BROADENING THE BASE OF APPLE GENETICS: A STUDY OF TRAITS USING MALUS PRUNIFOLIA

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Andrew Scheldorf

August 2023

BROADENING THE BASE OF APPLE GENETICS: A STUDY OF TRAITS USING MALUS PRUNIFOLIA

Andrew Scheldorf, Ph. D.

Cornell University 2023

Introgression of genes from the 25 to 33 species of apple provides the opportunity for unique traits in genetic improvement, such as disease resistance, increased antioxidants, unique phenolic compounds, and interesting flavor profiles and tannins. However, the use of wild species has been relatively limited, so the understanding of the inheritance and control of these traits are also limited. Introgression of wild species also offers the opportunity to understand the control of traits that are mostly fixed in the domesticated gene pool, because of domestication or later intentional breeding. However, the primary challenge with introgression is that fruit of wild species are often much smaller, with inferior fruit quality, potentially requiring generations of modified backcrossing for quality. New cultivars must be an improvement over what is currently on the market. This includes disease resistance, distinctive qualities, new flavor profiles, enhanced nutrition, long storage, or superior processing ability.

In this project polygenic controls of economically important traits such as fruit acidity, fruit size and shape, phenolic concentrations as well physiological controls of fruit storage were found from an interspecific population with *M. prunifolia* (PI

589816) as a parent. Evidence was found for a more complicated control of acidity in apple than currently described including transcription factors identified on chromosome 8 in previously identified regions. Genetic controls of fruit quality traits such as fruit size and phenolic concentration were identified and confirmed from previous studies. Genetic associations for fruit storage time and the associated pathways of senescence and ethylene biosynthesis in a unique background were found. Together this research discusses the potential usefulness of *M. prunifolia* (PI 589816) as a breeding parent, as well as importantly adding context to the traits that have been selected upon during domestication, and considerations for apple breeders when pursuing interspecific solutions in apple breeding.

BIOGRAPHICAL SKETCH

Andrew Delvin Scheldorf was born in the summer of 1995 in rural MN, the second child of Jeff and Teresa Scheldorf. Growing up in a many generations farm family in a farming community they were surrounded by plants from a young age. As a child they spent many hours in the garden with their mom, learning science in the form of gardening practices and old wife's tales. Much of their time was spent outside as they grew up, gardening and tending to a growing number of trees, bushes, brambles, and hardiness trials they convinced their family to set up with them. Deciding to study horticulture they enrolled at North Dakota State University in Fargo, ND in the Horticulture Science program. There they became involved in the Horticulture and Forestry Club, as well as in research, working as a lab and field technician for Dr. Todd West in the Woody Plant Improvement program and Dr. Harlene Hatterman-Valenti in the High Value Crops program. While attending NDSU they interned at Ball Horticulture as a Plug Production Technology Intern, traveling between production facilities in Denver Colorado, and Chicago Illinois to develop germination tracking software. They also studied at Ecole Supérieure d'Agricultures, in Angers, La Loire, France, and at Barrult Horticulture La possonaire, La Loire, France solidifying their love of the outdoors and of travel.

Upon graduating from NDSU they began work at Rose Innovations in Milwaukee Wisconsin, employed as the Assistant Hybridizer and Director of Secondary Breeding Programs while they waited to hear about their acceptance to Cornell University.

They were accepted to Cornell in 2018 in the Apple Breeding Program led by Dr. Susan Brown where their project started examining interspecific apple genetics and fruit quality traits. In 2020 to pursue different collaborations and skills they joined the newly formed Plant Physiology Lab for Climate Change under the direction of Dr.

Jason Londo. During their studies they have had the opportunity to create lasting connections and collaborations at Cornell and at institutions around the country, as well as foster and build a stronger community for LGBTQIA+ students acting in many DEI related student leadership positions.

For all the apples that weren't given a chance to grow

ACKNOWLEDGMENTS

I am a firm believer that nothing happens in a vacuum, and that I would not be here without a lot of help and assistance over my life.

I am immeasurably grateful to **Dr. Jason Londo** for the advice and guidance, as well as taking a risk by picking up a student who had a project that was primarily set, even if it wasn't the original plan for the lab. He gave me freedom to pursue not just science but advocacy and community work to build a stronger community in Geneva and in Ithaca. Completing my PhD with such a role model as my advisor taught me that not only do I deserve to be in plant science, but that representation matters, and it is important to be visible in a world that wants us to be small and invisible.

Endless thanks to **Hanna Martins**, our fantastic technician, who without her organization and endless drive the lab would surely fall apart. I can't say how many hours she has spent puzzling through protocols and programs with me, as well as collecting data and keeping things running when I am not in Geneva.

Thanks to **Felex Pike**, for the assistance in the field and hours of gross work in the root cellar and cooler. Working with you often kept me sane and encouraged me to keep going. Our conversations about all sorts of things made the repetition of measuring thousands of apples tolerable and often even enjoyable. I can't tell how important having you working alongside me was.

Thanks to my lab-mates, **Hongrui Wang** and **Kenneth Buck** for assistance harvesting and processing fruit, and special thanks to Hongrui for the assistance with RNAseq and data processing. Also big thanks to the Grape Breeding lab and the cast of characters led by **Dr. Bruce Reisch** for assistance and discussions over coffee in many morning meetings.

I would like to thank my committee members, **Dr. Ben Gutierrez**, **Dr. Kerik Cox**, and **Dr. Michael Gore** for the insights and challenges to my research. Special thanks to Dr. Ben Gutierrez who previously researched this population and provided me lab space, resources, and equipment to chase hairbrained ideas in the hope that they gave results.

I am grateful for my many collaborators in Geneva and Ithaca who ran experiments and sped up collection with their expertise and knowledge. Thanks to **Michael Brown** and **Dr. Greg Peck** in Ithaca for measuring the many phenolic concentrations in the population. Thanks to **Dr. Lailiang Cheng** as well for assisting in the measurement of acids and providing protocols and guidance on HPLC.

Many thanks to **Dr. Susan Brown** and **Kevin Maloney** in the Apple Breeding lab. Dr. Brown provided guidance and expertise that can only be given from a real expert in the field with years of experience, as well as taught me how to chase an idea that excites you and is interesting, even if it doesn't immediately give results. She also gave me a foothold in apple research that proved invaluable in my research. Keven was always there to provide support and assistance and put up with my often schedule disrupting plans and ideas, as well as providing knowledge and expertise on apples that was invaluable to my training.

Thanks to the USDA technicians **Amy Szewc-McFadden** and **Kas Deys**, who genotyped much of my population and provided many hours of expertise and technical skills. Thanks to **Tori Meakem**, USDA PGRU technician for her assistance in the lab, and for working through endless concentration and dilutions with me while I made a mess of her organized lab.

Finally, I am grateful to the countless people in my community and family both chosen and biological that have supported me and encouraged me to keep going, along with giving me a space to relax and be myself. **Mickie Quinn** and **Heather**

Sandford, who have been role models, teachers, co-conspirators, and endless sources of knowledge and wisdom, and who without, I know I wouldn't have made it this far. You two and our circle mean so much to me, and I am so wildly proud of everything we have accomplished. **Crissi Dalfonso** at the LGBT Resource center who created a welcoming and comfortable place to work, could deftly change hats as we both navigated our various roles in the community, and was always there to collaborate or assist in projects on and off campus. Thanks to **my parents**, who instilled a love of learning and a drive that was necessary to endure grad school. You and the rest of the family kept my life in perspective and kept me grounded and reminded me why I started the pursuit of a PhD.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

PREFACE

Apple (*Malus* spp.) is one of the most popular fruit crops produced and consumed worldwide, ranking third in global fruit consumption due to its high nutritional properties, cultural significance, and tolerance to the stresses and damages of shipping and storage and shipping (Spengler, 2019). The global annual production of apples has doubled from 41 million tons in 1990 to 2018 (FAO 2020). The apple species that is most familiar to the public is *Malus domestica* (Borkh.), which resulted from introgressive hybridization between several different *Malus* species, the primary progenitors being the Asian *Malus sieversii* (Ledeb.) M.Roem. and the European Crabapple *Malus sylvestris* (L.) Mill. (Duan et al., 2017). Initially, fruit size is believed to have been selected upon as the method of dispersal for apples changed from birds, which favor small fleshy fruits, to larger mammals. Three to four millennia ago, as supported by anthropological data (Spengler, 2019) it is believed that humans began to intentionally select specific apples and transport them along the Silk Road (Duan et al., 2017). During these initial stages of domestication, there is often a phase of unconscious selection across a wide pool of traits that results in a divergence from their wild progenitor species (Gross & Olsen, 2010). Recent research into the history of apple domestication shows evidence for this theory represented by gene fixation in the genomic regions associated with fruit quality traits and other commercially important genes. For example, red skin color, fruit size/weight, and acidity were selected from, and originated entirely, from one of the progenitor species, either *M. sieversii* or *M. sylvesteris*, demonstrating the potential to introgress traits unique to

other *Malus* species (X. Sun et al., 2020). Introgression of genes from the 25 to 33 species (there is disagreement about the number of *Malus* species (Brown, 2012)) of apple provides the opportunity for unique traits to be incorporated into our domesticated cultivars through genetic improvement, such as disease resistance, increased antioxidants, unique phenolic compounds, and interesting flavor profiles and tannins (Ma et al. 2017). However, the use of wild species has been relatively limited, and consequently, the understanding of the inheritance and control of these traits are also limited (Pereira-Lorenzo et al. 2018). "Introgression (or "introgressive hybridization") describes the incorporation (usually via hybridization and backcrossing) of alleles from one entity (species) into the gene pool of a second, divergent entity (species)" (Anderson & Hubricht, 1938). Introgression of domesticated crops with wild species also offers the opportunity to understand the control of traits that are mostly fixed in the domesticated gene pool, because of domestication or later intentional breeding. However, the primary challenge with introgression is that fruit of wild species are often much smaller and with inferior fruit quality, potentially requiring generations of modified backcrossing to regain quality. In order for new cultivars to be successful at market they must be an improvement in quality from what already exists (Brown and Maloney 2004).

