0.14	0.15	0.45
										Oct 2015	7	90	47.4	4.17	-0.14	0.18	0.28
Brix	Oct 2015	11	54	54.4	3.26	1.14	0.14	0.25		Sept 2014	7	26	3.4	3.04	0.17	0.10	0.40
										Sept 2015	9	37	15.6	4.23	0.98	0.17	0.31

B) Map	**Budd's Blues**									**Cumberland**							
Traits	Year	LG	Marker	Position (cM)	LOD	Additive	R^2	TR^2		Year	LG	Marker	Position (cM)	LOD	Additive	R^2	TR^2
Quinic										Oct 2014	2	15	30.2	3.13	-1.29	0.10	0.36
											7	49	84.7	3.44	-1.40	0.11	0.36
										Jul 2015	9	82	149.2	3.71	0.04	0.14	0.34
Citric	Oct 2015	8	70	68.6	3.13	-1.33	0.12	0.28		Jul 2015	3	90	159	3.71	-0.95	0.13	0.37
	Aug 2016	12	42	24.2	2.95	0.24	0.09	0.39		Aug 2016	10	42	65.8	6.19	0.31	0.19	0.45
		12	85	59.1	5.17	0.32	0.17	0.38									
Malic	Oct 2014	6	24	38	6.74	-0.11	0.43	0.65		Oct 2014	12_2	34	62.2	5.26	-0.10	0.18	0.35
	Jul 2015	4	57	70	4.30	-0.04	0.14	0.38		Jul 2015	2	37	68.3	2.98	0.03	0.09	0.40
		7	59	43.1	3.58	-0.03	0.12	0.37			12_2	5	7.6	3.08	0.03	0.11	0.37
	Aug 2016	9	44	79.1	3.94	0.70	0.15	0.25		Aug 2016	1	87	201.7	3.97	0.46	0.11	0.48
											12_2	4	5	8.62	-0.85	0.39	0.64
Benzoic	Oct 2015	9	45	81.4	3.47	-0.05	0.12	0.33		Jul 2015	5	53	119.2	3.35	-0.26	0.12	0.35
	Jul 2015	4	31	42.8	3.24	-0.31	0.12	0.29			5	92	166.2	4.10	0.29	0.15	0.32
		4	37	62.7	3.90	-0.34	0.15	0.29									
	Oct 2015	11	81	73.1	3.77	0.11	0.15	0.30									
	Aug 2016	9	63	99	3.51	0.00	0.16	0.32									
TA	Oct 2014	3	95	100.7	3.06	-0.34	0.13	0.17		Oct 2014	1	89	207.6	3.13	0.27	0.12	0.26
	Oct 2015	1	34	40.9	3.40	-0.27	0.12	0.29		Sept 2015	9	109	193.4	3.78	0.12	0.14	0.34
										Oct 2015	8	51	129	4.65	0.21	0.16	0.36
Tacy	Oct 2014	4	12	11.7	3.85	-0.15	0.13	0.35		Oct 2014	11	85	94	7.11	-0.21	0.26	0.42
											11	90	112.2	7.18	-0.20	0.25	0.40
										Sept 2015	8	4	5	3.74	0.24	0.25	0.42
Brix	Sept 2015	7	23	8.1	3.55	0.62	0.13	0.27		Sept 2015	12_2	40	72.6	3.97	-0.62	0.13	0.40
		10	80	66.4	4.17	1.05	0.30	0.51		Oct 2015	1	22	39.4	3.85	0.64	0.14	0.34

C) Map	**Budd's Blues**							**US89-3**						
Traits	LG	Marker	Position (cM)	LOD	Additive	R^2	TR^2	LG	Marker	Position (cM)	LOD	Additive	R^2	TR^2
Quinic								11	69	42.6	4.46	-3.69	0.22	0.64
Citric	8	21	33.9	5.81	3.39	0.36	0.61	12	92	76	4.20	3.75	0.28	0.52
Malic	10	23	32.1	5.21	-1.00	0.28	0.63	1	81	69.5	4.96	0.95	0.23	0.67
	10	42	40.6	5.38	-0.95	0.28	0.64	8	88	74.3	7.37	-1.51	0.48	0.60
								10	24	19.4	5.12	-1.01	0.24	0.67
Benzoic								1	39	46.7	4.83	-0.02	0.35	0.82
TA	8	57	73.3	7.00	-0.19	0.26	0.77	3	12	7.8	3.69	0.18	0.19	0.62
	9	79	107.8	6.63	0.19	0.25	0.79	T5/6_1	84	37.9	3.75	-0.17	0.21	0.62
	12	66	96.3	4.98	-0.17	0.16	0.77							
Tacy	8	11	17.7	3.72	-0.26	0.44	0.67	1	20	14.8	5.52	-0.30	0.19	0.76
								1	67	55.3	3.96	-0.15	0.13	0.76
								12	84	60	3.50	-0.14	0.11	0.76
Brix	8	4	1	3.53	-0.98	0.52	0.72	T5/6_1	94	38.5	3.25	1.09	0.32	0.54
								T5/6_1	119	43.7	5.15	1.03	0.41	0.63
								7	88	83.2	3.34	-0.66	0.18	0.59

D) Map	**Cumberland**								**US89-3**						
Traits	LG	Marker	Position (cM)	LOD	Additive	R^2	TR^2		LG	Marker	Position (cM)	LOD	Additive	R^2	TR^2
Quinic	2_2	18	24.1	3.77	3.58	0.19	0.66		2	65	60	3.31	-3.39	0.18	0.64
	11_1	15	38.4	3.32	3.20	0.16	0.66		10	46	63.8	3.44	-3.76	0.23	0.54
									10	52	75.5	5.44	4.39	0.34	0.64
Citric	5_1	17	56.5	4.09	-1.92	0.20	0.68		10	9	7.8	3.78	2.23	0.26	0.57
Malic	7	8	19.8	6.14	1.01	0.34	0.68		7_1	39	23.5	4.23	0.95	0.23	0.64
									10	12	12	5.25	1.03	0.31	0.64
									10	20	30.3	4.02	1.01	0.36	0.91
Benzoic	1_2	4	1	5.04	0.01	0.40	0.87		T5/6_1	117	69.4	4.51	-0.01	0.25	0.69
	7	18	31.8	6.43	-0.02	0.36	0.69								
TA	11_1	61	110	5.57	-0.26	0.38	0.85		2	53	48.7	4.84	0.21	0.17	0.78
									2	83	68.9	6.72	-0.26	0.31	0.75
									8	86	91.1	4.83	0.30	0.43	0.88
									8	88	102.9	5.31	0.31	0.22	0.75
									12	38	33.6	4.10	-0.16	0.16	0.75
Brix	8_2	41	108.9	5.91	-0.71	0.23	0.77		1	94	108.5	5.20	-0.75	0.30	0.66
									11	82	71.7	3.69	-0.56	0.18	0.67

Figure 2.1. Organic acid levels throughout fruit development in standard varieties Stevens and Mullica Queen, and parents Budd's Blues, Cumberland, US89-3, and Crimson Queen. All organic acid concentrations are measured in mg/g fresh fruit (means of 3 replications per individual). Fruit was collected in 2014 from just after petal drop through fully developed fruit. * represents dates with significant genotypic variation ($p < 0.05$). A) Quinic Acid, B) Citric Acid, C) Malic Acid, D) Benzoic Acid.

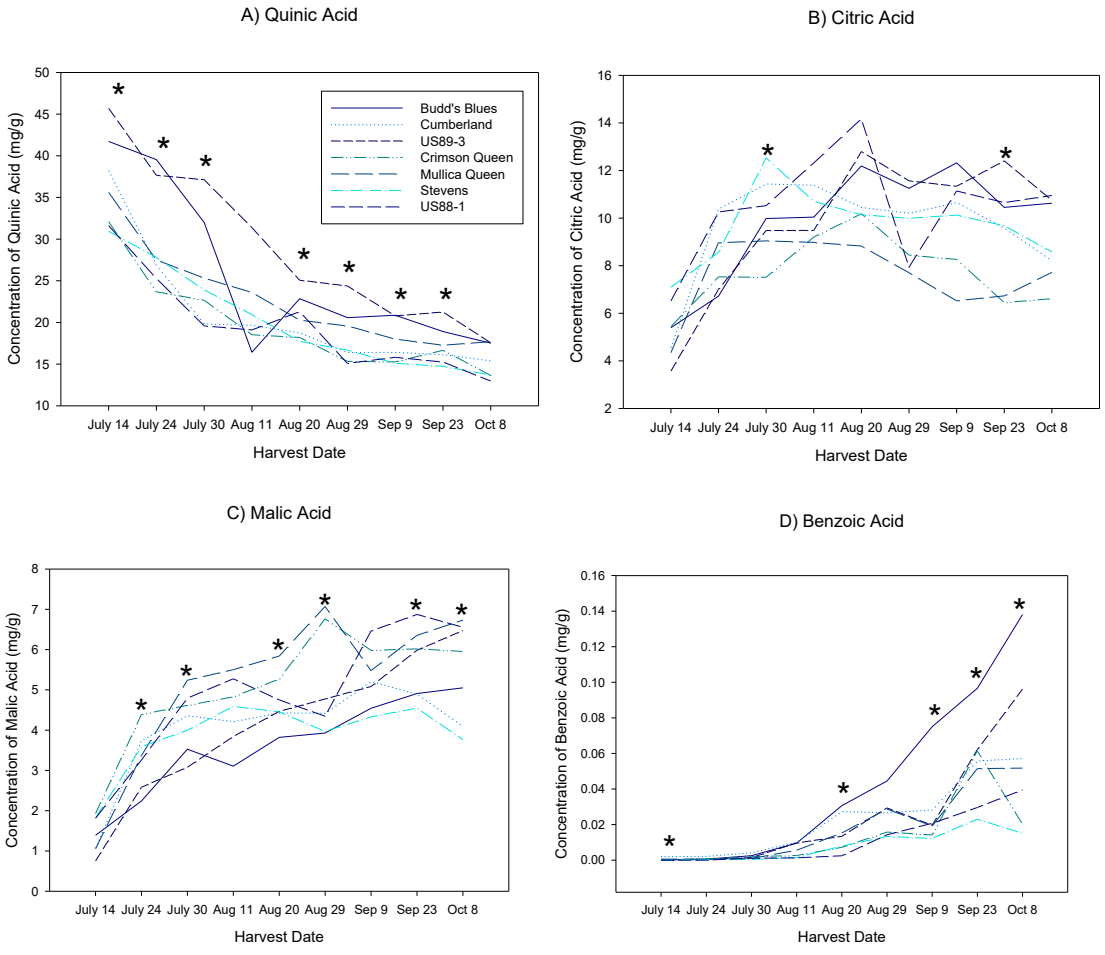
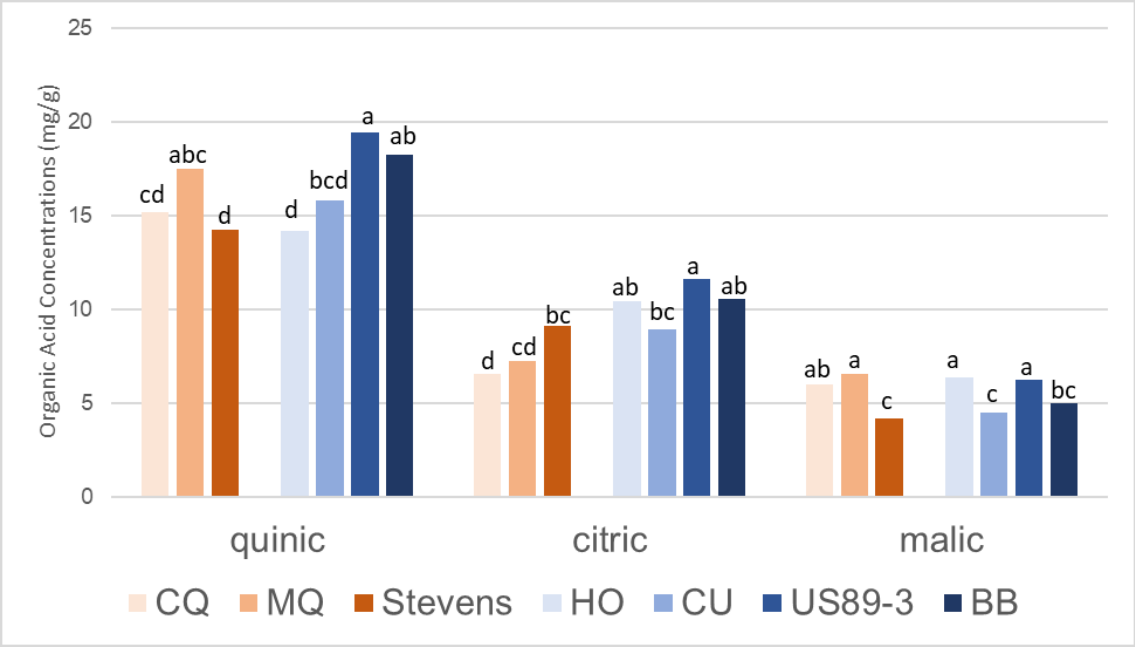


Figure 2.2. Average organic acid levels during cranberry harvest period (mean of September 23rd and October 8th samples). Orange grouped bars are varieties susceptible to fruit rot, while blue grouped bars are varieties resistant to fruit rot. LSD with alpha=.05 was used to determine the means separation for each group. A) Quinic, citric, and malic acids. B) Benzoic acid.

A)



B)

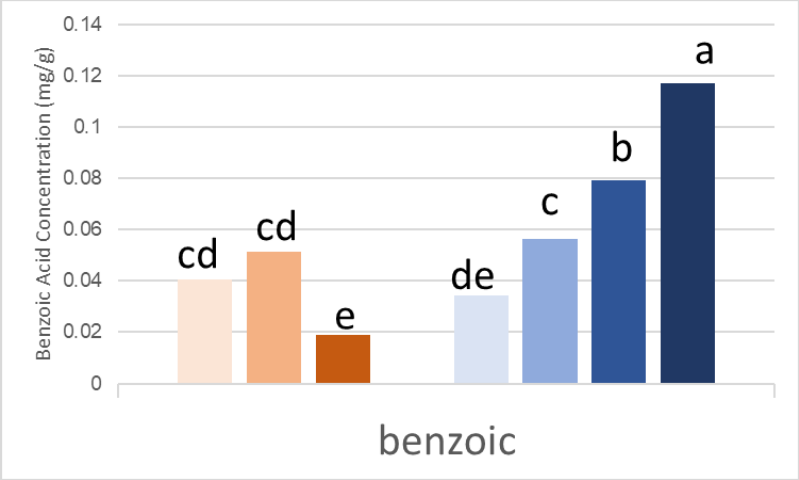
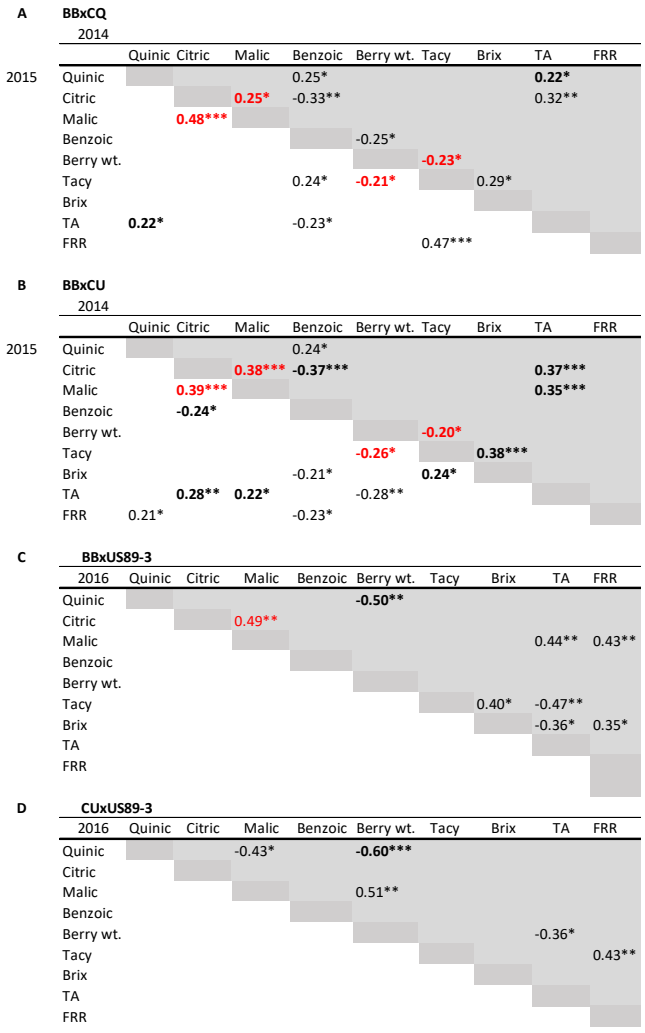


Figure 2.3. Correlation matrices of fruit chemistry traits (quinic, malic, citric, and benzoic acids, TAcy, brix, TA, and fruit rot resistance measured by percent rot) A) BB x CQ B) BB x CU. Fig. 2A and 2B, grey section is from October 2014 and white section is from October 2015; bolded text is for correlations identified in both years; and red font is for correlations identified in both BB x CU and BB x CQ. C) BB x US89-3 D) CU x US89-3. Data is from October 2016. Bolded text is correlation identified in both US89-3 half-sib crosses and red text is a correlation also identified in the other half-sib BB crosses. Significance levels are as follows: * p < .05, ** p < .005, and *** p < .0005.



Chapter 3: Genetics and Molecular Mapping of a Low Citric Acid Trait in Cranberry Fruit and Its Relationship to Titratable Acidity and Other Organic Acids

Introduction

Cranberry (*Vaccinium macrocarpon*, 2x=2n=24) is a temperate, woody perennial species native to North America and is a member of the Ericaceae family. US cranberry production accounts for a total utilized production of about $251 million, chiefly in Wisconsin, Massachusetts, New Jersey, Washington, and Oregon (USDA NASS, 2017). A major issue plaguing cranberry production in New Jersey, and more recently other regions of the United States as well, is fruit rot. Cranberry breeding and genetic research efforts thus far have focused on fruit rot resistance, yield, fruit size, vine vigor, and fruit quality traits, including total anthocyanins and soluble solids (Vorsa and Johnson-Cicalese, 2012; Johnson-Cicalese et al., 2015; Daverdin et al., 2017; Diaz-Garcia et al., 2018).

In recent years, increased health concerns have emerged regarding 'added sugar' in food products, including two popular cranberry products, cranberry juice cocktail and sweetened-dried cranberries. To balance the high acidity, cranberry products require considerable amounts of added sugar for palatability, relative to most other fruits. For product formulation, there is an optimal sugar-acid ratio, which is the ratio of sugars versus titratable acidity (TA), a measure of total acid (Bates et al., 2001). The Rutgers cranberry breeding program routinely measures TA in breeding populations, and some variation is apparent, but TA is typically greater than 2% (Vorsa and Johnson-Cicalese, personal communication). Thus, there is the potential to marginally reduce cranberry's acidity through traditional breeding and selection cycles, as a quantitative trait, with current varieties. For example, the cultivar Demoranville has a slightly lower TA at ≈2.0 (Vorsa,

2008). However, there has not been a qualitative trait for low citric acid cranberry characterized -- until the present study.

Within commonly consumed fruit crops, cranberry is unique due to having high concentrations of malic, citric, and quinic acids. Peach, *Prunus persica*, also has the same primary organic acids, but is significantly lower in citric acid upon ripening (Bae et al., 2014; Wu et al., 2012). Citric acid, the primary acid in *Citrus* spp., has been well characterized. For the low acid trait, measured by pH and citric acid concentration, molecular markers were developed using bulk segregant analysis and the inheritance of the trait in citrus was determined to be a single recessive gene (Fang et al., 1997). In addition, 3 QTLs for TA have been identified (Yu et al., 2016). Microarray data suggests that there is citrate metabolism via the gamma-aminobutyrate shunt which is linked with high acid levels (Cercos et al., 2006). Finally, there was the discovery of the Noemi bHLH transcription factor that controls both flavonoid pigments and fruit acidity in citrus (Butelli et al., 2019). In melons, the genetic mechanism for lowered acidity, primarily due to lowered citric acid concentrations, has been identified as a PIN H+/auxin transporter (Cohen et al., 2013).

Highbush blueberry (*Vaccinium corymbosum* L.), which is in the same genus as cranberry, has citric acid as the principal acid in fruit (Wang et al., 2019). However, not all blueberry species accumulate high levels of citric acid, contributing to perceptible flavor differences (Ehlenfeldt et al., 1994). Studies of acidity in cranberry fruit have identified multiple quantitative trait loci (QTL) for acidity on different chromosomes (Georgi et al., 2013, Diaz-Garcia et al., 2018, Chapter 2). In cranberry fruit, the accumulation of citric acid during fruit development shows a linear increase early in fruit development and

leveling off later in fruit development (Wang et al., 2017). However, prior research did not identify any individuals with exceptionally low citric acid.

Within the germplasm collection of cranberry at the Marucci Blueberry Cranberry Research Center in Chatsworth, NJ, an accession having a lower TA, approximately 1.5%, was identified. The low TA in this accession was due to a much lower level of citric acid, approximately 0.1%. This study is the first report and genetic characterization of a qualitatively low citric acid trait in cranberry. The objectives of this study are: 1) to characterize the genetics of the low citric acid trait, 2) to determine the relationship between low citric acid and TA, and the other major fruit acids, malic and quinic acids, and 3) to identify molecular markers linked to the trait for use in marker-assisted selection.

Materials and Methods

Plant Material

The wild accession carrying the low citric acid trait, NJ91-7-12, was collected from Burlington County, New Jersey. NJ91-7-12 was crossed to a series of commercial cultivars including Stevens, #35. Lemunyon, and Crimson Queen. These F1 hybrids expressed TA levels within the normal range (TA=2.3-2.5) for cranberry, as well as normal citric acid levels (Figure 1) and were subsequently used to develop the F_1 x F_1 populations described in this study. Three primary populations were evaluated, CNJ08-137 (CA137), CNJ08-80 (CA80), and CNJ08-64 (CA64). CA137 and CA64 are F_1 x F_1 populations while CA80 is a backcross to NJ91-7-12. The pedigrees of these crosses are described in Figure 2. These crosses were made in 2008, with seed germinated and planted into 4-inch pots in 2009. A series of additional F_1 x F_1 crosses were made in 2010, the CNJ10 series. For these smaller populations, crosses were made in 2010, with seed germinated and seedlings planted in

2011 in 4-inch pots. The pedigrees of these crosses are described in Figure 3. All of the populations were maintained in potted culture in a greenhouse, except for a period in spring/early summer in which they were taken outdoors for open pollination by bees for fruit set.

Fruit Evaluation

As the plants developed to flower and produce fruit, in 2014-16, mature red berries were collected in late August through early September. Fruit was only collected in 2014 for the CNJ10 populations. Initially, the fruit was analyzed only for organic acids [citric (CA), malic (MA), and quinic (QA) acids] whereas in 2016, the fruit was analyzed for TA and brix (soluble solids) as well.

Organic acid extraction and high performance liquid chromatography (HPLC) was done as in Wang et al. (2017). TA and brix were measured as in Vorsa and Johnson-Cicalese (2012), with modification due to smaller fruit sample sizes. Specifically, about 5 grams (2-5 fruit) of fruit was ground with a Precelly Evolution (Bertin Corp., Rockville, MD, USA) using ceramic beads at 7200 rpm. Distilled water, 40 mL, was used to dissolve the ground fruit and a 1 mL aliquot was set aside for organic acid analysis. TA was analyzed using a Metrohm Ti-Touch 916 (Metrohm AG, Riverview, FL, USA) using 0.05 N NaOH until a pH 8.2 endpoint. Brix (soluble solids) was analyzed using an ATAGO PR-32 digital refractometer (Atago USA, Inc., Bellevue, WA, USA).

Bulk Segregant Analysis with SSRs

DNA was extracted from leaf tissue using a modified CTAB method as in Daverdin et al. (2017). The parents (CNJ04-59-47 and CNJ04-30-34) of CA137 were screened for polymorphisms with 129 SSR markers from Georgi et al. (2013) and Schlautman et al.

(2015). PCR amplification of the SSRs was done on an ABI GeneAmp PCR 9700 and separation of SSR sizes was done on an ABI 3500 as in Georgi et al. (2013) (Applied Biosystems, Foster City, CA, USA). The SSR markers with polymorphisms in the parents were then used to screen high (> 2 mg/g) and low (< 2 mg/g) citric acid bulks. SSRs identified to have polymorphisms between the bulks were then used to screen individual progeny of entire populations for segregation of the low citric acid trait.

QTL Mapping using GBS

The CA137 population was further genotyped using genotyping by sequencing (GBS) to generate SNPs for fine mapping the citric acid trait. A dual enzyme restriction using Pst1 and Nde1 was used. Sequencing was completed by Genewiz, Inc. (South Plainfield, NJ, USA) using a 2x150bp output on an Illumina Hi-Seq. Sequencing data was de-multiplexed using STACKS. Reads were aligned to the cranberry reference genome of 124 contigs (Kawash et al., unpublished) and SNPs were called using bwa-mem and samtools (Li and Durbin, 2009; Li, 2011). Qualifying SNPs required a read support of four reads, and heterogeneity between 25% and 75%. Missing data was limited to only 10% of the population in a given marker, and markers that were homogenous through the population were also removed. Qualifying SNP markers were anchored in place, R/qtl was used to calculate genetic distance between markers and identify QTL.

Statistical Analyses

Segregation ratios were tested using a chi-squared test with a Yates correction factor. ANOVA was used to determine year-to-year variation. A means separation using the LSD option in PROC GLM was used to determine the significance of the genotypes.

PROC CORR in SAS was used to determine correlation between CA, MA, QA, and TA. (SAS Institute Inc., Cary, NC, USA)

Results

Segregation of Citric Acid

The frequency distribution (Figure 4A) for citric acid concentration in CA137 suggests a trimodal distribution, indicating that the *cita* allele is not fully recessive. In the CA64 population, there is an apparent quadrimodal distribution (Figure 4B), indicating that in addition to having the *cita* allele derived from NJ91-7-12 that affects citric acid concentration, there are two 'wild type' alleles, one derived from Stevens ($Cita_{st}$) and another from #35 ($Cita_{35}$), which are unequal in their influence on citric acid concentration. In population CA80, which is a F_1 (CNJ04-31-44) backcross to the low citric acid parent (NJ91-7-12), the frequency distribution appears bimodal, which would be expected if there are only two alleles segregating from CNJ04-31-44 (Figure 4C). Overall, the frequency distribution of fruit citric acid levels in these populations indicate the presence of a single locus with a low citric acid allele, which when homozygous, results in the low citric acid trait (< 2 mg/g) (Figure 4).

The observed segregation for citric acid levels in the CA64 and CA80 fit a model for a single locus recessive Mendelian trait. CA64 fit the expected segregation ratio of 1:3, low citric acid (< 2 mg/g) to normal-high citric acid (> 2 mg/g). The observed segregation of CA80 had an expected segregation ratio of 1:1, low (< 2.5 mg/g) to normal-high (> 2.5 mg/g) citric acid. The chi-squared values were not significant, indicating that observed segregation was not significantly different from the expected segregation for these two populations (Table 1). CA137, on the other hand, did not fit the expected segregation ratio

of 1:3, low citric acid (< 2 mg/g) to normal-high citric acid (> 2 mg/g), as the chi-squared value was significant at p = .035 (Table 1). In CA137, there was an underrepresentation of low citric acid individuals. These segregation ratios indicate that low citric acid is likely controlled by a single locus. The means of CA137 and CA64 were similar, with similar ranges. On the other hand, CA80 had a much lower mean, 3.7 mg/g with a smaller range. This is due to the lack of "homozygous normal" genotypes in CA80 where the trait is partially dominant/recessive (Table 1). In the smaller citric acid populations (CNJ10's), overall, there was segregation consistent with a single locus recessive Mendelian inheritance (Table 2). However, in one population (CNJ10-67) there was evidence of segregation distortion, with an overrepresentation of low citric acid individuals.

Relationship of Citric Acid to TA, Malic Acid and Quinic Acid

The CA137 and CA80 populations were both evaluated for titratable acidity (Table 1). Consistent with the citric acid data, the population mean for CA137 was higher than for CA80, 2.3% versus 1.8% TA. The TA range for CA137 (1.8% to 3.0%) was also greater than for CA80 (1.46% to 2.08%). Regression analysis shows that CA137 has a good linear fit of citric acid to TA, with an R-squared value of 0.71. CA80 also had a linear fit, but with a lower R-squared value of 0.33, likely due to the smaller range in citric acid concentration. The Pearson correlation found a significant positive correlation between CA and TA for populations CA137 and CA80 (Table 3). For CA137, there was also a significant positive correlation between TA and QA. In all three populations, there was a significant negative correlation between citric and malic acids (Table 3).

Bulk Segregant Analysis

Initial screening of the parents of CA137 (CNJ04-59-47 and CNJ04-30-34) revealed 66 polymorphic SSRs. The screening of the bulks, low versus high citric acid for CA137, identified a single polymorphic SSR (scf258d) with allele 217/217 co-segregating with the low citric acid phenotype. The SSR genotypes found in CA137 for scf258d were 217/217, 217/225, 217/229, and 225/229 (Figure 4). Scf258d is located at the end of VM9 on the initial cranberry SSR linkage map by Georgi et al. (2013). Using synteny, the end of VM9 was matched to the most recent SSR map by Schlautman et al. (2015), which placed the locus on MC1.Three more markers around the locus, scf68870, scf157322, and scf30816, from Schlautman et al. (2015) were tested for polymorphisms, and SCF157322, along with scf258d, were retained for further screening of other populations (Table 4). The 276/276 alleles for SCF157322 co-segregated for low citric acid. Table 5 shows the multiple SSR genotypes identified in each population.

In each population, possible recombinants, between the SSR markers and the acid phenotype, were identified, where the genotype was not consistent with the phenotype (Table 5). In CA64, there was one possible recombinant with low CA (< 2mg/g) and a scf258d genotype of 217/225, but had a scf157322 genotype of 276/276. There was also another individual with high CA (> 2mg/g) and a scf258d genotype of 217/217 and a scf157322 genotype of 276/276. In CA137, there were two possible recombinants with low CA, with a scf157322 genotype of 272/276, and a scf258d genotype of 217/229. There was one with low CA, and an scf258d genotype of 217/217, but a scf157322 genotype of 270/276. There were two possible recombinants with high CA, with a scf157322 genotype of 276/276 and a scf258d genotype of 217/217. In CA80, there was one recombinant with high CA and a scf157322 genotype of 276/276 and a scf258d genotype of 217/217. Overall,

there are a total of 7 putative recombinants for both of the SSR markers. The sensitivity and specificity for detecting an individual with homozygous *cita/cita* was calculated for each of the markers, shown in Table 5. Overall, scf258d had the best sensitivity at 94.5% and specificity of 97.5%, with scf157322 with a sensitivity of 94.4% and a specificity of 97.5%.

With these SSR genotypes, mean separation analysis revealed that there are significantly different phenotypic groups within each population (Figure 4). Overall, the heterozygous *cita* genotype (*Cita/cita*) had moderately lower citric acid than the homozygous *Cita* genotypes. It seemed like the parent from which the *Cita* allele came from also had an effect on the citric acid concentration. For example, the scf258d-225 allele ($Cita_{ST}$) originated from Stevens while the scf258d-229 allele ($Cita_{35}$) originated from #35. The scf258d-229 allele results in significantly lower citric acid concentrations than the scf258d-225 allele (Figure 4A). Currently three alleles have been identified for the *cita* locus having a dominance order of $Cita_{ST} > Cita_{35} > cita$.

QTL Analysis

Overall, a total of 756 million reads were obtained from sequencing. Of this, there was a mean of 6.2 million reads per individual. After SNP calling and clean up, a total of 2566 SNPs were used for QTL mapping in R/qtl. This resulted in one strong QTL, with a LOD of 21.3, that explained 57.8% of the variance (minimum LOD threshold of 6.84, Figure 6). This QTL was located on Contig 1: 48141640, with flanking markers at Contig 1:48141626 and Contig 1: 48655212.

Discussion

In this study, we focused on 3 populations of different genetic backgrounds introgressing the low acid germplasm accession NJ91-7-12 to characterize *cita* in cranberry. In these three populations, over 200 individuals were evaluated for citric acid concentrations, TA, and genotypes. The *cita* trait was found to consistent with a Mendelian inheritance of a single recessive locus through progeny testing and segregation analysis. Additionally, the *cita* alleles were co-dominant, as heterozygous carriers of the *cita* allele had significantly lower citric acid levels than non-carriers. Two SSR markers that co-segregate with the low citric acid trait were identified, scf258d - 217/217 and scf157322 - 276/276. Both of these markers have high sensitivity and specificity for detecting the *cita* trait.

However, there were some exceptions where the populations did not segregate as expected. CA137 and CNJ10-67 did not have the expected segregation ratio. For CA137, the segregation distortion could be due to lethality in some of the homozygous *cita* plants, as there were 127 seeds planted, and only 120 individuals surviving to fruit. Despite the segregation distortion, the SSRs still co-segregated for the low citric acid trait, indicating that the segregation distortion was due to both genotype and phenotype. For CNJ10-67, there were only 22 individuals evaluated even though there were 85 seedlings germinated. Thus, the segregation distortion of the overrepresentation of *cita* plants could be due to the low number of individuals sampled. Additionally, this population also seemed to segregate 1:1 for low citric acid. It is a possibility that there was cross pollination from another plant other than the intended pollen parent that would result in a different segregation pattern, which would be easily tested with DNA analysis.