The majority of introgression work has centered around incorporating disease resistance genes from wild sources, such as apple scab [*Venturia inaequalis* (Cooke)G)], Powdery Mildew [*Podosphaera leucotricha* (Ellis & Everh.)], fire blight [*Erwinia amylovora* (Burrill)], Blue Mold [*Penicillium expansum* (Link)] and Bitter Rot (*Colletotrichum acutatum* [J.H.Simmonds]) (Bus et al., 2019; O. Emeriewen et al.,

2014; O. F. Emeriewen et al., 2023; Gessler et al., 2006; Gessler & Pertot, 2012; Harshman et al., 2017; Jurick et al., 2011; Kellerhals et al., 2017; Khajuria et al., 2018; Knight & Alston, 1968; H. T. Liu et al., 2011; Luo et al., 2019, 2020; McClure et al., 2018; Norelli et al., 2017; Papp et al., 2020; Peil et al., 2020). The potential of different wild apple species to introduce disease resistance into the *M. domestica* gene pool has been successfully demonstrated by Sestras et al. 2011, using five wild species (*Malus coronaria, M. floribunda, M. niedzwetzkyana, M. zumi* and *M. prunifolia*) as well as six *M. domestica* cultivars ('Cluj 218/2', 'Frumos de Voinesti', 'Reinette Baumann', 'Rosu de Cluj', 'Jonathan', and 'Golden Delicious') to look at the inheritance of resistance to apple scab and powdery mildew. From these crosses they demonstrated that progeny derived from *M. prunifolia* and *M. coronaria* inherited scab resistance, while *M. coronaria* progeny also showed resistance to powdery mildew. While disease resistance sources have been identified and heritability of resistance has been explored, the time necessary to introgress these traits and maintain fruit quality is often seen as prohibitive. For example, 'Prima' the first cultivar released with apple scab resistance from *M. floribunda* 821, took 56 years from initial cross to cultivar release (Le Roux et al., 2012).

Due to the long generation time between crosses in apple, a deeper understanding of the genetic architecture of traits is desired for faster selection (Kenis et al., 2008). With the rapid development of next generation sequencing methods, a new age of genetic mapping has become the starting point for rapid selection and breeding for new cultivars (Peace et al., 2019). Genetic linkage maps have been made for *M. domestica* crosses, such as 'Fiesta' × 'Discovery" ((Liebhard et al. 2003); (Maliepaard

et al. 1998)) and ‘Honeycrsip’ ((Howard et al. 2017)) ‘Gala’ x ‘Jonathan’ (Oh et al., 2023), Jiguan’ x ‘Wangshanhong’ (Ma et al., 2016). Linkage maps have also been created with wild *Malus* species crossed with *M. domestica* cultivars, typically used in rootstock breeding with a focus on rootstock important traits; for example a cross made from rootstock cultivar Malling 9 (M.9) x *Malus robusta* 5 (R.5), enabled the identifications of the dwarfing trait derived from M.9 and resistance to fire blight, apple scab, and wooly apple aphid [*Eriosoma lanigerum* (Hausmann)] derived from the traits of R.5 (Celton et al. 2009)). Fernández-Fernández et al. (2008) used sequence-tagged site methods (STS) to develop a genetic map using *Malus pumilla* x ‘Totem’ and mapped scab resistance (*Vf),* mildew resistance (*Pl-2*), columnar habit, (*Co*), red tissues, and green flesh, as well as an ethylene receptor gene. In 2011 Moriya et al. (2011) used a progeny population of ‘JM7’ (*Malus prunifolia* × *Malus pumila* ‘Malling 9’) x *Malus sieboldii* ‘Sanashi 63’ to construct a linkage map to be used in rootstock breeding and marker assisted selection. This study was the first time a *M. sieboldii* individual was used in a mapping study. A genetic map of *Malus fusca* accession MAL0045 was made in 2020 using MAL0045 x *M. domestica* ‘Idared’ to elucidate a novel fire blight resistance locus, only 46.4% of the total SNP reads aligned to the ‘Golden Delicious’ reference genome demonstrating the large genetic difference between *M. domestica* and *M. fusca* (O. F. Emeriewen et al., 2020).

As a progenitor species, *M. sieversii* has been explored more recently with a focus on rootstock and growth-related traits. Using 110 F_1 individuals of ‘Red Fuji’ x ‘Hongro’ (*Malus sieversii*) a map was developed to identify 17 QTLs in *M. sieversii* for tree height, stem height, stem thickness, new shoot length, shoot diameter, internode

number, internode length, and lenticel density ((Liu et al. 2016). Disease resistance was again a focus with Desnoues et al. (2018), who used a population of 169 progeny and parents 'Royal Gala x *M. sieversi* "Kaz 95 18-07" (PI 613981) to identify 13 strain specific minor QTLs for fire blight resistance descending from *M. sieversii.* The first resistance genes to Blue Mold (*Penicillium expansum*) were also recently identified and originate from *M. sieversii,* though no cultivars have been released to date using these sources (Ballester et al., 2017; Lichtner et al., 2020; J. Sun et al., 2

The purpose of the research presented in the following chapters of this dissertation is to add to our understanding of economically important apple traits such as fruit acidity, fruit size and shape, phenolic concentrations as well as genetic and physiological controls of fruit storage. This research leveraged an existing bi-parental, full sibling population of *M. domestica* (NYCo1) x *M. prunifolia* (PI 589816), hereafter referred to as NYCo1 and 589816, grown on their own roots at the Lucy Research Farm located at Cornell Agritech in Geneva, NY (42.868536, -77.044668). The cross was made in 2013 by Dr. S. K. Brown as a part of the Cornell Apple Breeding Program. From this cross, 600 seeds were collected, stratified and planted once the radical emerged. These seedlings were grown in the greenhouse for 2 years before being planted outside.

Chapter 1 of this dissertation explores the trait of fruit acidity and its polygenic control, with evidence presented for a more complicated control of acidity than in apple than currently described. Chapter 2 of the dissertation examines variation in fruit quality traits and investigates traits of importance when considering interspecific apple breeding. Finally, Chapter 3 of the dissertation delves into the trait of fruit storage, it's

genetic control and the associated pathways of ethylene biosynthesis and of fruit senescence. Together this research aims to discuss the potential usefulness of *M. prunifolia* (PI 589816) as a breeding parent, as well as importantly adding context to the traits that have been selected upon during domestication, and considerations for apple breeders when pursuing interspecific solutions in apple breeding.

REFERENCES

Anderson, E., & Hubricht, L. (1938). Hybridization in Tradescantia. III. The Evidence for Introgressive Hybridization. *American Journal of Botany*, *25*(6), 396–402. https://doi.org/10.1002/j.1537-2197.1938.tb09237.x

Ballester, A. R., Norelli, J., Burchard, E., Abdelfattah, A., Levin, E., González-Candelas, L., Droby, S., & Wisniewski, M. (2017). Transcriptomic response of resistant (Pi613981–malus sieversii) and susceptible ("royal gala") genotypes of apple to blue mold (penicillium expansum) infection. *Frontiers in Plant Science*, *8*(November), 1–16. https://doi.org/10.3389/fpls.2017.01981

Brown, S. K., & Maloney, K. E. (2004). BIOTECHNOLOGY IN AGRICULTURE SERIES. In *Biotechnology In Agriculture Series* (pp. 475–511).

Bus, V. G. M., Scheper, R. W. A., Walter, M., Campbell, R. E., Kitson, B., Turner, L., Fisher, B. M., Johnston, S. L., Wu, C., Deng, C. H., Singla, G., Bowatte, D., Jesson, L. K., Hedderley, D. I., Volz, R. K., Chagné, D., & Gardiner, S. E. (2019). Genetic mapping of the European canker (Neonectria ditissima) resistance locus Rnd1 from Malus 'Robusta 5.' *Tree Genetics & Genomes*, *15*(2), 25. https://doi.org/10.1007/s11295-019-1332-y

Celton, J.-M., Tustin, D. S., Chagné, D., & Gardiner, S. E. (2009). Construction of a dense genetic linkage map for apple rootstocks using SSRs developed from Malus ESTs and Pyrus genomic sequences. *Tree Genetics & Genomes*, *5*(1), 93–107. https://doi.org/10.1007/s11295-008-0171-z

Desnoues, E., Norelli, J. L., Aldwinckle, H. S., Wisniewski, M. E., Evans, K. M., Malnoy, M., & Khan, A. (2018). Identification of Novel Strain-Specific and Environment-Dependent Minor QTLs Linked to Fire Blight Resistance in Apples. *Plant Molecular Biology Reporter*, *36*(2), 247–256. https://doi.org/10.1007/s11105-018-1076-0

Duan, N., Wang, N., Xiang, L., Mao, L., Legall, N., Mao, Z., Fazio, G., Wang, K., Sun, H., Wang, X., Wang, Y., Liu, J., Xu, Y., Jiao, C., Wang, D., Yin, C., Jiang, S., Fei, Z., Wan, S., … Chao, T. (2017). Genome re-sequencing reveals the history of apple and supports a two-stage model for fruit enlargement. *Nature Communications*, *8*(1). https://doi.org/10.1038/s41467-017-00336-7

Emeriewen, O. F., Richter, K., Berner, T., Keilwagen, J., Schnable, P. S., Malnoy, M., & Peil, A. (2020). Construction of a dense genetic map of the *Malus fusca* fire blight resistant accession MAL0045 using tunable genotyping-by-sequencing SNPs and microsatellites. *Scientific Reports*, *10*(1), 1–11. https://doi.org/10.1038/s41598-020-73393-6

Emeriewen, O. F., Richter, K., Wensing, A., Malnoy, M., & Peil, A. (2023). Mfu16 is an unstable fire blight resistance QTL on linkage group 16 of Malus fusca MAL0045. *Journal of Plant Pathology*. https://doi.org/10.1007/s42161-022-01296-8

Emeriewen, O., Richter, K., Kilian, A., Zini, E., Hanke, M. V., Malnoy, M., & Peil, A. (2014). Identification of a major quantitative trait locus for resistance to fire blight in the wild apple species Malus fusca. *Molecular Breeding*, *34*(2), 407–419. https://doi.org/10.1007/s11032-014-0043-1

Fernández-Fernández, F., Evans, K. M., Clarke, J. B., Govan, C. L., James, C. M., Marič, S., & Tobutt, K. R. (2008). Development of an STS map of an interspecific progeny of *Malus*. *Tree Genetics & Genomes*, *4*(3), 469–479. https://doi.org/10.1007/s11295-007-0124-y

Gessler, C., Patocchi, A., Sansavini, S., Tartarini, S., & Gianfranceschi, L. (2006). Venturia inaequalis resistance in apple. *Critical Reviews in Plant Sciences*, *25*(6), 473–503. https://doi.org/10.1080/07352680601015975