The *cita* allele also has a modest co-dominance effect, as individuals with one copy of the trait display lower levels of citric acid. However, there is an inverse relationship between citric and malic acids. If a decrease in citric acid did not result in an increase in malic acid, then it is likely that the TA would be lower. Overall, with the homozygous *cita/cita* genotype, the TA can be reduced about 20%. Perhaps there is the potential for further reducing the malic acid concentrations as well to further reduce the TA.

Additionally, there was a genotypic effect indicating partial dominance of 'wild-type' alleles. In CA137, there are alleles derived from #35 and Stevens. #35 has lower citric acid concentrations than Stevens as seen in Figure 1, which then results in significantly lower citric acid concentrations in the progeny of CA137 (Figure 5A). This could be due to the presence of a citric acid QTL located on chromosome 1 about 6.5 Mb away from the SSR markers. There could be a single gene accounting for the majority of the variation, with the nearby QTL modulating some of the variation.

In Chapter 2, a QTL for citric acid is located on VM9/MC1 on the Crimson Queen map is on same chromosome near the SSR markers identified in this chapter. This QTL accounts for about 20% of the variance in citric acid in the Budd's Blues x Crimson Queen mapping population. Crimson Queen is a progeny of Stevens x Ben Lear. It is possible that the variation from the $Cita_{Stv}$ allele is due to this QTL.

Cranberry fruit have a long generation time, about 3 years from seed to fruit. Thus, there is an urgent need to employ marker assisted selection to reduce the selection time from 3 years to 1 year per generation. This also cuts down on the number of plants housed in the greenhouse and the cost of maintaining the plants. The markers identified in this study will be used to identify seedlings carrying the low citric acid allele for marker assisted

selection. These markers will enable the screening and culling of large populations so that more genetic diversity can in incorporated for breeding for low acidity.

References

Bae, H., Yun, S.K., Yoon, I.K., Nam, E.Y., Kwon, J.H. and Jun, J.H., 2014. Assessment of organic acid and sugar composition in apricot, plumcot, plum, and peach during fruit development. *Journal of Applied Botany and Food Quality*, *87*.

Bates, R.P., Morris, J.R. and Crandall, P.G., 2001. *Principles and practices of small-and medium-scale fruit juice processing* (No. 146). Food & Agriculture Org.

Broman, K.W., Wu, H., Sen, S., Churchill, G.A., 2003. R/qtl: QTL mapping in experimental crosses. Bioinformatics 19, 889–890. https://doi.org/10.1093/bioinformatics/btg112

Butelli, E., Licciardello, C., Ramadugu, C., Durand-Hulak, M., Celant, A., Reforgiato Recupero, G., Froelicher, Y., Martin, C., 2019. Noemi Controls Production of Flavonoid Pigments and Fruit Acidity and Illustrates the Domestication Routes of Modern Citrus Varieties. Current Biology 29, 158-164.e2. https://doi.org/10.1016/j.cub.2018.11.040

Catchen, J.M., Amores, A., Hohenlohe, P., Cresko, W., Postlethwait, J.H., 2011. Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences. G3: Genes, Genomes, Genetics 1, 171–182. https://doi.org/10.1534/g3.111.000240

Cercós, M., Soler, G., Iglesias, D.J., Gadea, J., Forment, J., Talón, M., 2006. Global Analysis of Gene Expression During Development and Ripening of Citrus Fruit Flesh. A Proposed Mechanism for Citric Acid Utilization. Plant Mol Biol 62, 513–527. https://doi.org/10.1007/s11103-006-9037-7

Cohen, S., Tzuri, G., Harel-Beja, R., Itkin, M., Portnoy, V., Sa'ar, U., Lev, S., Yeselson, L., Petrikov, M., Rogachev, I., Aharoni, A., Ophir, R., Tadmor, Y., Lewinsohn, E., Burger, Y., Katzir, N., Schaffer, A.A., 2012. Co-mapping studies of QTLs for fruit acidity and candidate genes of organic acid metabolism and proton transport in sweet melon (Cucumis melo L.). Theor Appl Genet 125, 343–353. https://doi.org/10.1007/s00122-012-1837-3

Daverdin, G., Johnson-Cicalese, J., Zalapa, J., Vorsa, N., Polashock, J., 2017. Identification and mapping of fruit rot resistance QTL in American cranberry using GBS. Mol Breeding 37, 38. https://doi.org/10.1007/s11032-017-0639-3

Diaz-Garcia, L., Schlautman, B., Covarrubias-Pazaran, G., Maule, A., Johnson-Cicalese, J., Grygleski, E., Vorsa, N., Zalapa, J., 2018. Massive phenotyping of multiple cranberry populations reveals novel QTLs for fruit anthocyanin content and other important chemical traits. Mol Genet Genomics 293, 1379–1392. https://doi.org/10.1007/s00438-018-1464-z

Ehlenfeldt, M.K., Meredith, F.I., Ballington, J.R., 1994. Unique Organic Acid Profile of Rabbiteye vs. Highbush Blueberries. HortScience 29, 321–323. https://doi.org/10.21273/HORTSCI.29.4.321

Fang, D.Q., Federici, C.T., Roose, M.L., 1997. Development of molecular markers linked to a gene controlling fruit acidity in citrus. Genome 40, 841–849. https://doi.org/10.1139/g97-809

Georgi, L., Johnson-Cicalese, J., Honig, J., Das, S.P., Rajah, V.D., Bhattacharya, D., Bassil, N., Rowland, L.J., Polashock, J., Vorsa, N., 2013. The first genetic map of the American cranberry: exploration of synteny conservation and quantitative trait loci. Theor Appl Genet 126, 673–692. https://doi.org/10.1007/s00122-012-2010-8

Johnson-Cicalese, J., Polashock, J.J., Honig, J.A., Vaiciunas, J., Ward, D.L., Vorsa, N., 2015. Heritability of Fruit Rot Resistance in American Cranberry. Journal of the American Society for Horticultural Science 140, 233–242. https://doi.org/10.21273/JASHS.140.3.233

Li, H., 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27, 2987–2993. https://doi.org/10.1093/bioinformatics/btr509

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Michelmore, R.W., Paran, I., Kesseli, R.V., 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proceedings of the National Academy of Sciences 88, 9828–9832. https://doi.org/10.1073/pnas.88.21.9828

Schlautman, B., Covarrubias-Pazaran, G., Diaz-Garcia, L.A., Johnson-Cicalese, J., Iorrizo, M., Rodriguez-Bonilla, L., Bougie, Tierney, Bougie, Tiffany, Wiesman, E., Steffan, S., Polashock, J., Vorsa, N., Zalapa, J., 2015. Development of a high-density cranberry SSR linkage map for comparative genetic analysis and trait detection. Mol Breeding 35, 177. https://doi.org/10.1007/s11032-015-0367-5

United States Department of Agriculture- National Agriculture Statistics Service. 2017. *Wisconsin Ag News- Cranberries* [Data file] Retrieved from https://www.nass.usda.gov/Statistics_by_State/Wisconsin/Publications/Crops/2018/WI-Cranberries-Annual-06-18.pdf

Vorsa, N., 2008. Cranberry variety named 'NJS98-35.' USPP18911P2.

Vorsa, N., Johnson-Cicalese, J., 2012. American Cranberry, in: Badenes, M.L., Byrne, D.H. (Eds.), Fruit Breeding. Springer US, Boston, MA, pp. 191–223. https://doi.org/10.1007/978-1-4419-0763-9_6

Wang, Y., Johnson-Cicalese, J., Singh, A.P., Vorsa, N., 2017. Characterization and quantification of flavonoids and organic acids over fruit development in American cranberry (Vaccinium macrocarpon) cultivars using HPLC and APCI-MS/MS. Plant Science 262, 91–102. https://doi.org/10.1016/j.plantsci.2017.06.004

Wang, Y., Fong, S.K., Singh, A.P., Vorsa, N., Johnson-Cicalese, J., 2019. Variation of Anthocyanins, Proanthocyanidins, Flavonols, and Organic Acids in Cultivated and Wild Diploid Blueberry Species. HortScience 54, 576–585. https://doi.org/10.21273/HORTSCI13491-18

Wu, B.H., Zhao, J.B., Chen, J., Xi, H.F., Jiang, Q., Li, S.H., 2012. Maternal inheritance of sugars and acids in peach (P. persica (L.) Batsch) fruit. Euphytica 188, 333–345. https://doi.org/10.1007/s10681-012-0668-2

Yu, Y., Chen, C., Gmitter, F.G., 2016. QTL mapping of mandarin (Citrus reticulata) fruit characters using high-throughput SNP markers. Tree Genetics & Genomes 12, 77. https://doi.org/10.1007/s11295-016-1034-7

Table 3.1: Segregation test for the CA137, CA80, and CA64 populations. The mean and range are for citric acid (CA). % TA is measured in citric acid equivalents. Low citric acid is defined as < 2.5 mg/g citric acid concentration.

Cross	No. Low CA	No. Total	Test Ratio	X^2	p-value	CA Mean (mg/g)	CA Range (mg/g)	TA Mean	TA Range
CNJ08-137	20	120	1:3	4.444	0.035	6.448	1.11-12.18	2.27	1.80-2.98
CNJ08-80	12	24	1:1	0	1.000	3.649	0.807-8.56	1.75	1.46-2.08
CNJ08-64	23	74	1:3	0.086	0.769	7.028	.796-14.49	N/A	N/A

Table 3.2: Segregation of F2 and F1 self CNJ10 populations for low citric acid phenotype, with most populations having the expected ratios of 3:1 (high:low). Low citric acid is defined as < 2 mg/g fresh weight.

Cross	#Low CA	#Normal	Total	Test Ratio	X^2	DF	p-value
CNJ10-75	8	23	31	3:1	0.011	1	0.917
CNJ10-76	4	4	8	3:1	2.667	1	0.103
CNJ10-60	9	23	32	3:1	0.167	1	0.683
CNJ10-61	6	9	15	3:1	1.8	1	0.180
CNJ10-62	6	18	24	3:1	0	1	1.000
CNJ10-63	7	38	45	3:1	2.141	1	0.143
CNJ10-64	6	12	18	3:1	0.667	1	0.414
CNJ10-65	3	22	25	3:1	2.253	1	0.133
CNJ10-66	5	13	18	3:1	0.074	1	0.786
CNJ10-67	10	12	22	3:1	4.909	1	0.027
Total of X^2 s					14.69	10	>0.05
Pooled X^2	64	174	238	3:1	0.453	1	0.500
Heterogeneity X^2					13.58	9	0.138

Table 3.3: Correlation matrices showing relationships between quinic, malic, and citric acids and titratable acidity for CA137, CA80, and CA64. Only showing significant correlations with a $p < 0.05$.

	Population	Quinic	Malic	Citric	TA
Quinic	CA64				
	CA80				***<.0001
	CA137				**<.005
Malic	CA64				*<.05
	CA80				
	CA137				
Citric	CA64		-0.486***		
	CA80		-0.583**		
	CA137		-0.472***		
TA	CA64				
	CA80			0.465*	
	CA137	0.230*		0.541***	

Table 3.4: SSR marker primer sequences identified from bulk segregant analysis and a SNP marker that was identified by Kawash et al., unpublished.

Marker	Forward Primer Sequence	Reverse Primer Sequence	Physical Location
scf258d	CACGACGTTGTAAAACGACGTAACGCATTGGTCGGCTAT	GTTTCTTTAAGCCAAACCCAATCCAAC	Contig1:47998371
scf157322	TGTAAAACGACGGCCAGTAGTTATGAGGCTTACGAGGAG	GTTTCTTGATGGAACGATGAAACTGAT	Contig1:48039043
	SNP Location		
SNP CA_609	ATGGCTTAAGCCATTGTTCTCATTTCCGTTGTCTCAACGACCCATCATATCCATTTTTTGTAAAAGAACC GGTAAGAGAAAACATGCAGAGAAATGCACA**[A/G]**AAGCTAAACCTTAATCATTGGAAAACTGGACAA CTATAGAAGTTGGACCCAAATGAAATTTAGTAGAAACGAAGATATTATCAAATTACAAAGGCCCTGT		Contig1:47698903

Table 3.5: Segregation of genotypes for markers SNP CA_609, scf258d, and scf157322. The genotypes linked with low citric acid are SNP CA_609: G/G, scf258d: 225/225, and scf157322: 276/276. Sensitivity and specificity were calculated for detection of the *cita* individuals.

	Populations											
Markers	CA64				CA80		CA137				Sensitivity	Specificity
scf258d	217/217	217/223	217/225	223/225	217/217	217/231	217/217	217/225	217/229	225/229		
Low CA	23	0	1	0	11	0	18	0	2	0	0.945	0.975
Normal CA	1	16	19	14	1	10	2	31	37	26		
scf157322	276/276	260/276	260/260		276/276	270/276	276/276	260/270	260/276	270/276		
Low CA	24	0	0		11	0	16	0	0	3	0.944	0.975
Normal CA	1	30	19		1	10	2	17	42	35		
SNP CA_609	G/G	A/G			G/G	A/G	G/G	A/G	A/A			
Low CA	23	1			11	1	16	3	0		0.909	0.907
Normal CA	12	38			1	11	2	66	31			

Figure 3.1: Citric acid concentrations in mg/g fresh weight for parents (F1 hybrids CNJ04-59-47 and CNJ04-30-34) and grandparents (#35, Stevens, and NJ91-7-12) of CA137 (see Figure 2) with Crimson Queen (CQ) and Duke (*V. corymbosum*) as standards.

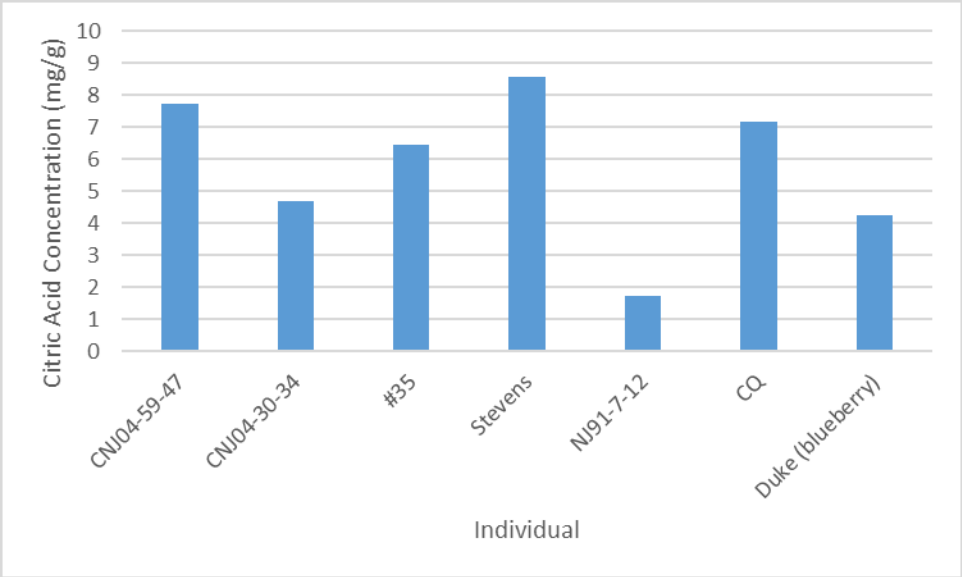


Figure 3.2: Pedigrees of CA137, CA80, and CA64. CA137 and CA64 are biparental crosses, while CA80 is a backcross to the low citric acid parent, NJ91-7-12.

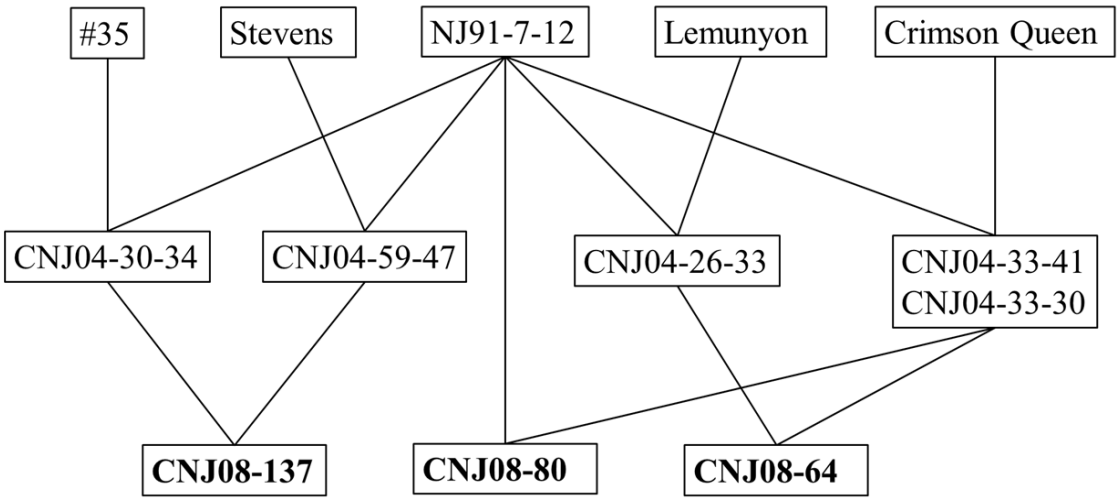


Figure 3.3: Pedigrees for populations evaluated in Table 2. Single lines indicate selfs or crosses between siblings.