Gessler, C., & Pertot, I. (2012). Vf scab resistance of Malus. *Trees - Structure and Function*, *26*(1), 1–14. https://doi.org/10.1007/s00468-011-0618-y

Gross, B. L., & Olsen, K. M. (2010). Genetic perspectives on crop domestication. *Trends in Plant Science*, *15*(9), 529–537. https://doi.org/10.1016/j.tplants.2010.05.008

Harshman, J. M., Evans, K. M., Allen, H., Wisniewski, M. E., & Norelli, J. L. (2017). Fire Blight Resistance in Wild Accessions of *Malus sieversii*. *Plant Disease*, *101*, 1738–1745. https://doi.org/10.1094/PDIS-01-17-0077-RE

Howard, N. P., van de Weg, E., Bedford, D. S., Peace, C. P., Vanderzande, S., Clark, M. D., Teh, S. L., Cai, L., & Luby, J. J. (2017). Elucidation of the 'Honeycrisp' pedigree through haplotype analysis with a multi-family integrated SNP linkage map and a large apple (*Malus ×domestica*) pedigree-connected SNP data set. *Horticulture Research*, *4*(1), 17003. https://doi.org/10.1038/hortres.2017.3

Jurick, W. M., Janisiewicz, W. J., Saftner, R. A., Vico, I., Gaskins, V. L., Park, E., Forsline, P. L., Fazio, G., & Conway, W. S. (2011). Identification of wild apple germplasm (Malus spp.) accessions with resistance to the postharvest decay pathogens Penicillium expansum and Colletotrichum acutatum. *Plant Breeding*, *130*(4), 481–486. https://doi.org/10.1111/j.1439-0523.2011.01849.x

Kellerhals, M., Schütz, S., & Patocchi, A. (2017). Breeding for host resistance to fire blight. *Journal of Plant Pathology*, *99*(Special Issue), 37–43. https://doi.org/10.4454/jpp.v99i0.3923

Kenis, K., Keulemans, J., & Davey, M. W. (2008). Identification and stability of QTLs for fruit quality traits in apple. *Tree Genetics & Genomes*, *4*(4), 647–661. https://doi.org/10.1007/s11295-008-0140-6

Khajuria, Y. P., Kaul, S., Wani, A. A., & Dhar, M. K. (2018). Genetics of resistance in apple against Venturia inaequalis (Wint.) Cke. *Tree Genetics & Genomes*, *14*(2), 16. https://doi.org/10.1007/s11295-018-1226-4

Knight, R. L., & Alston, A. F. H. (1968). Sources of Field Immunity to Mildew (*Podosphaera leucotricha*) in Apple. *Can. J. Genet. Cytol.*, *10*, 294–298. www.nrcresearchpress.com

Le Roux, P. M., Flachowsky, H., Hanke, M. V., Gessler, C., & Patocchi, A. (2012). Use of a transgenic early flowering approach in apple (Malus × domestica Borkh.) to introgress fire blight resistance from cultivar Evereste. *Molecular Breeding*, *30*(2), 857–874. https://doi.org/10.1007/s11032-011-9669-4

Lichtner, F. J., Gaskins, V. L., Cox, K. D., & Jurick, W. M. (2020). Global transcriptomic responses orchestrate difenoconazole resistance in Penicillium spp. causing blue mold of stored apple fruit. *BMC Genomics*, *21*(1), 1–15. https://doi.org/10.1186/s12864-020-06987-z

Liebhard, R., Koller, B., Gianfranceschi, L., & Gessler, C. (2003). Creating a saturated reference map for the apple (*Malus* × *domestica* Borkh.) genome. *Theoretical and Applied Genetics*, *106*(8), 1497–1508. https://doi.org/10.1007/s00122-003-1209-0

Liu, H. T., Zhang, Y. J., Li, C. M., Wang, Y., Liu, G. J., Chang, R. F., Zhao, Y. B., Xu, X. F., Han, Z. H., & Zhang, X. Z. (2011). Evaluation of the Resistance of Malus Germplasm to Bot Canker Caused by Botryosphaeria dothidea. *Journal of Phytopathology*, *159*(7–8), 511–515. https://doi.org/10.1111/j.1439-0434.2011.01794.x

Liu, Z. C., Bao, D. E., Liu, D. L., & Chen, X. S. (2016). Construction of a genetic linkage map and identification of QTL associated with growth traits in *Malus sieversii*. In *Journal of Environmental Biology* (Vol. 37).

Luo, F., Evans, K., Norelli, J. L., Zhang, Z., & Peace, C. (2020). Prospects for achieving durable disease resistance with elite fruit quality in apple breeding. *Tree Genetics and Genomes*, *16*(1). https://doi.org/10.1007/s11295-020-1414-x

Luo, F., Sandefur, P., Evans, K., & Peace, C. (2019). A DNA test for routinely predicting mildew resistance in descendants of crabapple 'White Angel.' *Molecular Breeding*, *39*(3). https://doi.org/10.1007/s11032-019-0933-3

Ma, B., Liao, L., Peng, Q., Fang, T., Zhou, H., Korban, S. S., & Han, Y. (2017). Reduced representation genome sequencing reveals patterns of genetic diversity and selection in apple. *Journal of Integrative Plant Biology*, *59*(3), 190–204. https://doi.org/10.1111/jipb.12522

Ma, B., Zhao, S., Wu, B., Wang, D., Peng, Q., Owiti, A., Fang, T., Liao, L., Ogutu, C., Korban, S. S., Li, S., & Han, Y. (2016). Construction of a high density linkage map and its application in the identification of QTLs for soluble sugar and organic acid components in apple. *Tree Genetics and Genomes*, *12*(1), 1–10. https://doi.org/10.1007/s11295-015-0959-6

Maliepaard, C., Alston, F. H., van Arkel, G., Brown, L. M., Chevreau, E., Dunemann, F., Evans, K. M., Gardiner, S., Guilford, P., van Heusden, A. W., Janse, J., Laurens, F., Lynn, J. R., Manganaris, A. G., den Nijs, A. P. M., Periam, N., Rikkerink, E., Roche, P., Ryder, C., … King, G. J. (1998). Aligning male and female linkage maps of apple (*Malus pumila* Mill.) using multi-allelic markers. *Theoretical and Applied Genetics*, *97*(1–2), 60–73. https://doi.org/10.1007/s001220050867

McClure, K. A., Gardner, K. M., Douglas, G. M., Song, J., Forney, C. F., DeLong, J., Fan, L., Du, L., Toivonen, P. M. A., Somers, D. J., Rajcan, I., & Myles, S. (2018). A Genome-Wide Association Study of Apple Quality and Scab Resistance. *The Plant Genome*, *11*(1), 1–14. https://doi.org/10.3835/plantgenome2017.08.0075

Moriya, S., Iwanami, H., Kotoda, N., Haji, T., Okada, K., Terakami, S., Mimida, N., Yamamoto, T., & Abe, K. (2011). Aligned genetic linkage maps of apple rootstock cultivar 'JM7' and *Malus sieboldii* 'Sanashi 63' constructed with novel EST-SSRs. *Tree Genetics & Genomes*, *8*(4), 709–723. https://doi.org/10.1007/s11295-011-0458-3

Norelli, J. L., Wisniewski, M., Fazio, G., Burchard, E., Gutierrez, B., Levin, E., & Droby, S. (2017). Genotyping-by-sequencing markers facilitate the identification of quantitative trait loci controlling resistance to Penicillium expansum in Malus sieversii. *PLoS ONE*, *12*(3). https://doi.org/10.1371/journal.pone.0172949

Oh, S., Ahn, S., Han, H., Kim, K., Kim, S. A., & Kim, D. (2023). Genetic linkage maps and QTLs associated with fruit skin color and acidity in apple (Malus × domestica). *Horticulture, Environment, and Biotechnology*. https://doi.org/10.1007/s13580-022-00473-z

Orcheski, B., Parker, R., & Brown, S. (2015). Pale green lethal disorder in apple (*Malus*) is caused by a mutation in the PHYLLO gene which is essential for phylloquinone (vitamin K 1) biosynthesis. *Tree Genetics & Genomes*, *11*(131). https://doi.org/10.1007/s11295-015-0956-9

Papp, D., Gao, L., Thapa, R., Olmstead, D., & Khan, A. (2020). Field apple scab susceptibility of a diverse Malus germplasm collection identifies potential sources of

resistance for apple breeding. *CABI Agriculture and Bioscience*, *1*(16), 1–14. https://doi.org/10.1186/s43170-020-00017-4

Peace, C. P., Bianco, L., Troggio, M., van de Weg, E., Howard, N. P., Cornille, A., Durel, C.-E., Myles, S., Migicovsky, Z., Schaffer, R. J., Costes, E., Fazio, G., Yamane, H., van Nocker, S., Gottschalk, C., Costa, F., Chagné, D., Zhang, X., Patocchi, A., … Vanderzande, S. (2019). Apple whole genome sequences: recent advances and new prospects. *Horticulture Research*, *6*(1), 59. https://doi.org/10.1038/s41438-019-0141-7

Peil, A., Emeriewen, O. F., Khan, A., Kostick, S., & Malnoy, M. (2020). Status of fire blight resistance breeding in Malus. *Journal of Plant Pathology*. https://doi.org/10.1007/s42161-020-00581-8

Pereira-Lorenzo, S., Fischer, M., Ramos-Cabrer, A. M., & Castro, I. (2018). Apple (*Malus* spp.) Breeding: Present and Future. In *Advances in Plant Breeding Strategies: Fruits* (pp. 3–29). Springer International Publishing. https://doi.org/10.1007/978-3-319-91944-7_1

Sestras, A. F., Pamfil, D., Dan, C., Bolboaca, S. D., Jäntschi, L., & Sestras, R. E. (2011). Possibilities to improve apple scab (*Venturia inaequalis* (Cke.) Wint.) and powdery mildew [*Podosphaera leucotricha* (Ell. et Everh.) Salm.] resistance on apple by increasing genetic diversity using potentials of wild species. In *AJCS* (Vol. 5, Issue 6).

Spengler, R. N. (2019). Origins of the apple: The role of megafaunal mutualism in the domestication of *Malus* and rosaceous trees. *Frontiers in Plant Science*, *10*(May), 1–18. https://doi.org/10.3389/fpls.2019.00617

Sun, J., Janisiewicz, W. J., Nichols, B., Jurick II, W. M., & Chen, P. (2017). Composition of phenolic compounds in wild apple with multiple resistance mechanisms against postharvest blue mold decay. *Postharvest Biology and Technology*, *127*, 68–75. https://doi.org/10.1016/j.postharvbio.2017.01.006

Sun, X., Jiao, C., Schwaninger, H., Chao, C. T., Ma, Y., Duan, N., Khan, A., Ban, S., Xu, K., Cheng, L., Zhong, G., & Fei, Z. (2020). Phased diploid genome assemblies and pan-genomes provide insights into the genetic history of apple domestication. *Nature Genetics*. https://doi.org/10.1038/s41588-020-00723-9

CHAPTER 1

ACIDITY GENETIC CONTROLS IN A *MALUS DOMESTICA* X *MALUS PRUNIFOLIA* POPULATION

INTRODUCTION

Much of the variation in the flavor of apples is determined by the interplay between perceived sweetness and acidity. The perception of acidity is primarily conveyed through combinations of different organic acids such as malic and citric acid (M. A. Khan et al., 2015), and when balanced with sweetness (brix), compose much of market acceptability. In apples these two acids have contrasting effects on the perceptions of sweetness, with citric acid masking the perception of sucrose and fructose and contributing to sour or bitter flavors, while malic acid enhances the perception of sucrose (Lobit et al., 2006). Despite some chemical diversity across cultivars (Ma, Chen, et al., 2015), malic acid is the primary acid in domesticated apples, constituting approximately 90% of the acid profile in *Malus domestica* fruit (Zhang et al., 2010). Considering that malic acid is the primary acid in *Malus,* fruit acidity is traditionally measured as pH or by titratable acidity, where a sample of juice is titrated with a known concentration of a base until a specific pH end point is reached, this volume of base used to reach this benchmark is then used to calculate the concentration of acid in the sample and is assumed to be the concentration of malic acid (Kumar et al., 2021).

Malic acid is synthesized through two different mechanistic pathways in apple fruit. The first pathway takes phosphoenolpyruvate (PEP), derived from sugar metabolism in early fruit development and converts it to oxaloacetate (OAA) in the

cytoplasm by PEP-carboxylase (PEPc), OAA is then reduced to malate in a reversable cytosolic NAD-dependent malate dehydrogenase (Tijero et al., 2021; Yao et al., 2009). In the second pathway malic acid is produced in the mitochondria through the TCA cycle, where CO_2 from respiration must be fixed (Etienne et al., 2013a; Tijero et al., 2021). Malic acid found in fruit is believed to primarily be produced in fruit tissue, through either of these pathways, or both depending on situation, with the potential for a limited amount of acid being translocated into the fruit from the leaves (Hale, 1957; Sweetman et al., 2009). It has been suggested that the most likely path for malic acid synthesis in apple fruit is by using the first pathway involving PEPc which is highly expressed during the cultivar 'Greensleeves' fruit development (M. Li et al., 2016).

Figure 1.1. Malic acid acts as an intermediate in several cellular processes like the Citric Acid cycle in the mitochondria where it is produced from fumerate by *fumerase*, or in the Calvin Cycle in the chloroplasts where it is produced from OAA in the path to make pyruvate (Sweetman et al., 2009).

The mechanism for concentrating malic acid in fruit was long believed to be influenced by a combination of external and internal conditions. There have been studies associating fruit growth stages, sucrose content, and malic acid content; potassium content and malic acid; as well as the amount of potassium fertilizer applied (Cummings & Reeves, 1971; Génard et al., 1991; Lobit et al., 2002). Inside cells acids are stored in the vacuole of fruit cells, with the specific mechanism for accumulating acid inside the vacuole is described as an "acid trap" where the malate or citrate that is produced in the neutral or slightly basic cytosol is in a negatively charged state. Once the acid flows across the tonoplast membrane into the vacuole where the pH is acidic,

the acid is pronated. Only the negatively charged acid forms can cross the membrane into the cytosol, so these positively charged molecules are “trapped” within the vacuole (Etienne et al., 2013b). Lobit et al., (2006) developed a model based on malate accumulation in peach that integrates prior studies on external effects as well as knowledge on the physiological mechanisms occurring in the plant. The model suggests that the accumulation of malate in the vacuole is an interplay between thermodynamic effects on both proton and malate tonoplast transporters.

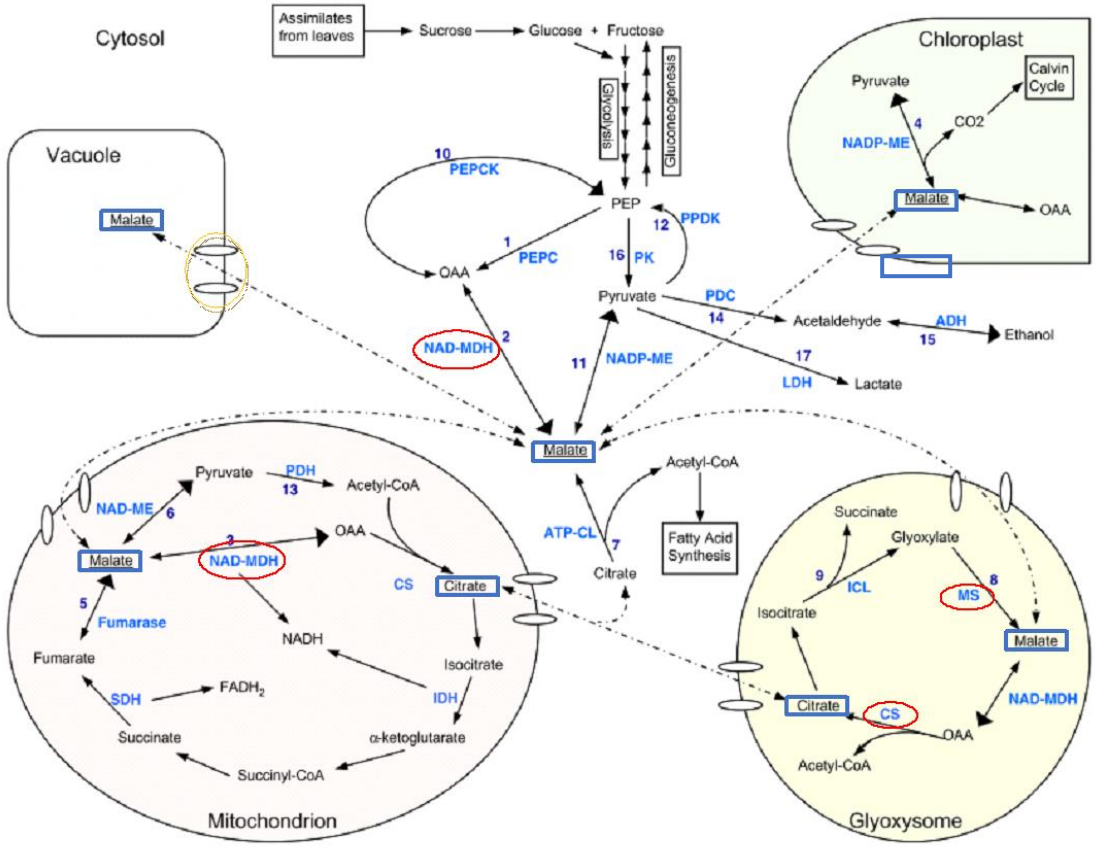


Figure 1.1 Adapted from (Sweetman et al., 2009) proposed metabolic pathways in fruits cells. Enzymes discussed in text are in bold and circled in red in figure, Circled in yellow is the location of the vacuole “acid trap” discussed, boxed in blue are sources of malic acid and citric acid; ADH, alcohol dehydrogenase; ATP-CL, ATP-citrate lyase; **CS, citrate synthase**; ICL, isocitrate lyase; IDH, isocitrate dehydrogenase; LDH, lactate dehydrogenase; **MS, malate synthase**; NAD-ME, NAD-linked malic enzyme; NAD-MDH, **NAD-linked malate dehydrogenase**; NADP-ME, NADP-linked malic enzyme; NADP-MDH, NADP-linked

malate dehydrogenase; PDC, pyruvate decarboxylase; PDH; pyruvate dehydrogenase; PEPC, phosphoenolpyruvate carboxylase; PEPCK, phosphoenolpyruvate carboxykinase; PK, pyruvate kinase; PPDK, pyruvate orthophosphate dikinase; SDH, succinate dehydrogenase.

Table 1.1 non-exhaustive list of genes currently proposed to control malic acid concentration in apple fruit.

Name	**Physical Location**	**Citation**
Ma1/Ma	Chromosome 16	(Liebhard et al., 2003); (Xu et al., 2011); Khan et al. 2013
Ma3	Chromosome 8	Verma et al. 2019
Ma4	Chromosome 6	Ban and Xu 2020
Ma5	Chromosome 1	Rymenants et al. 2020
Ma6	Chromosome 4	Ban and Xu 2020
Ma10	Chromosome 17	Ma et al. 2019b

QTLs and gene functions for malic acid

The concentration of malic acid in apple fruits was initially hypothesized to be controlled by a single gene that was first proposed and mapped using pH values in a *Malus domestica* population (n=152) of 'Prima' x 'Fiesta'. The trait segregated as a single dominate locus located on LG 16, which was designated as *Malic acid* (*Ma*) with the recessive form designated *ma* (Maliepaard et al., 1998). The *Ma* candidate locus was further validated in a cross between 'Fiesta' x 'Discovery'. In this mapping population, 42% of variation in malic acid could be attributed to the allele located at the *Ma* locus (Liebhard et al., 2003). The *Ma* region was further investigated using two half-sibling populations (n=438) with one parent being the cultivar Royal Gala, and the reciprocal parent being one of two genotypes of the wild apple species *M. sieversii* (PI613988 and PI613971) (n= 438). Using these populations, the locus was

reduced to a region of approximately 150 kb in the 'Golden Delicious' reference genome and this region contained 44 predicted candidate genes (Xu et al., 2012). The *Ma* region was further delineated using four interspecific populations of *Malus sieversii* (Ledeb.) (GMAL 4592, GMAL 4592, GMAL 4595, GMAL 4596) by 'Royal Gala' using a total of n=724 fruiting individuals. With this increased population size, the locus was narrowed to 65-82 kb, containing 12-19 predicted genes, notably the region contained an Aluminum-Activated Malate Transporter (ALMT)(Bai et al., 2012).