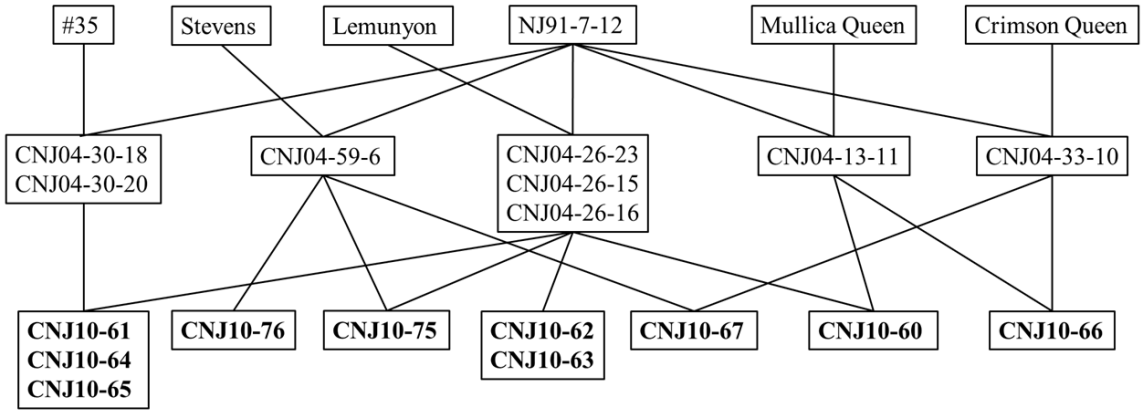
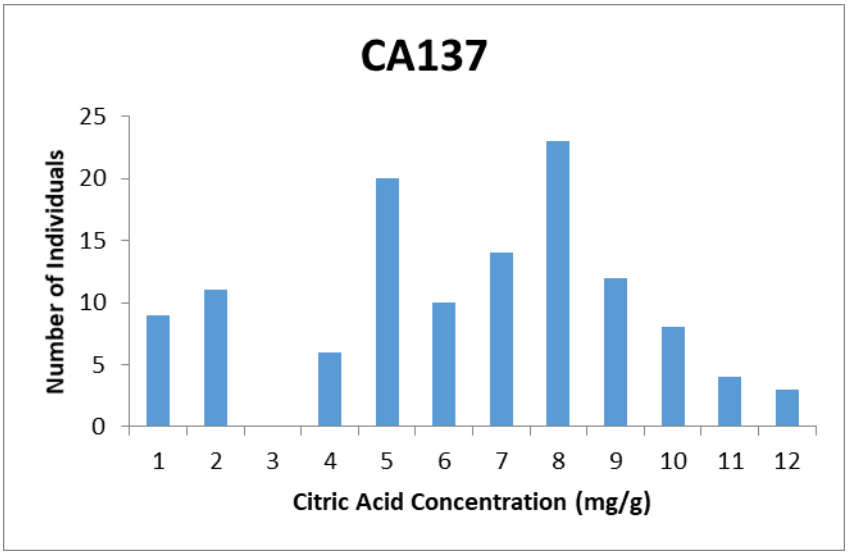
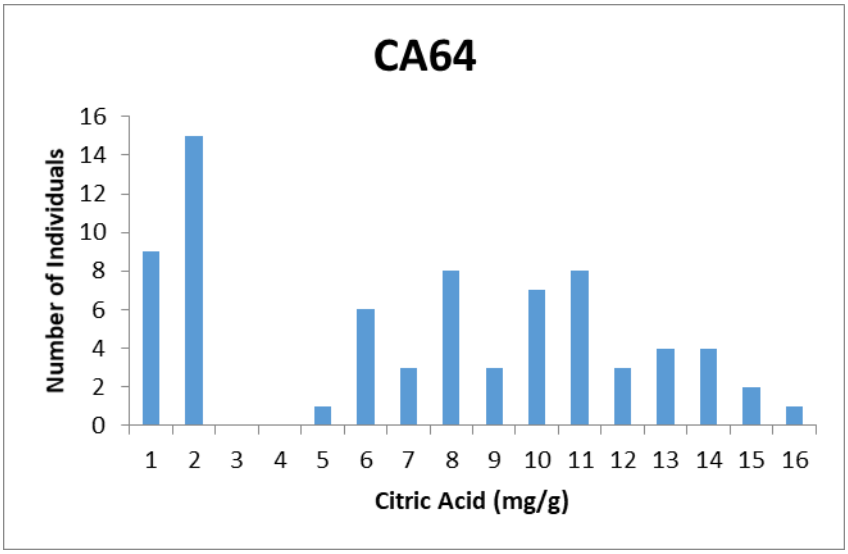


Figure 3.4: Frequency distribution of citric acid concentration in mg/g fresh weight for A) CA137, B) CA64, and C) CA80. Low citric acid was classified as less than 2 mg/g.

A)



B)



C)

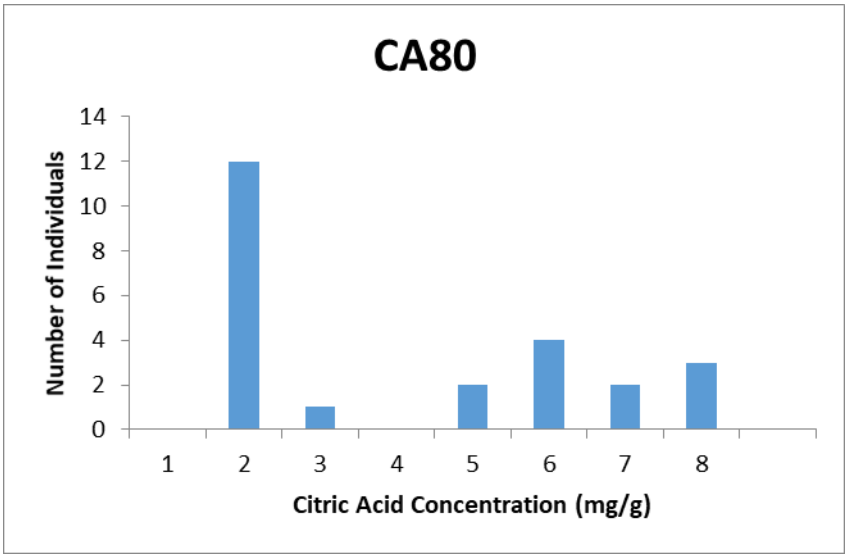
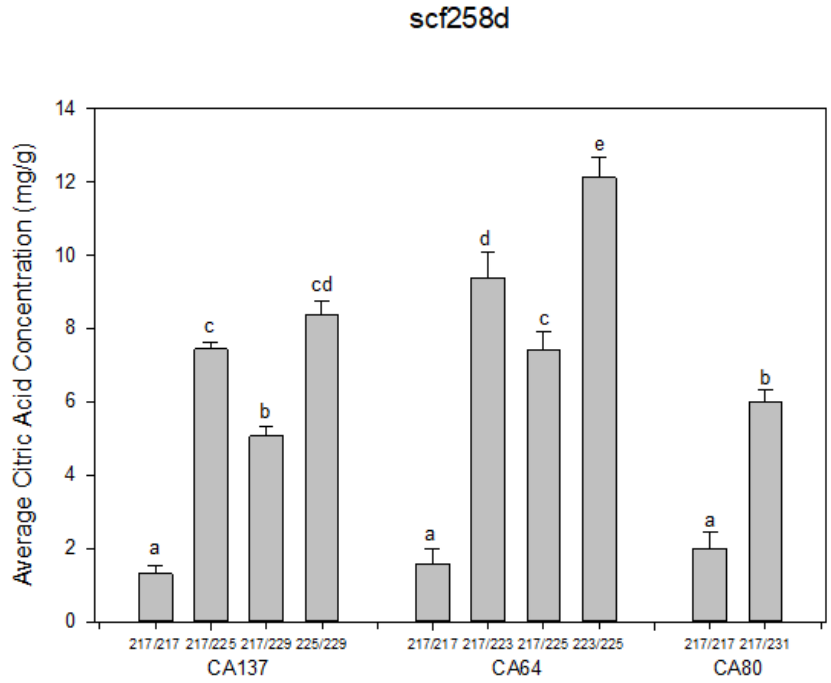

CA80
Number of Individuals
14
12
10
8
6
4
2
0
1
2
3
4
5
6
7
8
Citric Acid Concentration (mg/g)

Figure 3.5: Average citric acid concentration for each genotype for the SSR markers scf258d and scf157322.

A)



B)

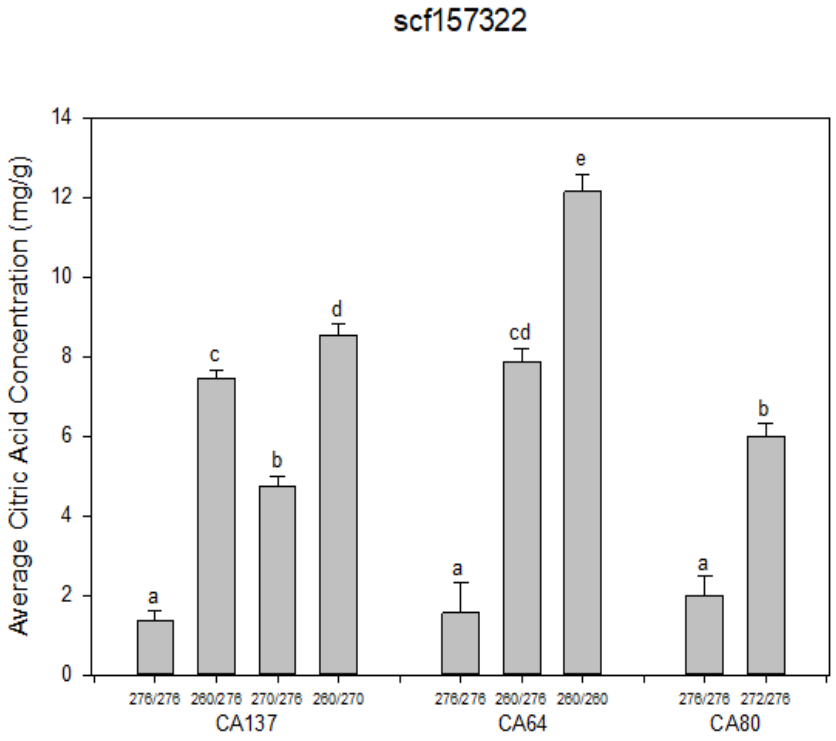
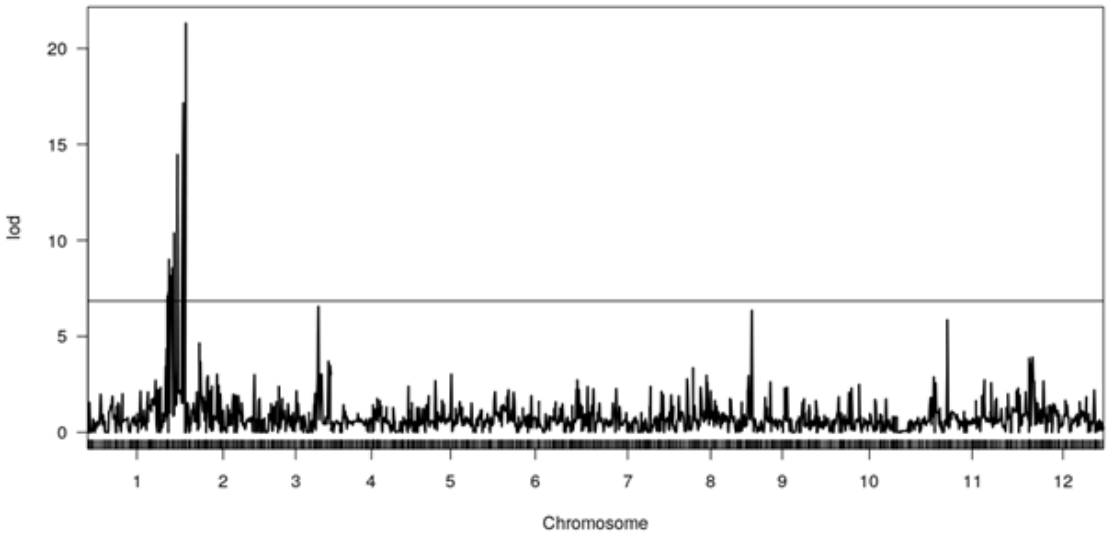


Figure 3.6: LOD score chart of the citric acid QTL from CA137. The threshold LOD score is 6.84 and there is a single large effect QTL at Contig 1: 48141640, which is located on chromosome 1.



Chapter 4: Genetics of a Low Malic Acid Trait and the Development of Molecular Markers in Cranberry Fruit

Introduction

While cranberry is well known for its health benefits, significant amounts of added-sugars are necessary for palatablity of cranberry products. 95% of the US cranberry crop is processed into fruit juice blends, sweetened dried cranberries, and other consumer products. However, cranberry products typically contain up to 40% of added-sugars (Ocean Spray, 2019) The palatability of fruit juices is related to the sugar:acid ratio (Bates et al., 2001). The optimal sugar-acid ratio used to formulate these products uses brix (an industry a measure of sugar content) and titratable acidity (TA), generally measured as citric acid equivalents in cranberry. Some minimum level of acid or acids, typically between 0.5-1.0 percent is desired for expression of fruit flavor.

Cranberry fruit has three organic acids that contribute to TA: malic, citric, and quinic acids (Cunningham et al., 2003). Malic acid is a significant contributor to cranberry's TA. Thus, reducing the malic acid content in cranberry should decrease TA. By lowering the malic acid in cranberry, the amount of added sugar can be decreased and the percentage of cranberry juice can be increased in cranberry juice products. This will enhance the flavonoid constituents of cranberry products which are considered to provide health benefits of cranberry consumption. It should be noted however, that the perception of 'sourness' and astringency depends on specific acids, various acid compositions, and varies among different people (Rubico and McDaniel, 1992).

Malic acid, a dicarboxylic acid, has a pKa of 3.4 and 5.2 and is approximately 90% as sour as citric acid (Kader, 2008). A common and naturally occurring compound, malic

acid contributes to the refreshing tart taste of various fruit species, particularly associated with apple, in which malic is the primary fruit acid. In apples, the low acid trait, as measured by pH and TA, is recessively inherited (Xu et al., 2012). Fine mapping of the low malic acid trait revealed aluminum-activated malate transporter-like genes, which maintain malate homeostasis (Bai et al., 2012). In tomato, there is also evidence that acidity is also controlled by aluminum-activated malate transporter-like genes (Ye et al., 2017). Previously, a low citric acid trait in cranberry was characterized (Chapter 2) and thus, this study focuses on characterizing the phenotypic variation and genetics of the low malic acid trait in cranberry. Evaluation of malic acid throughout the cranberry fruit development period revealed that malic acid increases early after fruit set and levels off during fruit ripening (Wang et al., 2017). In addition, QTLs were identified for malic acid on chromosomes 1, 6, and 12 in various breeding populations segregating for fruit rot resistance (Chapter 2). In populations with the low citric acid trait (*cita)*, there was an inverse relationship between citric and malic acids; as citric acid concentrations decreased, malic acid concentrations increased (Chapter 2).

A wild accession with reduced malic acid concentration was identified in the cranberry germplasm collection at the Marucci Blueberry and Cranberry Research and Extension Center, Chatsworth, NJ (Cunningham and Vorsa, unpublished data). As the first characterization of the genetics of a qualitative low malic acid fruit trait (*mala*) in cranberry fruit, the objectives of this study are threefold: 1) describe the inheritance of the *mala* allele, its effect on TA, and its relationship to citric and quinic acids, 2) identify and develop molecular markers closely linked to *mala* for use in marker-assisted selection, and 3) determine the effect of the *mala* allele and genotypes in different genetic backgrounds.