The *Aluminum-Activated Malate Transporter* (*ALMT*) gene family is an important family of anion channel proteins that are involved in integral physiological processes in plants such as metal tolerance, stomatal aperture function, hormonal and cellular signaling, and pH regulation (Barbier-Brygoo et al., 2011; Linlin et al., 2018; Roelfsema & Hedrich, 2005; Sasaki et al., 2004). Ma et al. (2015b) examined the ALMT gene family in apple, searching the draft genome of 'Golden Delicious' (Velasco et al., 2010) and the coding sequence of the *ALMT* family from *Arabadopsis thaliana* (Kovermann et al., 2007) and identified three gene family members, designated *ALMTI*, *ALMTII* and *ALMTIII*. An *ALMTII* family gene was identified within the *Ma1* region previously identified, as well as 21 other homologous genes located on chromosomes 1, 2, 3, 6, 11, 13, 14, 15, and 16. By screening 353 *Malus* accessions for associations between any of these homologues and acidity, it was found that the only gene that correlated with acidity was the gene located in the *Ma1* region (MDP0000252114). To further lend mechanistic support to *Ma1* as the candidate gene determining malic acid level, *Ma1* was also localized to the tonoplast membrane using

ectopic expression in yeast lines (Ma, Liao, et al., 2015). However, variation in acidity has been repeatedly observed when contrasting genotypes that are fixed for the *Ma1* locus, indicating that *Ma1* is likely not the only gene controlling acidity in *Malus*.

The recessive allele of *Ma1, ma1,* is believed to have formed from a natural mutation that leads to a premature stop codon truncating the terminus by 84 amino acids, rendering the transporter nonfunctioning (Bai et al., 2012). This truncation was further elucidated by Li et al. (2020) by comparing the functionality of *Ma1* and *ma1* expressed both in *Xenopus laevisoocytes* and *Nicotiana benthamiana* cells as well as with RNA interference (RNAi) suppression of *Ma1* expression in apple and phenotyping 186 apple accessions from 17 *Malus* species. The *ALMT* gene family member located at the *Ma1* locus was confirmed as an anion channel localized to the tonoplast membrane and named *ALMT9* (C. Li et al., 2020). This protein functions by driving the accumulation of malic acid inside the vacuole from the cytoplasm where it is then pronated. This mechanism agrees with the proposed acid trap mechanism mentioned previously. Complicating our understanding of *Ma1* is *Pale Green Lethal* (*pgl*) a recessive physiological disorder that causes the plant to be unable to produce phylloquinone resulting in morbidity at the early cotyledon stage, this gene was identified only 1.6 cM from Ma1 and is common in breeding material (Orcheski et al., 2015).

A second major malic acid concentration QTL on chromosome 8, which accounted for 46% of total variation was identified in a mapping population derived from the cultivars ‘Fiesta’ by ‘Discovery’. Together this QTL on chromosome 8 and *Ma1* (on chromosome 16) accounted for nearly all variation in acidity in this

population (Liebhard et al., 2003). Using a multi-family population that contained modern cultivars ‘Arlet’, ‘Aurora Golden Gala’, ‘Cripps Pink’, ‘Delicious’, ‘Enterprise’, ‘Honeycrisp’, ‘Splendour’ and advanced breeding selections ‘WA 5’ and ‘W1’ from Washington State University’s Apple Breeding program, Verma et al., (2019) also identified this second locus and named it *Ma3* with the recessive form (*ma3*) conferring lower acidity. When combined, *Ma1* and *Ma3* are important for titratable acidity in this mapping population, and an additive dosage model for these two QTLs accounts for the majority of the variation, for each addition of a high acid allele (*Ma1 or Ma3),* malic acid concentration increased on average 1.8mg/L. Acidity during storage was also evaluated, finding that the individuals with higher acid alleles (*Ma1, Ma3*) degraded acidity faster in storage, eventually settling on a stable level of acidity (Verma et al., 2019).

Using the same population of ‘Royal Gala’ x *M. sieversii* as Xu et al. (2011), and a bulk phenotyping approach, bins of low, regular, and high acid fruit were designated based off of fruit TA below, within, and above the acceptable range for consumer preference (3-10 mg ml-1), When pooled, two recessive loci for the high acidity phenotype were identified. These loci, one on chromosome 4 (*Ma6)* and one on chromosome 6 (*Ma4*), accounted for 20.6% and 28.5% individually (and respectively) and 50.7% combined. SSR markers were also developed for these two high acid loci (Ban & Xu, 2020).

Another minor effect locus was identified by Rymenants et al. (2020) using a trained panel of evaluators for acidic taste, and a population (n=661) of three full sibling families (‘Nicoter’ x ‘Cripps Pink’, ’Zari’ x ‘Fuji’, and ‘B3F44’ x ‘B3FS1’) at

harvest and after storage. The locus identified was located on LG 1 and has been designated *Ma5*.

The next minor malic acid gene identified (*Ma10*) explained 7.5% of acidity in homozygous recessive *Ma* individuals examined using transcriptomic data. Contrasting two homozygous individuals with differing acid concentration helped to identify this locus. The leading candidate gene in this locus encodes a vacuolar *H*+-*ATPase* (*VHA*) gene that has functions in transporting malic acid into the vacuole using energy from ATP. Ma reported that *Ma10* is physically close to a QTL for acidity located on LG 17 reported by Kenis et al. (2008) using progeny (n= 250) of 'Telamon' x 'Braeburn'. (Ma, Liao, et al., 2019), however Kenis et al. reported a QTL for acidity located on Chromosome 16, this discrepancy could be due to alignment or assembly error or updates to the reference genomes.

Finally, Jia et al., (2018) identified four QTL regions on linkage group 8 and 16, using two pedigree populations of 1800 and 2637 F1 hybrids from the cultivars 'Jonathan' x 'Golden Delicious' and *Malus asiatica* (Nakai**)** 'Zisai Pearl' x Red Fuji'. The QTL region on LG 16 contained the *ALMTI* that is the *Ma1* locus. The three QTLs identified on chromosome 8 (designated qtl08.1, qtl08.2 and qtl08.3) were all physically closely linked and contained the *MYB44*, *PP2CH*, and *SAUR37* genes respectively. An epigenetic cascade was suggested involving *SAUR37* binding to *PP2CH* and *PP2CH* acting to dephosphorylate *ALMTII* to control malic acid accumulation.

Transcription Factors

Transcription Factors (TF) have been associated with the control of fruit acidity as well. The MYB genes comprise one of the largest TF families in the plant kingdom and play integral roles in plant metabolism, plant development and hormone synthesis. *MYB1*/*10* most notably controls the biosynthesis of anthocyanin in apple but has also been connected to the regulation of proton pumps that influences the accumulation of malate in the vacuole of the cell, lowering the pH inside the vacuole. *MYB1*/*10* binds to three *VHA*s (*VHA*, *MdVHA-B1* and *MdVHA-B2*) to activate its expression, energetically pumping malic acid into the vacuole (Hu et al., 2016; Jia et al., 2021). These proton pumps are also used to transport anthocyanin into the vacuole when induced by exposure to light (Hu et al., 2016; Takos et al., 2006).

Using transcriptomic data of tissue cultured cells of the *M. domestica* cultivar 'Orin', another MYB transcription factor *MYB73* was found to bind to and activate malate transporters. Notably this MYB was observed directly binding to the *ALMT9* (*Ma1*) promotor region, as well as *MdVHA-A* (*vacuolar ATPase subunit A*) and *MdVHP1* (*vacuolar pyrophosphatase 1*) increasing their expression, resulting in higher malate accumulation in the vacuole of overexpressing apple callus (Hu et al., 2017). Variation in *MYB123* itself has been reported to moderate *Ma1* expression, with one form, designated *MYB123*, binding to the promotor region of *ALMT9* causing standard expression, and another form, designated *myb123*, binding and resulting in no expression, implicating a more complicated epistatic control of acidity in *Malus* than initially believed (Zheng et al., 2023b).

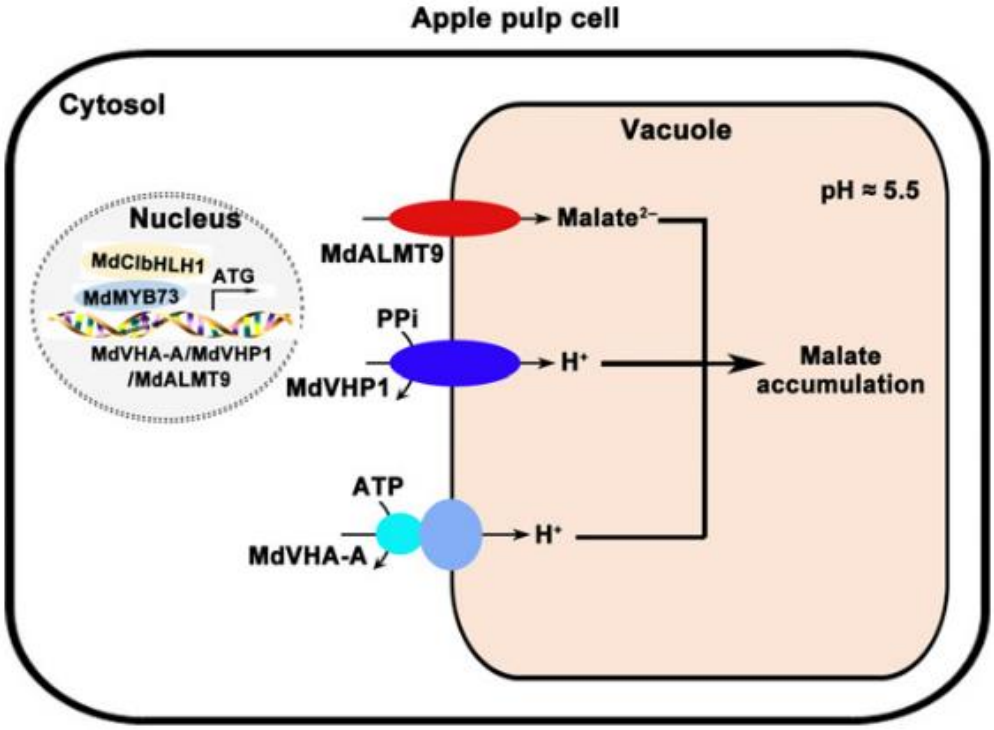


Figure 1.2. Adapted from (Hu et al., 2017) A working model demonstrating that the *bHLH* transcription factor (TF) *MdClbHLH1* interacts with the MYB TF *MdMYB73* and enhances its activity upon downstream target genes *MdALMT9*, *MdVHA-A* and *MdVHP1*, leading to increasing malate content and vacuolar acidification in apple. *VHA-A* and *VHP1* both act as proton pumps, energetically moving protons into the vacuole to maintain a charge gradient to maintain the acid-trap.