Materials and Methods

Plant Material

The wild, low malic acid germplasm accession, NJ93-57, was collected in Suffolk County, New York in 1993. An initial cross, NJ93-57 x cv. Mullica Queen (MQ) was made in 2004 to generate the CNJ04-52 population. From this population, one low TA individual (CNJ04-52-46) was selfed to generate the CNJ08-100 (hereafter MA100) population, and another (CNJ04-52-54) was crossed with the *cita* germplasm accession (NJ91-7-12) (Chapter 3) to generate CNJ08-103 and CNJ08-98 populations (Figure 1). The progeny from these populations were evaluated for organic acid concentrations and progeny with the lowest malic acid content, CNJ08-103-20 and CNJ08-98-80, were selected for subsequent crosses (Figure 2). In 2012, CNJ08-103-20 was self-pollinated to give rise to population CNJ12-155 (CM155, named CM for containing both *cita* and *mala* alleles). The cross CNJ08-98-80 x CNJ08-103-20 gave rise to population CNJ12-151 (CM151). CM155 is a population derived from self-pollinated CNJ08-103-20, and CM151 is CNJ08-98-80 x CNJ08-103-20, half-sibs of CM155 (Figure 1).

Seeds were germinated in the greenhouse after stratification in a refrigerator for approximately 3 months, seedlings transplanted into 4 inch pots, and maintained in the greenhouse and grown for at least three years. Populations began to flower and fruit during years 2014-2017. Hand self-pollinations were carried out during 2014-2016. During the spring/summer 2017flowering plants were taken outside for open-pollination by bees. After fruit set was completed, plants were brought into the greenhouse. Fruit was collected

during late August and September in 2014-2017 for analysis of organic acids and TA. Leaf tissue was collected in 2017 for DNA extractions.

Organic Acid and Titratable Acidity Analysis

Organic acids were extracted and analyzed with high performance liquid chromatography (HPLC) as in Wang et al. (2017), with modifications as follows: About 5 grams of fruit were used per sample, depending on availability. The fruit and water (1g fruit:10ml water) were homogenized using a laboratory blender. A portion of the supernatant was used for HPLC analyses and the rest of the slurry was used for TA and brix (soluble solids). The supernatant was heated to 90 degrees C for 10 minutes, and frozen until analyzed on HPLC. TA and brix was performed similar to Vorsa and Johnson-Cicalese (2012). TA was quantified by titrating to an endpoint of pH 8.2 with .05 N NaOH using a Metrohm Ti-Touch 916. The final percent TA was calculated using citric acid equivalents. Brix was measured with a Atago PR-32 digital refractometer.

Genotyping

DNA was extracted from leaf tissue with a modified CTAB protocol and GBS libraries were generated as in Daverdin et al. (2017). The GBS libraries for all the populations used Msp1 and Pst1 as the restriction enzymes. After the initial analysis, an additional GBS library was generated for CM155 using Nde1 and Pst1 to yield more reads in the region of interest. Prepared DNA libraries were sent to Genewiz LLC (Plainfield, New Jersey, USA) for sequencing on Illumina Hi-seq. The first sequencing run was on a 2x100 bp configuration, with later sequencing runs on a 2x150 bp configuration.

Bulked Segregant Analysis with SSRs

Bulked segregant analysis with the same SSRs as in Chapter 3 was conducted.

QTL Identification

Barcoded samples were de-multiplexed using STACKS and aligned to the cranberry reference genome (Kawash et al., unpublished) with bwa-mem (Catchen et al., 2011; Li and Durbin, 2009; Kawash et al., unpublished). Samtools was used to call SNPs (Li, 2011). Qualifying SNPs required a read support of 4 reads, and heterogeneity between 25% and 75%. Missing data was limited to only 10% of the population in a given marker, and markers that were homogenous through the population were also removed. R/qtl was used to calculate genetic distance between markers and identify QTL (Broman et al., 2003). Genome wide significance of LOD scores was calculated at $p < 0.05$ through 1000 permutations.

Statistical Analyses

Correlations were analyzed using PROC CORR in SAS 9.4. Regression analysis was completed using PROC GLM in SAS 9.4. Means separation using PROC GLM with the lsd means function. Chi-square values were calculated using chisq.test in R. R/sm was used to generate density distributions.

Results

Inheritance of Low Malic Acid Trait

The frequency distribution for malic acid for all three crosses combined suggest that there is a threshold at < 2.5 mg/g malic acid to be considered low malic acid (Figure 3A). The density distribution for malic acid of the three crosses suggest a bimodal distribution, which indicate the presence of a single locus controlling for low malic acid (Figure 3B). In the populations, there is segregation for low malic acid at < 2.5 mg/g and normal malic acid > 2.5 mg/g. Further analysis of the segregation (Table 1), MA100

segregated 3:1 ($X^2 = 0.01$, $p = 0.919$) for normal malic acid to low malic acid (< 2 mg/g). CM155 also segregated 3:1 for malic acid ($X^2 = 0.877$, $p = 0.349$). However, the observed segregation for malic acid in CM151 was significantly different from the expected 3:1 ratio ($X^2 = 4.25$, $p = 0.039$.

Relationship of Malic Acid, Citric Acid, and TA

As shown in Table 1, the population means for malic acid concentration of all three populations was similar, 4.2-5.5 mg/g. However, the range was lower in the CM151 and CM155 populations. For citric acid, the average concentration and range was lower in the CM151 and CM155 populations than in the MA100 population. This is likely due to the introgression of the low citric acid allele into the CM151 and CM155 populations. Overall, the range for TA was similar for all three populations, with the exception of MA100, which had a slightly higher maximum TA. In all three populations, there was no significant variation between years of harvest for citric acid, malic acid, and TA (unpublished data). However, there was significant variation between years for quinic acid.

For MA100 and CM151, there was a positive correlation between citric and malic acids, with MA100 having a correlation of 0.76 ($p < 0.0001$) and CM151 with a correlation of 0.24 ($p<0.05$) (Table 2). In MA100, there was a significant positive correlation of 0.593 ($p<0.005$) between citric and quinic acid. For all three populations, there were significant positive correlations for quinic, malic, and citric acids versus TA, for malic versus quinic acid (Table 2).

Multiple regression analysis for MA100, CM151, and CM155 showed that the variation in malic and citric acids contributed the most to the variation in TA. For MA100, citric and malic acids accounted for 81% of the variation in TA. For CM151, citric and

malic acids accounted for 92% of the variation in TA. For CM155, citric and malic acids accounted for 96% of the variation in TA.

Bulk Segregant Analysis with SSRs

Initial screening of the parents with SSRs of MA100, NJ93-57 and MQ, resulted in 75 markers with polymorphisms. When bulks of high vs low malic acid were screened, 5 markers had polymorphisms between the bulks. However, when the full population was genotyped, there was no apparent co-segregation with low malic acid. None of the markers with polymorphisms were located on Chromosome 4.

Genetic Mapping

The first sequencing of the GBS libraries (MA100 and the first half of CM155, n=56) with a 2x100bp format resulted in a total of 43,517 Mbases with a mean quality score of 34.87. The second sequencing run (CM151 and the second half of CM155, n=75) with a 2x150bp format resulted in a total of 108,584 Mbases with a mean quality score of 37.21. De-multiplexing with STACKS, the samples yielded an average of 7 million reads per individual. In the first sequencing run, there were 431 million reads and 6.4 million reads per sample. The second sequencing run had 719 million reads and 7.6 million reads per sample. After aligning the reads to the current cranberry reference genome (Kawash et al., unpublished), calling SNPs using samtools, and filtering, a total of 13698 SNP markers were identified.

For MA100, QTL were identified using scanone, and peaks that surpassed the LOD threshold of 9.4 were deemed significant. The QTL was located at marker 4_36181956 on chromosome 4. Further analysis with R/qtl determined the QTL interval to be 267 kb in size from 36106859 to 36371940 (Figure 4A). This QTL on chromosome

4 had a LOD score of 11.6 and accounted for 81% of the phenotypic variance in the population. For CM151 and CM155, the data were combined for QTL analysis. Peaks that surpassed the LOD threshold of 10.21 were deemed significant. The QTL was located at marker 4_36106859. Further analysis determined the QTL interval to be 437 kb in size from 36063067 to 36499933. This QTL was also on chromosome 4 and had a LOD score of 24.0 and accounted for 72% of the phenotypic variance in the populations. All significant peaks occurred on chromosome 4.

These two QTLs were SNP markers that co-segregate with the low malic acid trait (Figure 5). For marker 4_36106859 (MA_859) and marker 4_36181956 (MA_956), the TT genotype co-segregated with low malic acid. These two markers are 11kb apart. Additionally, the malic acid level phenotypes grouped by the SNP genotypes are significantly different, with homozygous, *mala/mala,* having the lowest, heterozygotes, *Mala/mala,* as an intermediate, and homozygous *Mala/Mala* having the highest malic acid, indicating that the low malic acid trait displays co-dominance.

Discussion

Cranberry fruit has been well studied for its various health benefits due to the high level of human health promoting compounds such as flavonols (Wang et al., 2017). However, their 'superfruit' status is damaged due to the added-sugars necessary to balance their high acid content (Bates et al., 2001). Here, we identify a low malic acid trait, locus and allele that confers a decreased TA to below 1%, on par with more acidic fruit such as strawberries (Kallio et al., 2000). Cranberry is a woody perennial, producing fruit 2-3 years after germination of seed (Vorsa and Johnson-Cicalese, 2012). Due to this

long generation time, marker assisted selection (MAS) would be useful to decrease selection time and allow seedlings of no value to be culled, saving space and money.

In this study, we identified two SNP markers that co-segregate with the low malic acid trait that are 11kb apart. Through progeny testing, we determined that the low malic acid trait originally found in the germplasm accession NJ93-57 was heterozygous. MA100 segregates with a single, co-dominant gene in a Mendelian pattern. In two advanced populations, CM155 and CM151, for malic acid, both of these populations were consistent with a 3:1 segregation pattern of normal malic acid level to low malic acid level.

The *mala* allele also has a significant effect on TA. In comparison to *cita*, in its homozygous form, it can decrease the TA more significantly than cita can (Chapter 3). The *mala* allele also does not just affect malic acid concentrations, it also significantly decreases citric acid concentration as well. In the populations studied in this chapter, there is a positive correlation between citric and malic acids, as malic acid concentrations decrease, so do citric acid concentrations. However, there is a caveat, the mala plants have a dwarfed growth habit (Figure 6), which make them not commercially viable.

In this study, the locus for low malic acid was located on chromosome 4. In Chapter 2, there were QTL also located on chromosome 4, however, they were not consistent year-to-year. The QTL that were consistent year-to-year were on chromosomes 9, 12, and 10. This indicates that there are multiple loci controlling malic acid concentrations, as well as the qualitative locus we identified in this study.

Within 1Mb of the malic acid QTLs, there are multiple BLAST hits for malate dehydrogenase. In tomato, reduction in the malate dehydrogenase enzyme through

siRNA resulted in increased malic acid concentrations in the fruit (Centeno et al., 2011). While this is a possible candidate gene for low malic acid, there needs to be further sequencing and experimentation to determine if malate dehydrogenase is the causative gene.

A total of three populations and 119 unique individuals were evaluated in this study. Even though the population sizes were not very large, combining the analysis with three populations with related parentage helped us identify linked markers for the *mala*. Also, because *mala* segregates in a Mendelian pattern, the low number of individuals were sufficient to identify co-segregation of the low malic acid trait with SNP markers. In future studies, the markers developed in this study will need to be verified in populations of different genetic backgrounds.

The identification and characterization of this low citric acid trait in cranberry allows breeders to decrease the acidity of cranberry to reduce added-sugars in cranberry products. The SNP markers identified have been developed into KASP assays to start screening different populations as well as the germplasm collection for low malic acid. The implementation of the SNP markers will allow for marker assisted selection to accelerate the breeding cycles in cranberry.

References

Bai, Y., Dougherty, L., Li, M., Fazio, G., Cheng, L., Xu, K., 2012. A natural mutation-led truncation in one of the two aluminum-activated malate transporter-like genes at the Ma locus is associated with low fruit acidity in apple. Mol Genet Genomics 287, 663–678. https://doi.org/10.1007/s00438-012-0707-7

Bates, R.P., Morris, J.R. and Crandall, P.G., 2001. *Principles and practices of small-and medium-scale fruit juice processing* (No. 146). Food & Agriculture Org

Broman, K.W., Wu, H., Sen, S., Churchill, G.A., 2003. R/qtl: QTL mapping in experimental crosses. Bioinformatics 19, 889–890. https://doi.org/10.1093/bioinformatics/btg112

Catchen, J.M., Amores, A., Hohenlohe, P., Cresko, W., Postlethwait, J.H., 2011. Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences. G3: Genes, Genomes, Genetics 1, 171–182. https://doi.org/10.1534/g3.111.000240

Centeno, D.C., Osorio, S., Nunes-Nesi, A., Bertolo, A.L.F., Carneiro, R.T., Araújo, W.L., Steinhauser, M.-C., Michalska, J., Rohrmann, J., Geigenberger, P., Oliver, S.N., Stitt, M., Carrari, F., Rose, J.K.C., Fernie, A.R., 2011. Malate Plays a Crucial Role in Starch Metabolism, Ripening, and Soluble Solid Content of Tomato Fruit and Affects Postharvest Softening. The Plant Cell 23, 162–184. https://doi.org/10.1105/tpc.109.072231

Cunningham, D.G., Vannozzi, S.A., Turk, R., Roderick, R., O'Shea, E., Brilliant, K., 2003. Cranberry Phytochemicals and Their Health Benefits, in: Nutraceutical Beverages, ACS Symposium Series. American Chemical Society, pp. 35–51. https://doi.org/10.1021/bk-2004-0871.ch004

Daverdin, G., Johnson-Cicalese, J., Zalapa, J., Vorsa, N., Polashock, J., 2017. Identification and mapping of fruit rot resistance QTL in American cranberry using GBS. Mol Breeding 37, 38. https://doi.org/10.1007/s11032-017-0639-3

Kader, A.A., 2008. Flavor quality of fruits and vegetables. Journal of the Science of Food and Agriculture 88, 1863–1868. https://doi.org/10.1002/jsfa.3293

Kallio, H., Hakala, M., Pelkkikangas, A.-M., Lapveteläinen, A., 2000. Sugars and acids of strawberry varieties. European Food Research and Technology 212, 81–85. https://doi.org/10.1007/s002170000244

Li, H., 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27, 2987–2993. https://doi.org/10.1093/bioinformatics/btr509

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Original Dried Cranberries | Ocean Spray® [WWW Document], n.d. URL https://www.oceanspray.com:443/sitecore/content/OS/Brand Sites/OceanSpray/Master/Intl Master/Data/Products/Craisins Dried Cranberries/Craisins Original Dried Cranberries (accessed 3.10.19).

Rubico, S.M., McDaniel, M.R., 1992. Sensory evaluation of acids by free-choice profiling. Chemical Senses 17, 273–289. https://doi.org/10.1093/chemse/17.3.273

Vorsa, N., Johnson-Cicalese, J., 2012. American Cranberry, in: Badenes, M.L., Byrne, D.H. (Eds.), Fruit Breeding. Springer US, Boston, MA, pp. 191–223. https://doi.org/10.1007/978-1-4419-0763-9_6

Wang, Y., Johnson-Cicalese, J., Singh, A.P., Vorsa, N., 2017. Characterization and quantification of flavonoids and organic acids over fruit development in American cranberry (Vaccinium macrocarpon) cultivars using HPLC and APCI-MS/MS. Plant Science 262, 91–102. https://doi.org/10.1016/j.plantsci.2017.06.004

Xu, K., Wang, A., Brown, S., 2012. Genetic characterization of the Ma locus with pH and titratable acidity in apple. Molecular Breeding 30, 899–912. https://doi.org/10.1007/s11032-011-9674-7

Ye, J., Wang, X., Hu, T., Zhang, F., Wang, B., Li, C., Yang, T., Li, H., Lu, Y., Giovannoni, J.J., Zhang, Y., Ye, Z., 2017. An InDel in the Promoter of *Al-ACTIVATED MALATE TRANSPORTER9* Selected during Tomato Domestication Determines Fruit Malate Contents and Aluminum Tolerance. The Plant Cell 29, 2249–2268. https://doi.org/10.1105/tpc.17.00211

Table 4.1: Segregation ratios and summary statistics for MA100, CM151, and CM155. Low malic acid (MA) is < 2mg/g fresh fruit. MA and CA are concentrations in mg/g fresh fruit. TA is % acid in citric acid equivalents.

Cross	CNJ08-100	CNJ12-151	CNJ12-155
# Low MA	8	6	12
# Normal	25	43	26
Total	33	49	38
Test Ratio	3:1	3:1	3:1
X^2	0.01	0.877	4.252
p-value	0.919	0.349	0.039

Cross	CNJ08-100	CNJ12-151	CNJ12-155
MA mean	5.06	5.54	4.22
MA range	1.27-8.98	0.38-8.9	.08-6.8
CA mean	10.54	6.31	4.70
CA range	4.82-16.64	1.3-11.0	1.3-10.3
TA mean	2.03	1.77	1.39
TA range	0.73-3.26	0.69-2.42	0.60-2.05

Table 4.2: Correlation matrix of quinic, malic, citric, and TA.