Citric Acid

The primary acid in domesticated *Malus* fruit is malic acid. However, in some individuals and especially wild species, other acids, primarily citric acid, can be found in significant levels. Citric acid is produced in the tricarboxycylic acid cycle (TCA) in the mitochondria by *mitochondrial citrate synthase* (*CS*). Like malic acid it is primarily stored in the vacuoles of cells and the concentration may be constrained by activity of citrate transporters on the tonoplast (Ma, Ding, et al., 2019).

Comparison of fruit acidity profile between wild species *Malus sargentii* (Rehder) and *Malus niedzwetzkyana* [Dieck) C.K. Schneid] revealed citric acid was present in *M.*

sargentii, but not in *M. niedzwetzkyana.* Further investigation found that in *M. sargentii, CS* proteins were more prevalent than in *M. niedzwetzkyana*, indicating that *CS* potentially plays an important role in controlling citric acid production in *M. sargentii* fruit (Ma, Ding, et al., 2019).

Ma et al. (2018) surveyed a group of 101 apple accessions (53 *M. domestica* cultivars and 58 wild relatives) for organic acids using high performance liquid chromatography (HPLC). It was found that 53% of the wild apple relatives accumulated high levels of citric acid, whereas the domesticated cultivars accumulated very little to undetectable levels. The two main progenitor species of the domesticated apple, *M. sieversii* (Ledeb.) and *M. sylvestris* [(L.) Mill.] contained little to no citric acid, suggesting that the low levels of citric acid in domesticated apples is a result of the progenitors also not accumulating citrate. The genetic control of citric acid concentration in *Malus* fruit remains unknown. Ma et al., (2018) proposed a hypothesis that a single dominate gene controls citric acid in wild Malus species based on the absence of citric acid in *M.* domestica cultivars. The genetic basis of citric acid in *Malus* remains elusive and requires more research in order to determine its genetic controls.

Domestication effects

Fruit acidity appears to have been selected upon during the domestication process, with domesticated apples having low frequency of homozygously dominant (*MaMa*) alleles and high frequency of the presumed loss of function form of the *Ma* locus (*mama*). This pattern is contrasted in the wild *Malus* species surveyed to date,

which have a high frequency of dominate homozygosity at the *Ma* locus (*MaMa*), and higher acid concentration (Ma, Chen, et al., 2015).

Evidence for selective sweeps were identified in domesticated apple using deep resequencing of 117 wild and cultivated accessions, and the origin of the sweeps from either progenitor species (*M. sieversii*, *M. sylvestris*) were able to be identified (Duan et al., 2017). The sweeps correlating to acidity descended from *M. sylvestris* only and contained the *Ma1* locus and genes annotated as *ALMT*s, *malate dehydrogenase*, and *citrate synthase* enzymes, explaining why domesticated apples have acid levels closer to *M. sylvestris* (Duan et al., 2017). Selection for lower acid by humans when compared to wild species is further supported by Spengler (2019); archaeobotanical evidence indicates that wild mammals in the Pleistocene preferred higher acid apples while human records give evidence of selection for larger and sweeter fruit.

Apple acidity is an important trait determining fruit quality and flavor. The trait appears to be primarily controlled by the *Ma1* locus, with several minor genes modifying and affecting concentration. However, the complete genetic control of malic acid concentration continues to be elusive, with QTLs explaining low to significant amounts of variability depending on the genetic background of the accessions screened. Adding to the difficulty of clarifying the controls of malic acid is that the *Ma1* locus is closely linked (1.6 cM) to the *pgl* gene, a gene which causes seedling lethality, that is prevalent in at least 35% of germplasm including founders such as ‘Golden Delicious’ ‘Jonathan’ and ‘Fuji’ resulting in disruption in progeny populations (Orcheski et al., 2015).

Objectives

The objectives of this study were to identify genes that may be unique to wild species apples using a segregating population with the parents *Malus domestica* (hereafter NY Co1) in the Cornell Apple breeding program and *Malus prunifolia* PI 589816 (hereafter 589816). Apple acidity is a trait that has been selected upon during domestication and subsequent intentional breeding, but it is not completely understood. This study provides the opportunity to examine a wild species apple for utility in future breeding by examining a mapping population derived from it for unique fruit acidity controls, in conjunction to further exploring our understanding of apple fruit acidity and its genetic control.

MATERIALS AND METHODS

Plant Material

This research was completed using a bi-parental, full sibling population of *M. domestica* (NYCo1) x *M. prunifolia* (PI 589816) grown on their own roots at the Lucy Research Farm located at Cornell Agritech in Geneva, NY (42.868536, -77.044668). The cross was made in 2013 with 600 seeds collected, stratified, and planted once the radical emerged. 150 seedlings were identified following germination carrying the trait of pale green lethal, which is expected in material heterozygous for the disorder, resulting in a population of 445 individuals (Orcheski et al., 2015). Surviving seedlings were grown in the greenhouse for 2 years before being planted outside with orchard spacing of 12’ x 3’ Parent 1, NYCo1, is heterozygous for *Co*, the columnar gene, and produces large (>80 mm) yellow fruit with a pink blush where exposed to

the sun with minimal russet outside of the stem well. Parent 2, *M. prunifolia* PI 589516, has standard architecture and produced small (<15 mm) yellow fruit, primarily with a waxy cuticle, but occasionally displays partial to full russet.

Genetic Data

DNA was extracted from this population in a collaboration between the USDA-ARS PGRU and the Apple Diversity Group of Dalhousie University (Migicovsky et al., 2016). Briefly, single nucleotide polymorphisms (SNP) were identified using genotyping-by-sequencing (GBS). Libraries were constructed using *Ape*KI restriction enzyme and Illumina HiSeq 2000 (Illumina Inc., San Diego CA, USA) sequences ere completed at the Genomic Diversity Facility, Cornell University, Ithaca NY, USA (Gutierrez et al., 2018).

All individuals were scored using genotyping-by-sequencing (Elshire et al., 2011) using the methods described by (Gutierrez et al., 2018). Sequence tags were aligned to the *M. domestica* GDDH13 Whole Genome v1.1 (https://www.rosaceae.org/) using Bowtie 2 (Langmead & Salzberg, 2012) with parameters D, R, N, L and i set to 30, 5, 1, 15, and S, 1, 0.25 to reduce misalignment. The Tassel 5 (Glaubitz et al., 2014) GBS pipeline was used for SNP calling. Ten replicates of parents NYCo1 and PI 589816 were merged into single genotypes. VCFtools (Danecek et al., 2011) was used to filter data; twenty-six individuals were removed for low mean depth; sites were filtered for read depth of 8 and minor allele frequency of 0.20; and genotypes were filtered for 80% missing data. Genotypes were filtered based on Mendelian error calculated in PLINK 1.9 (Purcell et al., 2007) and

set to missing. SNP markers were filtered based on chi-square tests for segregation distortion and > 0.95 similarity using JoinMap4.1.

Genetic Association

Marker-trait association analyses between 51442 Genotype-by-Sequencing markers and measured traits were performed using TASSEL v5 software (Bradbury et al. 2007). Using the Q method, a general lineal model (GLM) was done using numeric data joined to genotype data and principal component analysis (PCA) as population membership estimates, along with Ma1 and Ma3 genotypes, and tree architecture as covariates (Ma1Ma1 = 1, Ma1ma1 =0; Ma3ma3 = 1, ma3ma3 = 0; Coco = 1, coco =0). Data was exported from TASSEL and results were visualized using the 'qqman' R package (D. Turner, 2018) as Manhattan plots and QQ plots. Significant SNPs were determined using the Bonferroni-corrected threshold for GWAS (Log of Odds > 7).

DNA Extraction for Marker Based Genotyping

Fresh new leaf tissue was collected in the summer of 2021 from all members of the population and the parents and placed into 96 well tubes containing disruption beads. The tissue was kept cold on ice until they could be placed into -80 C to freeze. Once the tissue was frozen the samples were placed into a tissue homogenizer (miniG Genogrinder, SPEX) and shaken until pulverized. DNA was extracted using the Qiagen DNAeasy 96 Plant Extraction Kit using the standard protocol.

Juice Collection

Fruit were harvested in 2019, 2020, and 2021 from the Lucy farm at Cornell Agritech, located in Geneva NY (42.868536, -77.044668). Trees that appeared unhealthy or had insufficient fruit were excluded from collection. A small (n=18)

number of trees fruited in 2018 but data wasn't analyzed. Fruit was harvested starting at the beginning of September and extending to the end of October each year. Ripeness was determined by evaluating the ease of abscission from the tree as well as ground color change, defined by a change in peel color from the color in the stem well to a brighter green or yellow. At least 30 fruit were randomly harvested from each tree, harvesting from different regions of the tree canopy. The fruit was placed into paper bags to allow gas exchange and labeled with the row number and tree number; this number was then used to track the fruit through data collection. In 2021 five fruit from every individual harvested were checked for maturity using the starch pattern index (SPI) assay. Cortex SPI was determined by staining the stem side of a cross section taken at the equator of the apples with iodine solution (0.22 w/v iodine, 0.88% w/v potassium iodine) and visual rating using the Cornell Starch Iodine rating where 1 indicated 100% staining and 8 indicated 0% staining, with higher levels of staining indicating more starch and conversely less ripe. (Blanpied and Silsby, 1992)

The fruit were stored at 4° C in a commercial storage room under ambient atmospheric gases for 1-3 weeks until processing. Processing was completed within 2 weeks following the date of harvest. A subsample of 10 fruit that were deemed to be representative of the fruit collected were juiced with a juicer (ACME Supreme Juicerator 6001) using a Kim-wipe as a disposable filter to reduce fruit flesh in the juice sample. After juicing, 50-mL aliquots were made and stored at -20 °C until juice chemistry was able to be evaluated.