	Population	Quinic	Malic	Citric	TA
Quinic	MA100				
	MA151				***<.0001
	MA155				**<.005
Malic	MA100	0.523**			*<.05
	MA151	0.220*			
	MA155	0.429**			
Citric	MA100	0.593**	0.758***		
	MA151		0.241*		
	MA155				
TA	MA100	0.523**	0.859***	0.817***	
	MA151	0.312**	0.712***	0.738***	
	MA155	0.529***	0.828***	0.606***	

Figure 4.1: Pedigrees for CM100, CM151, and CM155, red indicates *cita* background, blue indicates *mala* background, and purple indicates both.

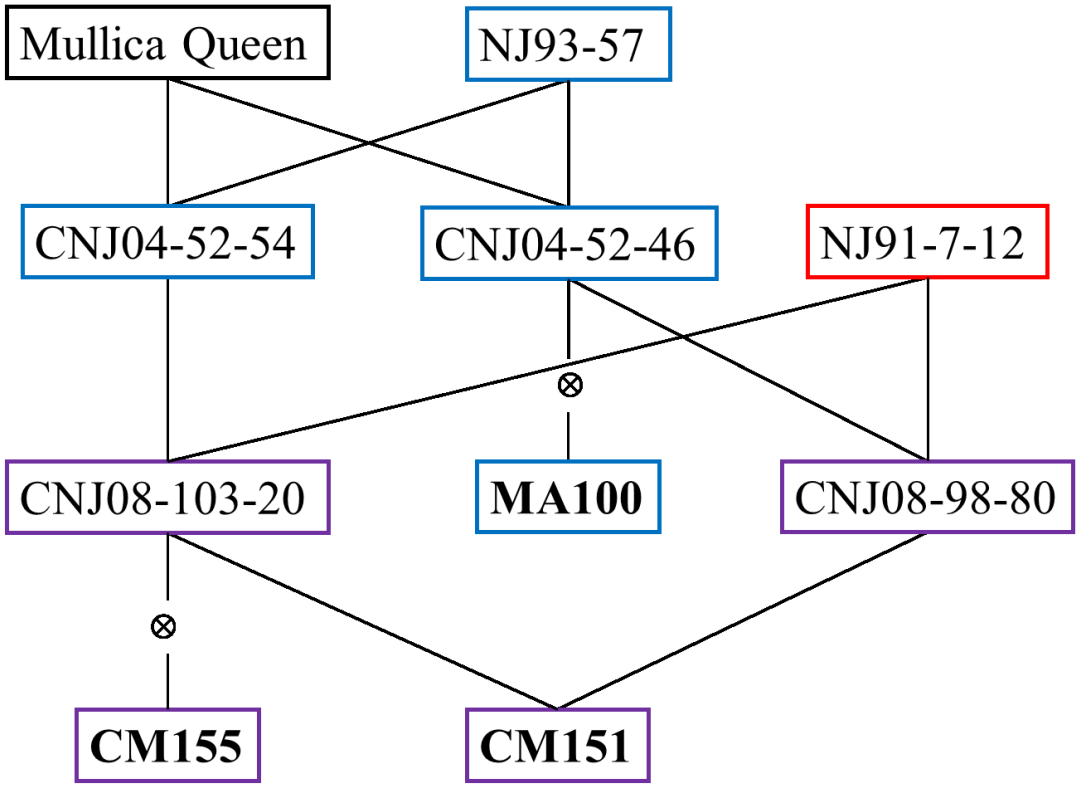


Figure 4.2: Malic acid concentrations of parents and standards, with blueberry as comparison. Malic acid was measured as mg/g fresh weight.

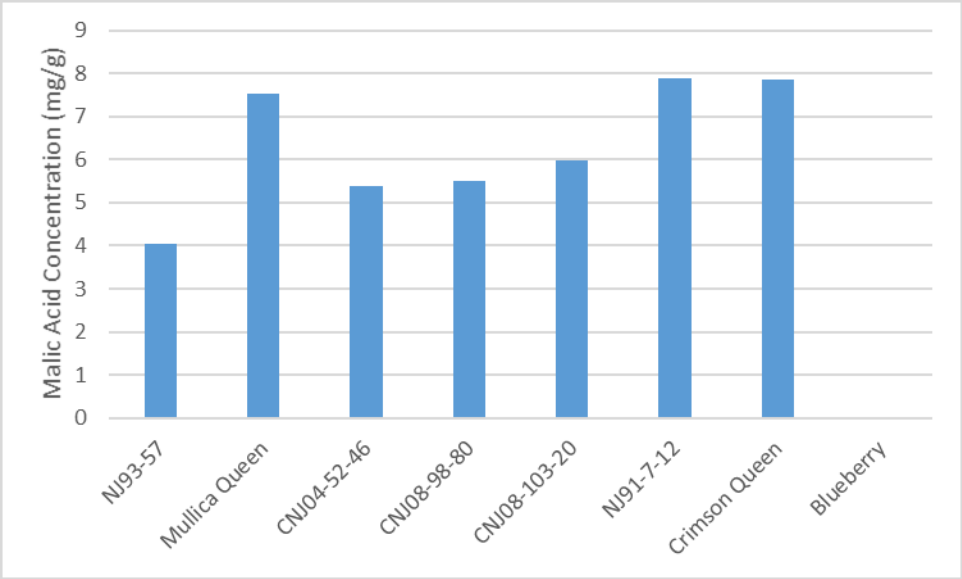
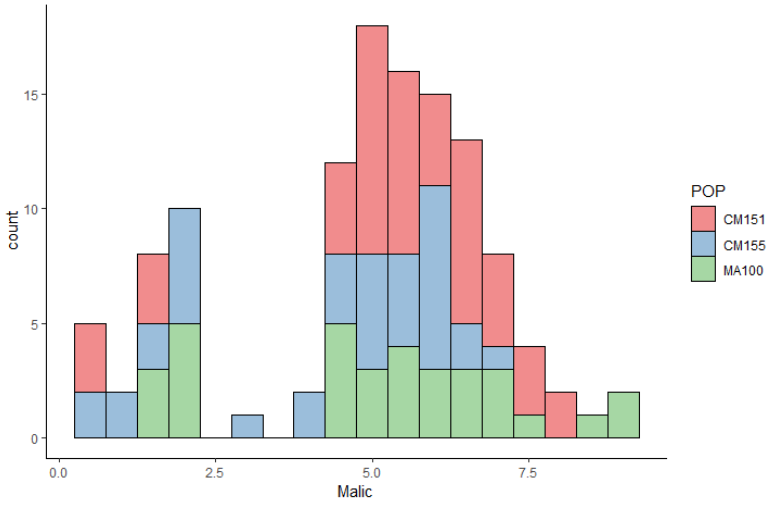


Figure 4.3: A) Combined histogram of CM151, CM155, and MA100. B) Density distribution of CM 151, CM155, and MA100. Malic acid was measured as mg/g fresh weight. Low malic acid was considered to be < 2.5 mg/g.

A)



B)

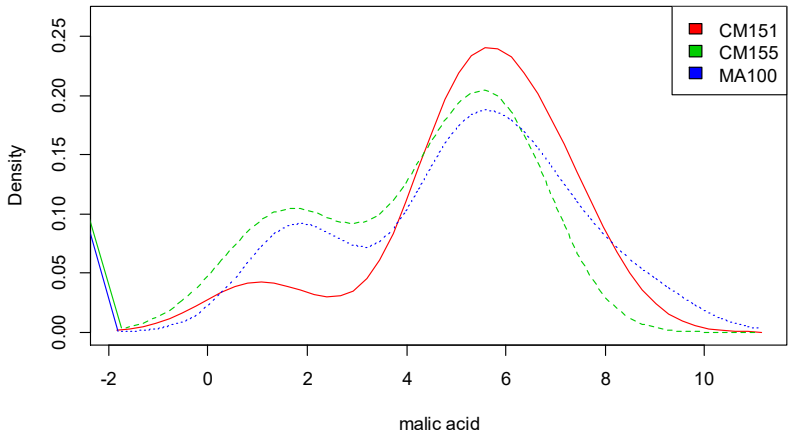
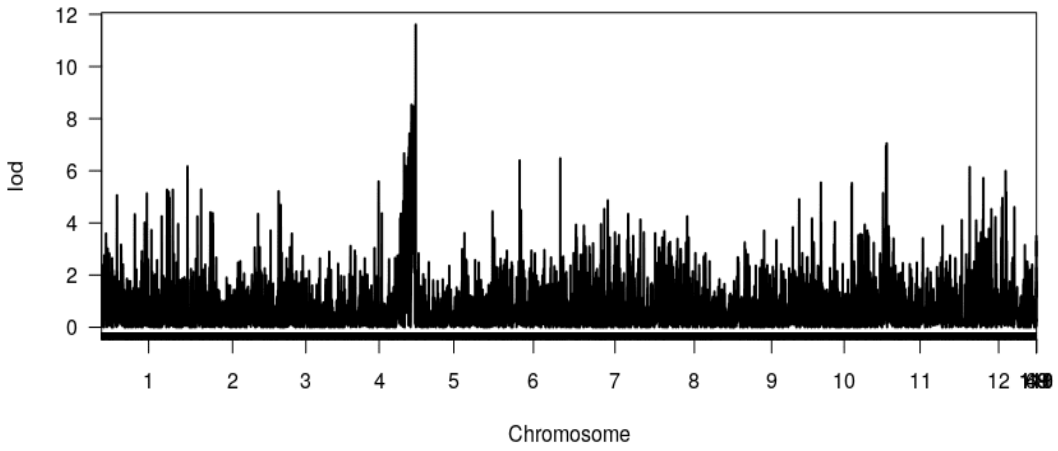


Figure 4.4: A) QTL peak for CM100 for low malic acid. B) QTL peak in CM151 and CM155 for low malic acid.

A)



B)

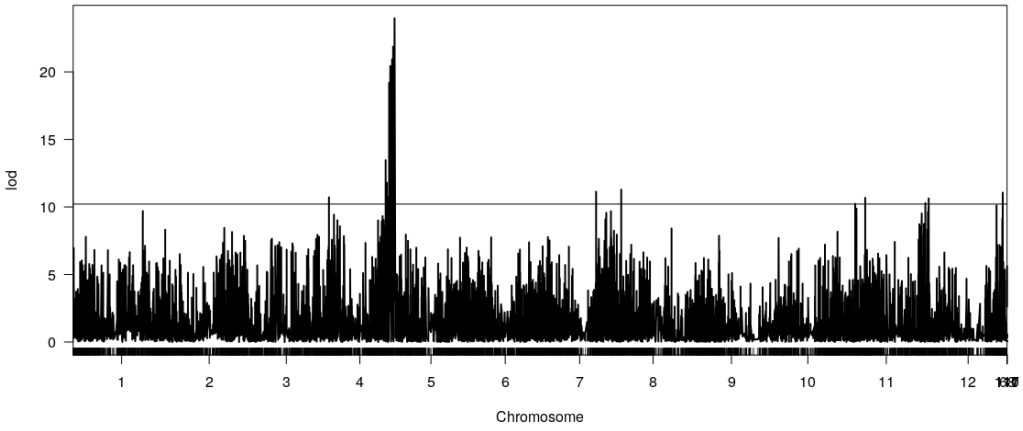
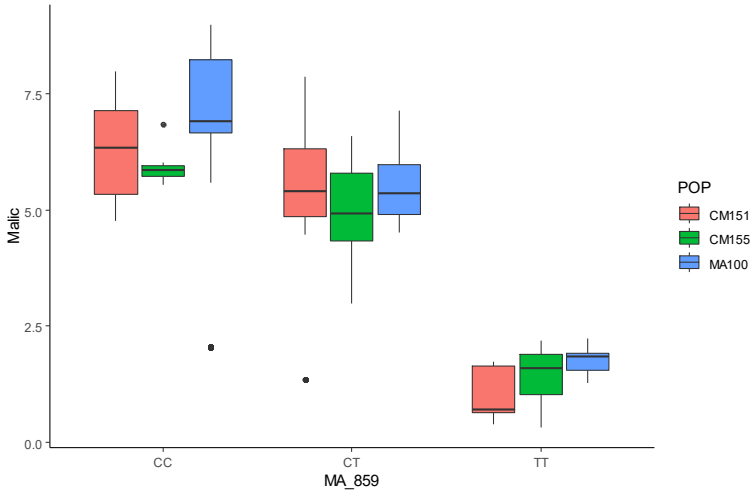


Figure 4.5: Distribution of the low malic acid trait in CM151, CM155, and MA100 along with the SNP genotypes for A) MA_859 and B) MA_956. Malic acid is quantified as mg/g fresh weight. The TT genotype in MA_859 and MA_956 are associated with low malic acid (< 2 mg/g).

A)



B)

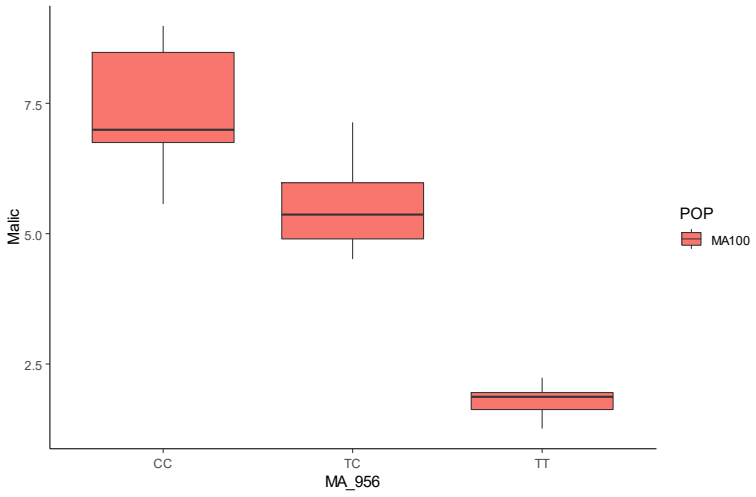


Figure 4.6 : A) Dwarfed growth habit phenotype of a *mala/mala* individual. B) Normal phenotype.

A)

B)

Chapter 5: Validation of the *cita* and *mala* Loci and Their Interactions and Effect on Titratable Acidity in Cranberry Fruit

Introduction

Citric acid, malic acid, and titratable acidity (TA) significantly affect fruit quality and taste, particularly the sourness, in most fruit species (Kader, 2008). In cranberry fruit, the acidity is of particular importance because of the extremely high TA found, ~2.5% and over 2x the concentration found in most edible fruits (Sadler and Murphy, 2010). Added-sugars are necessary to be added to cranberry products for palatability (Bates et al., 2001). However, there has been labeling regulations from the USDA to indicate added-sugars to products to promote healthier eating, as excess sugar consumption is thought to lead to increased obesity rates and metabolic diseases (Pomeranz, 2012). Thus, it is important to reduce added-sugars to cranberry products to increase their consumer acceptance and increase the health benefits.

Previously in chapter 2, the QTLs for citric and malic acids in cranberry have been found on different chromosomes, chromosomes 1, 9, 10, 11, and 12. Then, in chapters 3 and 4, a low citric acid trait (*cita*) and a low malic acid trait (*mala*) have been characterized, found to be inherited as a single locus Mendelian-inherited trait, and located to chromosomes 1 and 4, respectively. The markers identified for the *cita* trait are located on the same chromosome as the QTL for citric acid on chromosome 1 in Crimson Queen. While the remaining QTLs identified were on different chromosomes than *cita* and *mala.* Additionally, two SSRs and two SNPs associated with *cita* and *mala*, respectively, have been identified through bulked segregant analysis and QTL analysis.

In efforts to further reduce the TA in cranberry fruit, here we combine the *mala* and *cita* traits in segregating populations.

Here, we characterized four related populations containing alleles from *cita* and *mala* to determine the effect of both these alleles on TA as well as citric and malic acids. The objectives of this study are to determine: 1) if *cita* and *mala* segregate independently, 2) if there is interaction or additivity between *mala* and *cita* alleles as reflected in TA, citric acid, and malic acid, 3) the efficacy and validity of the molecular markers previously identified.

Materials and Methods

Plant Material

The wild accessions, NJ93-57 and NJ91-7-12 were collected in southern New Jersey and Long Island, NY, respectively. Crosses were made in 2004, 2008, and 2012 as shown in Figure 1. F_1 crosses were made as follows: NJ93-57 x Mullica Queen (MQ), NJ91-7-12 x MQ, and NJ93-57 x Crimson Queen to yield the populations CNJ04-52, CNJ04-13, and CNJ04-34, respectively. The second cycle crosses were either a self (CNJ08-30-20) or a cross to NJ91-7-12 (CNJ08-103, CNJ08-98, and CNJ08-90). Finally, CNJ12-92 (CM92), CNJ12-93 (CM93), and CNJ12-155 (CM155) were generated by intercrossing the second cycle progeny as in Figure 1A. CNJ12-151 (CM151) was a self-pollination of an F_1 (CNJ08-103-20). The genotypes of the crosses are shown in Figure 1B. The plants were germinated in the greenhouse, transplanted into 4 inch pots, and retained in the greenhouse. During the summer, flowering plants were taken outside for open-pollination by bees. Fruit was collected in 2014-2017 for acidity analyses and leaf tissue was collected in 2017 for DNA extractions.

Organic Acid and Titratable Acidity Analysis

Organic acid analysis and titratable acidity were performed as in Chapter 4.

Marker Screening

Individuals were genotyped using SSRs as in Georgi et al. (2014). SNP genotyping was conducted using custom KASP (kompetitive allele specific PCR) assays developed by LGC Biosearch Technologies. SSR primer sequences and SNP regions are shown in Table 1.

Statistical Analyses

PROC GLM in SAS was used for means separation, regression analysis, and to determine year-to-year variation between the genotypes. PROC CORR was used for correlations between traits. R/chisq.test was used for chi-square analysis of segregation ratios. R/sm was used to generate density distributions.

Results

Citric and Malic Acid Phenotypes and Relationship to Quinic Acid, Brix, and TA

For CM92 and 93, the density distributions for quinic acid, malic acid, TA, and brix all appear to follow a normal distribution (Figure 2A, B, D, E). The distributions for citric acid appeared to follow a bimodal distribution (Figure 2C). For CM151 and CM155, the distribution seemed to be bimodal for quinic acid, malic acid, and TA (Figure 2A, B, D). The distribution for brix seemed to follow a normal distribution while the distribution for citric acid appeared to follow a bimodal distribution for CM155 and almost a trimodal distribution for CM151 (Figure 2B, C).

Closer analysis of the distributions revealed that there was segregation for low citric acid (< 2 mg/g) in all four populations while there was segregation for low malic

acid (< 2 mg/g) in CM151 and CM155, shown in Table 2. Both CM151 and CM155 fit a 9:3:3:1 segregation ratio for a dihybrid cross with two independently segregating traits (CM151: chisq=6.33, p=0.097, CM155: chisq=4.85, p=0.18). Both CM92 and CM93 fit a 1:1 segregation ratio for low citric acid (CM92: chisq=0.13, p=0.73, CM93: chisq=1.6, p=0.21).

In the correlation analysis between citric acid, malic acid, quinic acid, TA, and brix (Table 3), there were significant positive correlations of both malic and citric acids and TA in all four populations. For CM92 and CM93, there is a small negative correlation between TA and quinic acid while there was a small positive correlation between quinic acid and brix. For CM151 and CM155, there was a significant positive correlation between quinic acid and malic acid, as well as quinic acid and TA. CM151 had significant positive correlations between brix and malic acid as well as brix and TA. CM155 had a significant positive correlation between citric and quinic acids.

KASP and SSR Marker Validation

The SNP markers for malic acid (MA_271 and MA_476) were first identified in the previous chapter (Chapter 4). These SNPs were converted into KASP assays for high-throughput analysis of progeny in different populations. Because CM151 and CM155 also segregate for citric acid, a SNP (CA_609) was associated with low citric acid (Kawash et al., unpublished). The two SSRs were identified in Chapter 3 through bulk segregant analysis.

The association of the low citric and low malic acid traits were evaluated in two new populations, CM92 and 93, to determine the sensitivity and specificity of these markers (Table 4). The heterozygous individuals were determined based on expected

phenotypic segregation. For citric acid, the sensitivity and specificity was over 90% for all 3 markers, with scf258d having the best sensitivity and specificity. For malic acid, while CM92 and CM93 did not segregate for low malic acid, the SNP markers were used to identify the efficacy of the markers in false positives. As Figure 4B shows, MA_271 had a lot of false positives, resulting in a specificity of 81%. However, MA_476 showed consistency against false positives and had a high sensitivity and specificity to detecting *mala* (Figure 4B).

Interaction of cita and mala alleles

Overall, individuals with homozygous *cita* have significantly lower citric acid concentrations (< 2 mg/g) (Figure 3B, E). The presence of one allele of *mala* has a significant effect on reducing overall TA and citric acid (Figure 3B, E, C, F). However, the presence of two *cita* alleles seems to have an additive effect on increasing malic acid concentrations (Figure 3A, D). However, even though the presence of *cita* alleles increases malic acid concentrations, cumulatively, the *cita* and *mala* alleles seem to have an additive effect on decreasing overall TA (Figure 3C, F).

Because CM151 and CM155 are dihybrid crosses, they are exceptional populations to observe the different combinations of *mala* and *cita* genotypes and their effects on citric acid, malic acid, and TA. There is no significant effect of the heterozygous *mala* genotype on malic acid concentrations (Figure 4A). However, overall, there is a significant effect of at least one copy of the *cita* allele on citric acid concentration, indicating that the allele is indeed co-dominant. Interestingly, the presence of homozygous *mala* alleles (*mala/mala*) significantly decreases the citric acid concentration. There is a significant effect of the heterozygous *mala* allele on decreasing

malic acid concentrations, except for when there is also the presence of homozygous *cita* alleles (Figure 4b). In fact, the presence of at least one *cita* allele actually increases the malic acid concentration when the *mala* locus is *Mala*/-.

Utilizing all the phenotype and genotype data, a multiple regression analysis was performed to determine the total proportion of observed TA variation could be accounted for by these three acids. For CM151 and CM155 combined, the variables most significant for TA were citric, malic, and quinic acids and genotype, with a R^2 of 0.92. For CM92, the variables most significant were malic acid, quinic acid, and genotype, with a R^2 of 0.65. For CM93, the variables most significant were citric acid, malic acid, and genotype, with a R^2 of 0.72. When all the data was pooled, citric acid, malic acid, and genotype were most significant and the R^2 was 0.86.

Discussion

In this study, we evaluated 300 individuals with 2 SSRs and 3 KASP markers to validate the markers and determine the effect of the *cita* and *mala* alleles in different combinations. Segregation analysis revealed that *cita* and *mala* segregated independently, which is expected since the loci are located on two different chromosomes (Chapters 3 and 4). The SSR marker scf258d and the KASP marker MA_476 showed the best sensitivity and specificity for determining low citric or low malic acid individuals.

CM151 and CM155 were dihybrid crosses, thus had segregation for all possible combinations of *cita* and *mala.* This allowed us to determine the interaction of all the combinations of these two traits. There was an epistatic effect of the presence of at least one copy of the *cita* allele in increasing malic acid concentration. However, overall, when the interactions for TA are analyzed, there is a significant effect of the *cita* allele in

decreasing overall TA. This makes sense because while *cita* may increase malic acid, it also decreases citric acid, and TA is affected by both of these acids. While homozygous *cita* and homozygous *mala* individuals yielded the lowest TA, they are not currently commercially viable due to the possible linkage of the homozygous *mala* genotype to a dwarfed phenotype (Chapter 4).

To generate commercially viable populations, we explored the CM92 and CM93 populations which are heterozygous *cita* heterozygous *mala* x homozygous *cita* homozygous *Mala,* and thus do not have a dwarfed phenotype. This resulted in segregation of 1:1 for low citric acid (*cita/cita)* to heterozygous citric acid (*Cita/cita)* and heterozygous malic acid *(Mala/mala)* and normal malic acid (*Mala/Mala)*. CM92 and CM93 displayed the same effect observed in CM151 and CM155 of the *cita* allele increasing malic acid concentration while overall decreasing TA and citric acid concentration. Overall, homozygous *cita* and heterozygous *mala* individuals have an average TA of 1.50. The presence of *mala* and *cita* alleles had an additive effect on TA.

Even though linked markers are not the best for marker assisted selection (MAS), in the short term, these markers can be used for screening our breeding populations derived from the two low acid progenitors. However, finding the genes controlling citric and malic acid accumulation would allow greater ability for MAS and gene editing applications in the future.

References

Bates, R.P., Morris, J.R. and Crandall, P.G., 2001. *Principles and practices of small-and medium-scale fruit juice processing* (No. 146). Food & Agriculture Org.

Georgi, L., Johnson-Cicalese, J., Honig, J., Das, S.P., Rajah, V.D., Bhattacharya, D., Bassil, N., Rowland, L.J., Polashock, J., Vorsa, N., 2013. The first genetic map of the American cranberry: exploration of synteny conservation and quantitative trait loci. Theor Appl Genet 126, 673–692. https://doi.org/10.1007/s00122-012-2010-8

Kader, A.A., 2008. Flavor quality of fruits and vegetables. Journal of the Science of Food and Agriculture 88, 1863–1868. https://doi.org/10.1002/jsfa.3293

Pomeranz, J.L., 2012. The Bittersweet Truth About Sugar Labeling Regulations: They Are Achievable and Overdue. American Journal of Public Health 102, e14–e20. https://doi.org/10.2105/AJPH.2012.300732

Sadler, G.D., Murphy, P.A., 2010. pH and Titratable Acidity, in: Nielsen, S.S. (Ed.), Food Analysis, Food Analysis. Springer US, Boston, MA, pp. 219–238. https://doi.org/10.1007/978-1-4419-1478-1_13

Table 5.1: SSR marker primer sequences and SNP location sequences with physical location on genome sequence of cranberry. Scf258d, scf157322, and SNP CA_609 were linked with the *cita* locus while SNP MA_271 and SNP MA_476 were linked with the *mala* locus.

SSRs	Forward Primer Sequence	Reverse Primer Sequence	Physical Location
scf258d	CACGACGTTGTAAAACGACGTAACGCATTGGTCGGCTAT	GTTTCTTTAAGCCAAACCCAATCCAAC	Contig1:47998371
scf157322	TGTAAAACGACGGCCAGTAGTTATGAGGCTTACGAGGAG	GTTTCTTGATGGAACGATGAAACTGAT	Contig1:48039043
KASPs	**SNP Location**		
SNP CA_609	ATGGCTTAAGCCATTGTTCTCATTTCCGTTGTCTCAACGACCCATCATATCCATTTTTTGTAAAAGAACC GGTAAGAGAAAACATGCAGAGAAATGCACA**[A/G]**AAGCTAAACCTTAATCATTGGAAAACTGGACAA CTATAGAAGTTGGACCCAAATGAAATTTAGTAGAAACGAAGATATTATCAAATTACAAAGGCCCTGT		Contig1: 47698903
SNP MA_271	GCCTTGGAATTTGAACTTGATACCTCCTTGGCAAGCCCGGTCAATCTTCGATCTCAACCGTTACATCCC GCAATCAACGGTTCAGATTCACAAATGAATT**[T/C]**TTACCGGTCAAAGTGCATTATTATTGGTGATAAT GCATCTGCGAATCGAAATTGTTGATTGCCAAGATGAAAGGTTAGGATTGGTTGGCAGGGCCTACCG		Contig 4: 35386437
SNP MA_476	CGCAGGGAACAAAAATCTCCATTGTAGGAGATAAAACTAGAACAATCAAAAATGAAAAACGGAAAAA CAGAAAACAGTAAACCTCCGGCAGTAGATGCAG**[A/C]**TTCACAGGAACCCGTTGCCACTGGATCTGAT TACCTATTCGCCAAAAAACAAGGCCAAGCCCATCTTGATCCAAATAATTGATTCCATTCAATTTGGCAG		Contig 4: 36038844

Table 5.2: Chi-square analysis of CM151, CM155, CM92, and CM93 for goodness of fit to expected segregation ratios.

Cross	CM151	CM155	CM92	CM93
# Low MA	3	11	NA	NA
# Low CA	8	7	60	39
# MA and CA	3	0	NA	NA
# Normal	35	19	64	51
Total	49	37	124	90
Test Ratio	3:3:1:9	3:3:1:9	1:1	1:1
X^2	6.329	4.850	0.129	1.6
p-value	0.097	0.183	0.719	0.206

Table 5.3: Pearson's correlation coefficients and significance between quinic acid, malic acid, citric acid, TA, and brix for CM92, CM93, CM151, and CM155. Quinic, malic, and citric acids were measured in mg/g fruit weight, TA was in % citric acid equivalents, and brix was in % soluble solids.

Population		Quinic	Malic	Citric	TA	Brix
CM92	Quinic				***<.0001	
CM93					**<.005	
CM151					*<.05	
CM155						
CM92	Malic	-0.235**				
CM93		-0.208*				
CM151		0.553***				
CM155		0.791***				
CM92	Citric	-0.280***	0.092			
CM93		-0.143	0.068			
CM151		0.24	0.248			
CM155		0.342*	0.148			
CM92	TA	-0.393***	0.547***	0.503***		
CM93		-0.241**	0.558***	0.653***		
CM151		0.547***	0.743***	0.772***		
CM155		0.828***	0.860***	0.587**		
CM92	Brix	0.232**	0.000	0.163**	-0.126	
CM93		0.285**	-0.02815	-0.07764	-0.0366	
CM151		0.160	0.442**	0.260	0.327*	
CM155		0.178	0.204	0.345*	0.334	

Table 5.4: A) Phenotypes and marker genotypes (SSRs and KASPs) for the *cita* trait in CM92, CM93, CM155, and CM151. B) Phenotypes and marker genotypes (KASPs) for the *mala* trait in CM92, CM93, CM155, and CM151. Sensitivity and specificity were calculated based on ability to detect *cita/cita* or *mala/mala* individuals.

A)

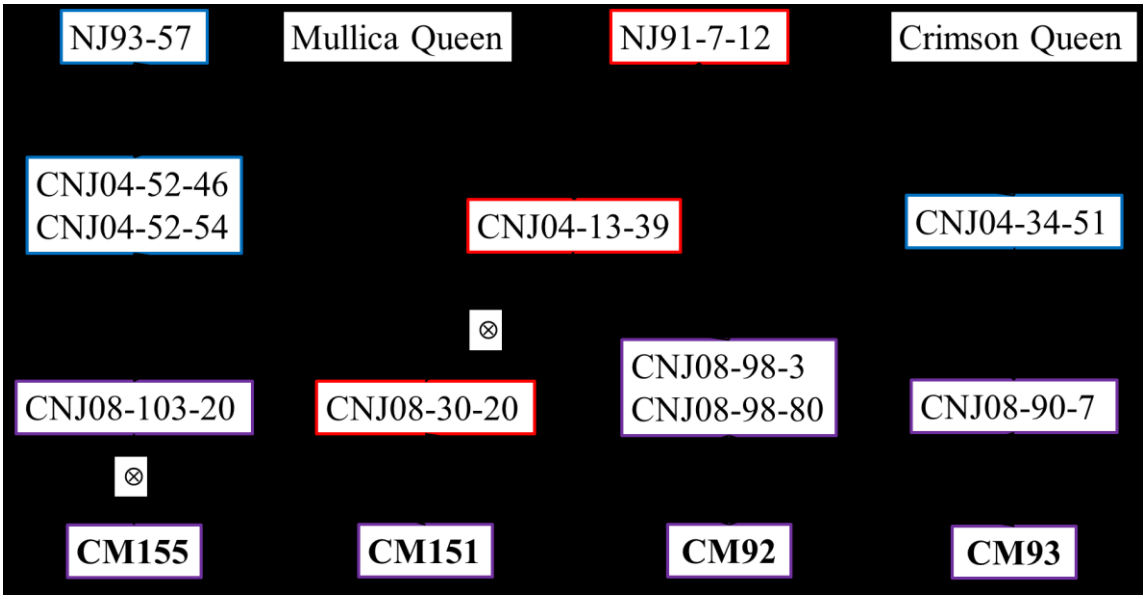
	Populations											
Markers	CM92		CM93		CM155			CM151			Sensitivity	Specificity
scf258d	217/217	217/227	217/217	217/227	217/217	227/227	217/227	217/217	227/227	217/227		
Low CA	55	4	38	1	7	0	2	9	0	0	0.940	0.994
Heterozygous	0	53	1	45	0	2	15	0	2	18		
Normal CA					0	5	5	0	9	3		
scf157322	276/276	272/276	276/276	270/276								
Low CA	54	4	37	2							0.938	0.971
Heterozygous	0	54	3	48								
SNP CA_609	G/G	A/G	G/G	A/G	G/G	A/G	A/A	G/G	A/G	A/A		
Low CA	53	5	36	3	6	2	0	9	0	0	0.912	0.982
Heterozygous	2	52	1	49	0	16	2	0	20	3		
Normal CA					0	5	5	0	3	12		

B)

	Populations													
Markers	CM92			CM93			CM155			CM151			Sensitivity	Specificity
SNP MA_271	T/T	T/C	C/C	T/T	T/C	C/C	T/T	T/C	C/C	T/T	T/C	C/C		
Low MA	0	0	0	0	0	0	11	0	0	5	1	0	0.941	0.811
Heterozygous	21	28	8	14	15	16	0	12	5	1	22	5		
Normal MA	9	26	22	6	6	31	0	5	3	0	6	9		
SNP MA_476		A/C	C/C		A/C	C/C	A/A	A/C	C/C	A/A	A/C	C/C		
Low MA		0	0		0	0	11	0	0	5	1	0	0.941	1.000
Heterozygous		43	14		13	32	0	12	5	0	23	5		
Normal MA		18	39		29	16	0	5	3	0	6	9		

Figure 5.1: Pedigree of CM155, CM151, CM92, and CM93. Red boxes indicate *cita* in background, blue boxes indicate *mala* in background, and purple boxes indicate both *cita* and *mala* alleles in background.