Juice titratable acidity (TA) was determined by titrating samples of 10 ml of juice in a 50 ml dilution with 0.1 N NaOH to pH 8.2 using an autotitrator (Metrohm

848 Titrino Plus and Metrohm 869 Com-pact Sample Changer, Herisau, Switzerland). The fruit TA (g/L) was calculated based on the formula by Nielsen, (2017), where the equivalent weight of malic acid is 67.04 g. pH was taken of pure juice using a pH meter (Unitrode pH meter Metrohm,Herisau, Switzerland).

Ma1 and Ma3 Genotyping

A cleaved amplified polymorphic sequence marker (CAPS1455) targeting base 1455 in the open reading frame of the *Ma1* gene localized to Chromosome 16 (was used to genotype trees in the population as described in (Bai et al., 2012). Briefly, polymerase chain reaction (PCR)-amplified products were digested overnight with BspHI (New England Bio Laboratories, Ipswich, MA) in a 37 °C water bath. Digested products were visualized on a 1.5% agarose gel, and *Ma1* genotypes were determined based on band patterning. The PCR program included 2 min at 98 °C, 35 cycles of 10 s at 98° C, 15 s at 55 °C, and 90 sat 72 °C, and then a final 5 min at 72 °C. The reactions were conducted in 20 µL volumes containing 1 unit PrimeSTAR MAX DNA Polymerase (R045A; Takara/Clontech, Mountain View, CA), 0.5 mM of each primer, and approximately 30 ng of genomic DNA in a gradient thermal cycler (MastercyclerVREP; Eppendorf, Hamburg, Germany). Restriction digestion was performed for at least 12 h at 37 °C in 20-µL reactions that contained 10mL PCR products, 2units of BspHI restriction enzyme (New England Biolabs Ipswich, MA), and 1NEBuffer 4 (New England Biolabs). After sample incubation, 10 µL of sample 3 µL of loading dye and 2 µL of SYBR were injected into each well of a 1.5% (w/v) agarose gel. After 1 h of electrophoresis, the banding patterns in the gel was

illuminated using the GelDoc Go Imaging System and were visually scored and the Ma1 alleles for each accession were recorded.

Trees were also genotyped for *Ma3* using the Q8 marker located on Chromosome 8 with primer sequences. The Q8 marker primer sequences were 50-AAAAATT-GAAACTTGTGGATCGTT-30 (forward primer) and 50-AAAT-CAAAAGCATACCACCACA-30 (reverse primer). The marker was PCR-amplified using 1OneTaqDNA Polymerase (NewEngland BioLabs) with the following conditions: 2 min at 98 °C, 35 cycles of 30 s at 94 °C, 30 s at 54 °C, and 45 s at 68 °C. The PCR products were visualized on 1.5% agarose gel. After 40 minutes of electrophoresis the banding patterns in the gel was illuminated using the GelDoc Go Imaging System and were visually scored, and the *Ma3* alleles for each accession were recorded (Kumar et al., 2021).

Citric Acid Phenotyping

Considering that *Malus* wild species have been found to contain higher levels of acids other than malic (primarily citric acid), an experiment was designed to quantify the possibility that the discrepancies in acid profile and expected phenotype in this population were caused by high levels of citric acid in the fruit juice. A subset of 10 individuals including the two parents of the population and progeny individuals from the highest lowest and average areas of the distribution were used to do preliminary examinations of the likelihood that citric acid was accumulating in high levels in this population, these preliminary examinations are described below.

Enzymatic Assay

In collaboration with the Cornell craft beverage analysis lab located in Geneva, NY an enzymatic kit from Biosystems was used to measure the citrate levels in this subset of this population. This kit has a lower limit of 11 mg/L and an upper limit of 400 mg/L. Since the concentration of citric acid in these samples was unknown, full-strength tests and dilution tests were done to confirm presence or concentration of citric acid. As this was an exploratory experiment, results were not analyzed or reported.

HPLC

To measure the concentration and ratio of malic more accurately compared with citric acid, high performance liquid chromatography was attempted using the same subset of samples. HPLC was attempted using slightly modified protocols from (Flores et al., 2012; Scherer et al., 2012). As this was an exploratory experiment, results were not analyzed or reported.

GCMS

Gas chromatography mass spectrometry has the capability of measuring small concentrations of polar and potentially volatile substances. GCMS was undergone in collaborating lab of Dr. Lailang Cheng using derivatization and measurement protocol described by Wang et al., (2010). These measurements found no citric acid peak located in any of the samples, however, did find large variation in malic acid levels. As this was a preliminary examination for citric acid concentration the results were not analyzed or reported here.

Statistical Analysis

Statistical analysis was done using JMP Pro 16 to do Analysis of Variance (ANOVA), and t-tests to test for difference between genotypes and genotypic classes.

RESULTS

Allelic Demographics

The population was genotyped using markers for *Ma1* and *Ma3*. NYCo1, the *M. domestica* parent was found to be homozygous dominant for *Ma1* (*Ma1Ma1*) indicating a high acid phenotype and the PI589816, the *M. prunifolia* parent was found to be heterozygous at *Ma1* (*Ma1ma1*), containing one allele for high acid, and one for low. The expected ratio for the progeny based on these parental types is 1:1 homozygous dominant: heterozygous (*Ma1Ma1:Ma1ma1*). The progeny had an actual ratio of 164 homozygous dominant individuals to 259 heterozygous individuals or a ratio of 1: 1.58.

NYCo1 was found to be homozygous recessive for *Ma3* (*ma3ma3*) and PI589816 was found to be heterozygous at *Ma3* (*Ma3ma3*). The expected ratio for the progeny using these parental types would be 1:1 heterozygous: homozygous recessive (*Ma3ma3:ma3ma3*). The progeny had an actual ratio of 213 homozygous recessive individuals to 210 heterozygous individuals or a ratio of 1.01:1.

The most common genotype combinations of *Ma1* and *Ma3* was *Ma1ma-ma3ma3* (144; 33%), followed by *Ma1ma-Ma3ma3* (129; 30%), *Ma1Ma1-ma3ma3* (83; 19%), and *Ma1Ma1-Ma3ma3* (80; 18%). These combined genotypes with the highest proportion containing a single high acid allele in *Ma1* and no high acid alleles in *Ma3*, and the second highest proportion containing only a single high acid gene in

both predict that the population should have a acidity profile that is weighted more to the lower acid, however the distribution of TA and pH both show a marked shift to the high acid side of the distribution, with more individuals closer to PI589816 side than NYCo1. This would suggest that PI589816 contains some minor high acid genes that could not be captured in the association of this study, or that there are modifiers acting on the known acidity genes causing up regulation.

Fruit Maturity, Titratable Acidity, and pH

In 2019, 176 individuals fruited, in 2020, 288 individuals fruited, in 2021, 306 individuals fruited. The SPI for the 306 individuals evaluated in 2021 ranged from 6 to 8 with an average SPI of 7.85 and 86% of individuals above or equal to 7. Since sugar, starch, and acid levels have shown an interplay it is important that the harvest method was validated to ensure that fruit were not harvested at widely different maturity levels.

All three years TA and pH showed a distribution that was primarily located between the two parents of the population with a single peak, located closer to PI589816 than NYCo1.

In 2019 TA ranged from 1.54 $g*L^{-1}$ to 23.3 $g*L^{-1}$. The mean TA was 9.5 $g*L^{-1}$ with 41% of individuals falling below the mean and 59% falling above the mean. pH ranged in 2019 from 3.13 to 4.53. The mean was 3.57 with 46% of individuals falling below the mean and 54% falling above the mean (**Figure 1.3A** and **Figure 1.4A**).

Comparing means between the two possible genotypes at *Ma1* (*Ma1Ma1*, *Ma1ma1*) did not show a significant difference on TA $\alpha = .05$ (*p*-value .2696). However, the means were significantly different between the two *Ma3* genotypes

possible (*Ma3ma3*, *ma3ma3*) (*p*-value = .0220). Between *Ma1* genotypes there was a significant difference in pH in 2019 (*p*-value = .0285). There was not a significant difference in pH between the *Ma3* genotypes (*p*-value = .2642).

2020 TA ranged from .777 g*L $^{-1}$ to 23.901 g*L^{-1} with a mean of 9.09 g*L^{-1} with 38% falling below the mean and 62% falling above the mean. pH ranged in 2020 3.08 to 4.35, with a mean of 3.61 with 49% falling below the mean, and 51% falling above the mean (**Figure 1.3B** and **Figure 1.4B**). The *Ma1* genotype showed a significant effect on TA at $\alpha = .05$ (*p*-value = .0226). *Ma3* genotype also showed a significant difference on TA (*p*-value = .0091). *Ma1* genotype had a significant effect on pH (*p*-value = .0450) while *Ma3* genotype did not (*p*-value = .068). The TA values ranged from 1.13 to 24.44 with a mean of 8.29 with 47% falling below the mean and 53% falling above the mean, pH ranged in 2021 2.69 to 4.30 with a mean of 3.23 with 46% % falling above the mean and 54% falling below the mean (**Figure 1.3C** and **Figure 1.4C**).