A)



B)

Population	Maternal	Paternal	Maternal Genotype	Paternal Genotype
CM92	CNJ08-30-20	CNJ08-90-7	*cita/cita Mala/Mala*	*Cita/cita Mala/mala*
CM93	CNJ08-30-20	CNJ08-98-3	*cita/cita Mala/Mala*	*Cita/cita Mala/mala*
CM151	CNJ08-98-80	CNJ08-103-20	*Cita/cita Mala/mala*	*Cita/cita Mala/mala*
CM155	CNJ08-103-20	selfed	*Cita/cita Mala/mala*	N/A

Figure 5.2: Density distributions showing phenotypic distributions for CM151, CM155, CM92, and CM93. The x-axis is for each trait, where quinic acid, malic acid, and citric acid are in mg/g fresh fruit, TA is in % citric acid equivalents, and brix is in % soluble solids.

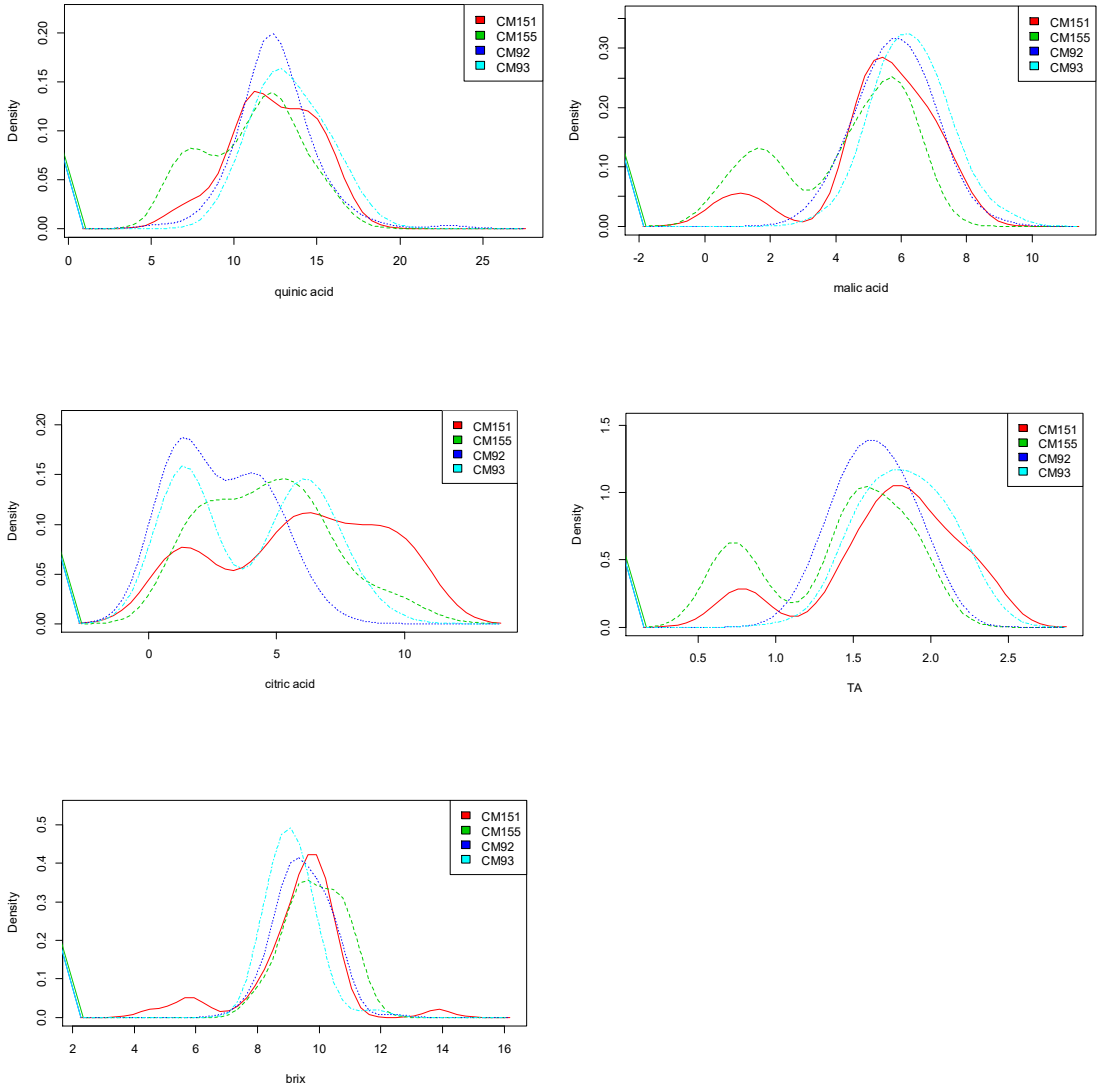
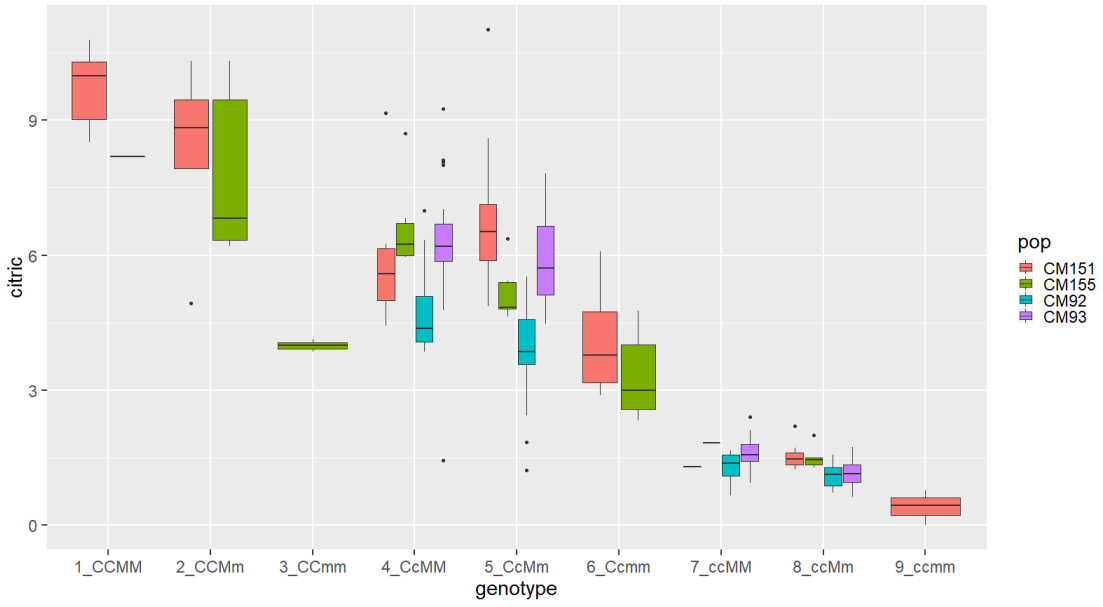
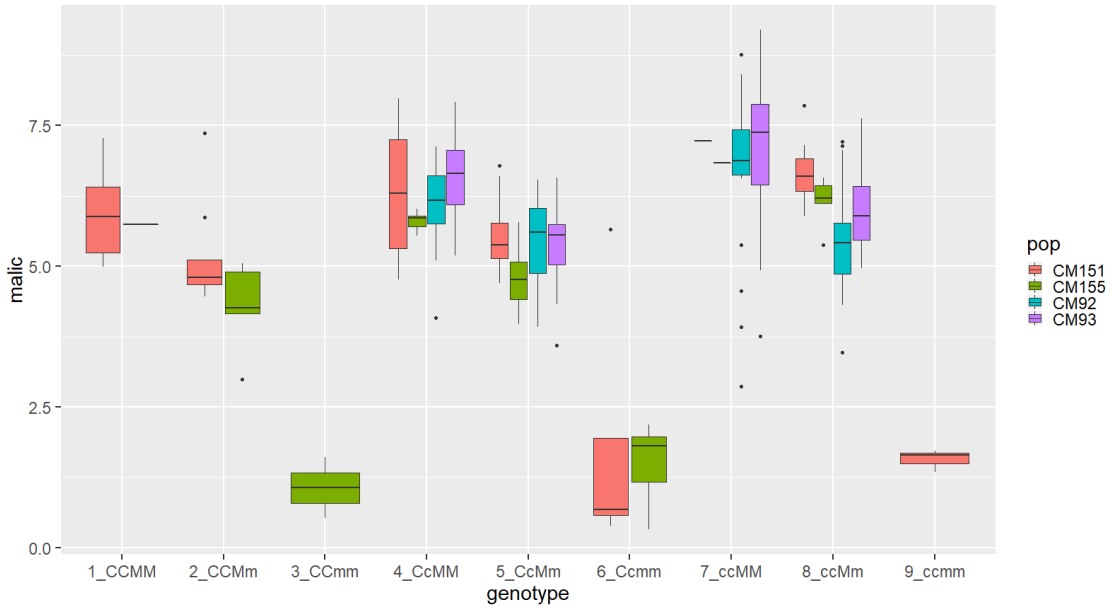


Figure 5.3: Distribution of phenotypes A) citric acid, B) malic acid, and C) TA based on the consensus genotype from genotyping with SSR and KASP markers. C is for *Cita*, c is for *cita*, M is for *mala*, and m is for *mala.*

A)



B)



C)

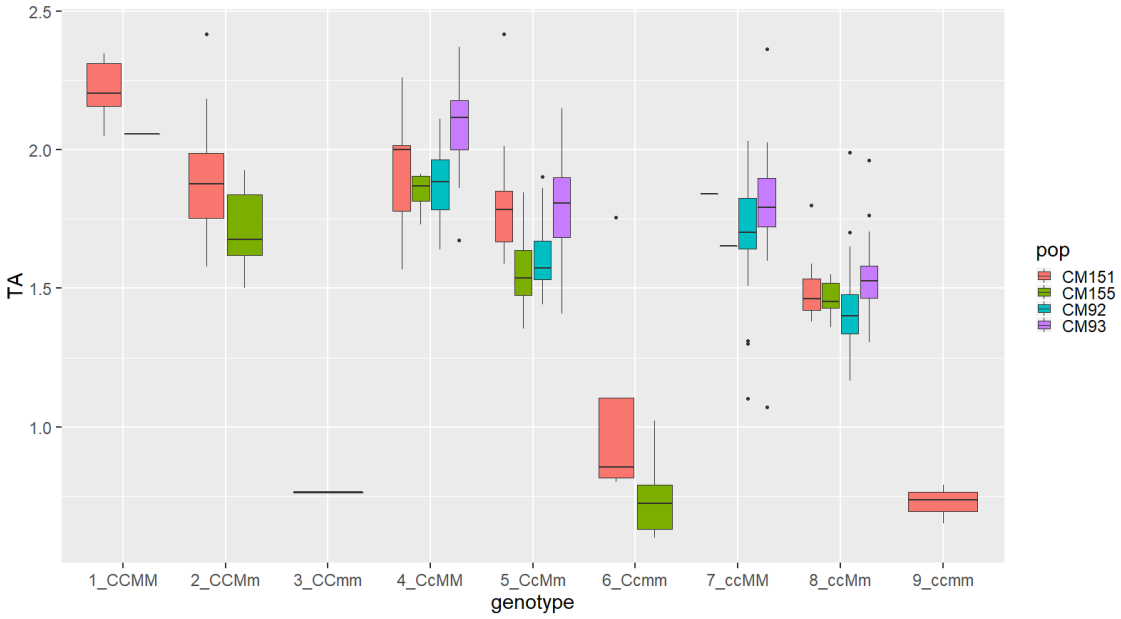
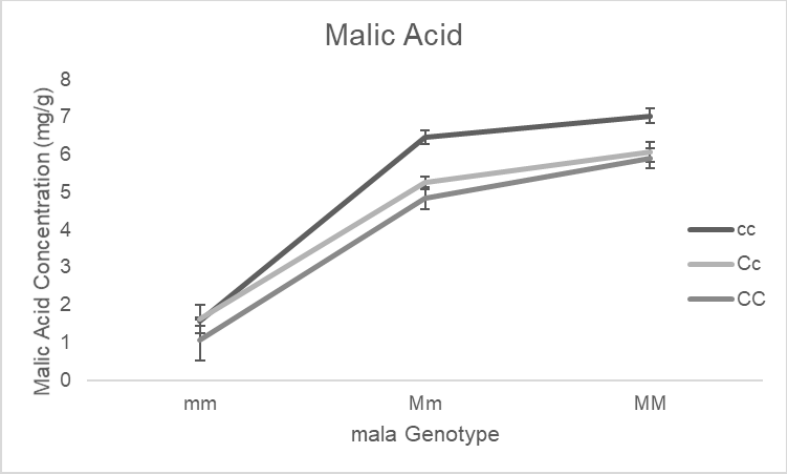
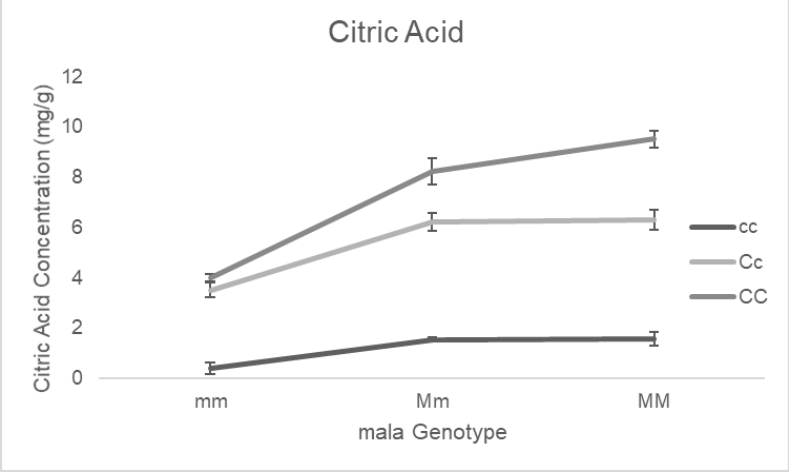

2.5
2.0
1.5
1.0
TA
1_CCMM
2_CCMm
3_CCmm
4_CcMM
5_CcMm
6_Ccmm
7_ccMM
8_ccMm
9_ccmm
genotype
pop
CM151
CM155
CM92
CM93

Figure 5.4: Interaction between *mala* and *cita* genotypes for CM151 and CM155 combined. CC is Cita/Cita, Cc is Cita/cita, cc is cita/cita. MM is Mala/Mala, Mm is Mala/mala, mm is mala/mala. A) The effect of *mala* and *cita* on malic acid concentration in mg/g fruit weight. B) The effect of *mala* and *cita* on citric acid concentration in mg/g fruit weight. C) The effect of *mala* and *cita* on TA, as measured as citric acid equivalents. Error bars are calculated as standard error.

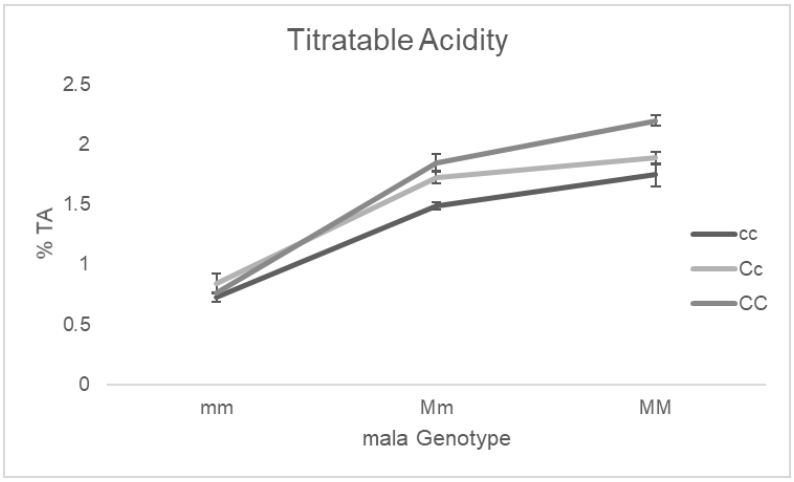
A)



B)



C)


Titratable Acidity
2.5
2
1.5
1
0.5
0
% TA
mm
Mm
MM
mala Genotype
cc
Cc
CC