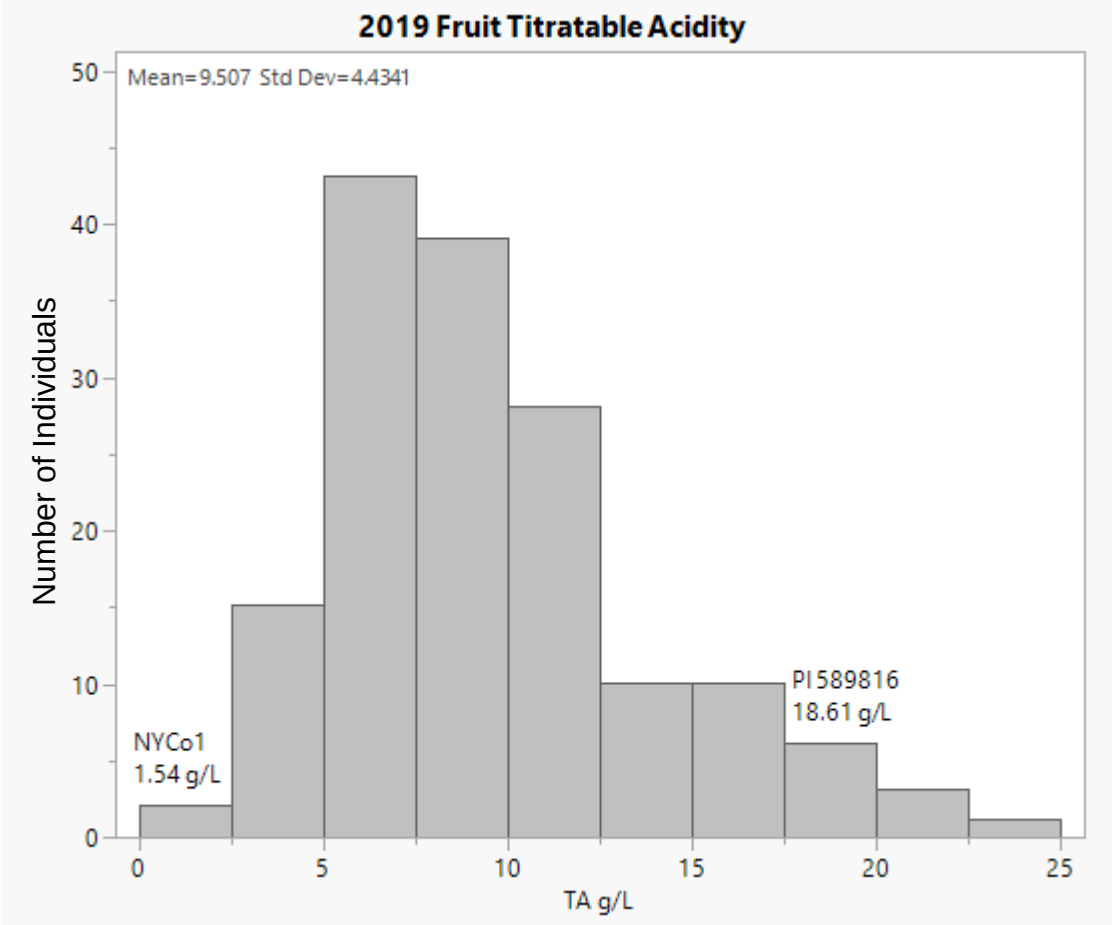
Ma1 genotype did not show a significant effect on TA or pH (*p*-values of .2121 and .1574 respectively). *Ma3* genotype did show a significant effect on TA (*p*-value = .0124) but did not show a significant effect on pH (*p*-value = .1537).

A generally accepted range for acidity of dessert type apples commonly seen on the market is .06 g*L^{-1} – 5.7 g*L^{-1}. Counts were made of the number of accessions that fall within this range each year, with 27 in 2019, 63 in 2020, and 73 in 2021. On average across the years *Ma1Ma1* genotypes had .705 g*L^{-1} higher TA compared to *Ma1ma1* genotypes, and *Ma3ma3* genotypes had 1.462 g*L^{-1} higher TA compared to ma3ma3 genotypes.

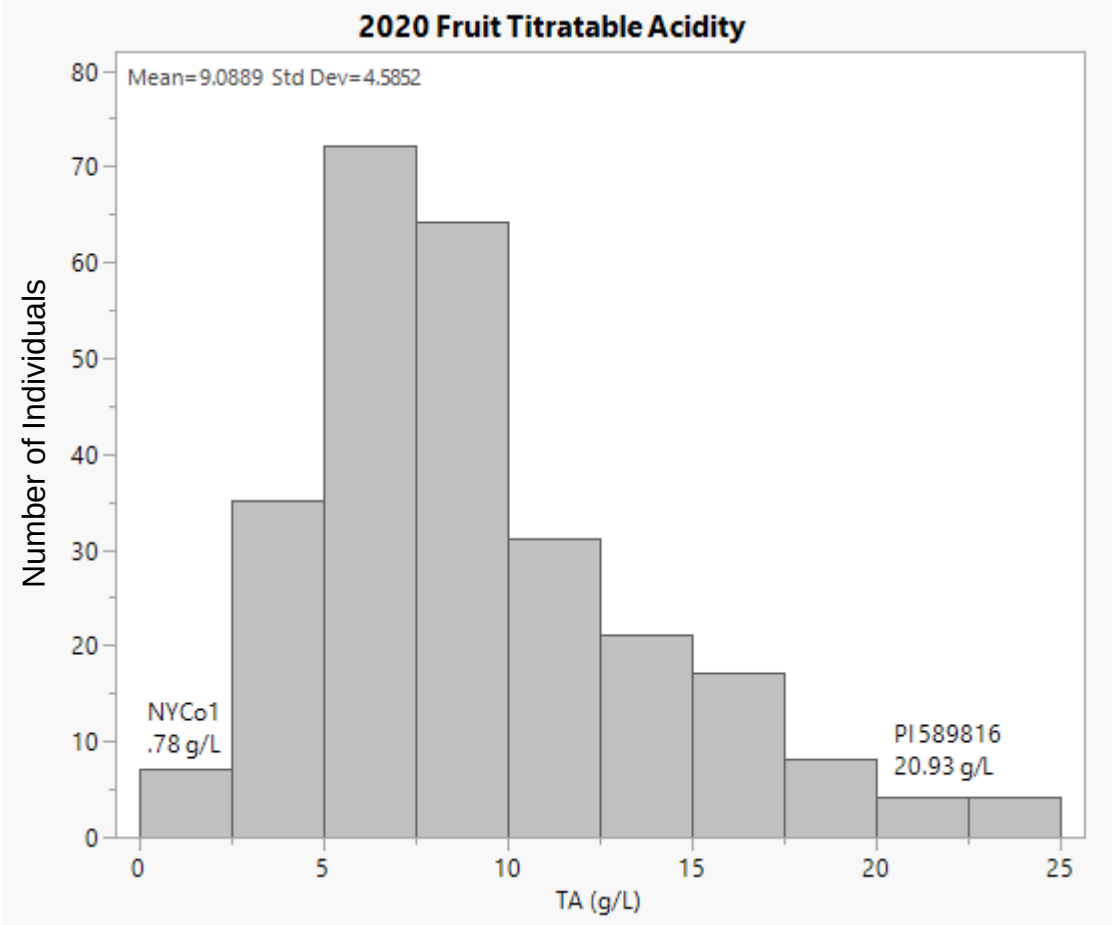
The combined possible genotypes (*Ma1ma-ma3ma3, Ma1ma-Ma3ma3, Ma1Ma1ma3ma3, Ma1Ma1-Ma3ma*) were compared using ANOVA. In 2019, the differences across genotypes were not significant at $p < 0.05$, (p-value = .0595), however it is important to note that 2019 had the lowest sample number of individuals. In contrast, genotypic means were significant at $p < 0.05$ in both 2020 (p-value = .0001) and in 2021 (p-value = .0332).

Further ANOVA analysis examining the 75 trees which fruited and were sampled in all three years revealed no significant difference for TA values, but pH was significantly different in 2019 and 2021. These two years also had the largest difference in number of trees harvested (2019; 176 and 2021; 306) Considering the stability displayed in the TA data, it can be believed that there were not large differences in juice chemistry across years.

A.


2019 Fruit Titratable Acidity
Mean=9.507 Std Dev=4.4341
Number of Individuals
50
40
30
20
10
0
0
5
10
15
20
25
TA g/L
NYCo1
1.54 g/L
PI 589816
18.61 g/L

B


2020 Fruit Titratable Acidity
Mean=9.0889 Std Dev=4.5852
Number of Individuals
80
70
60
50
40
30
20
10
0
NYCo1
.78 g/L
PI 589816
20.93 g/L
0
5
10
15
20
25
TA (g/L)

C

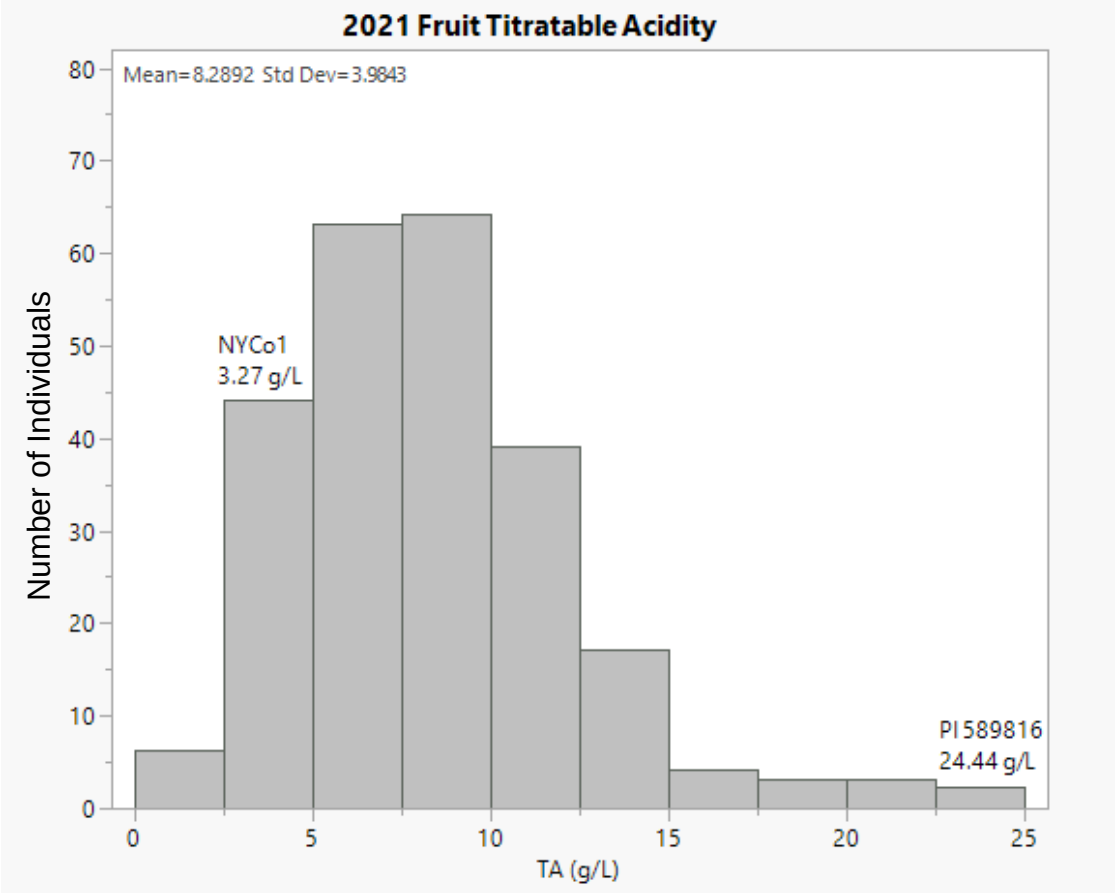
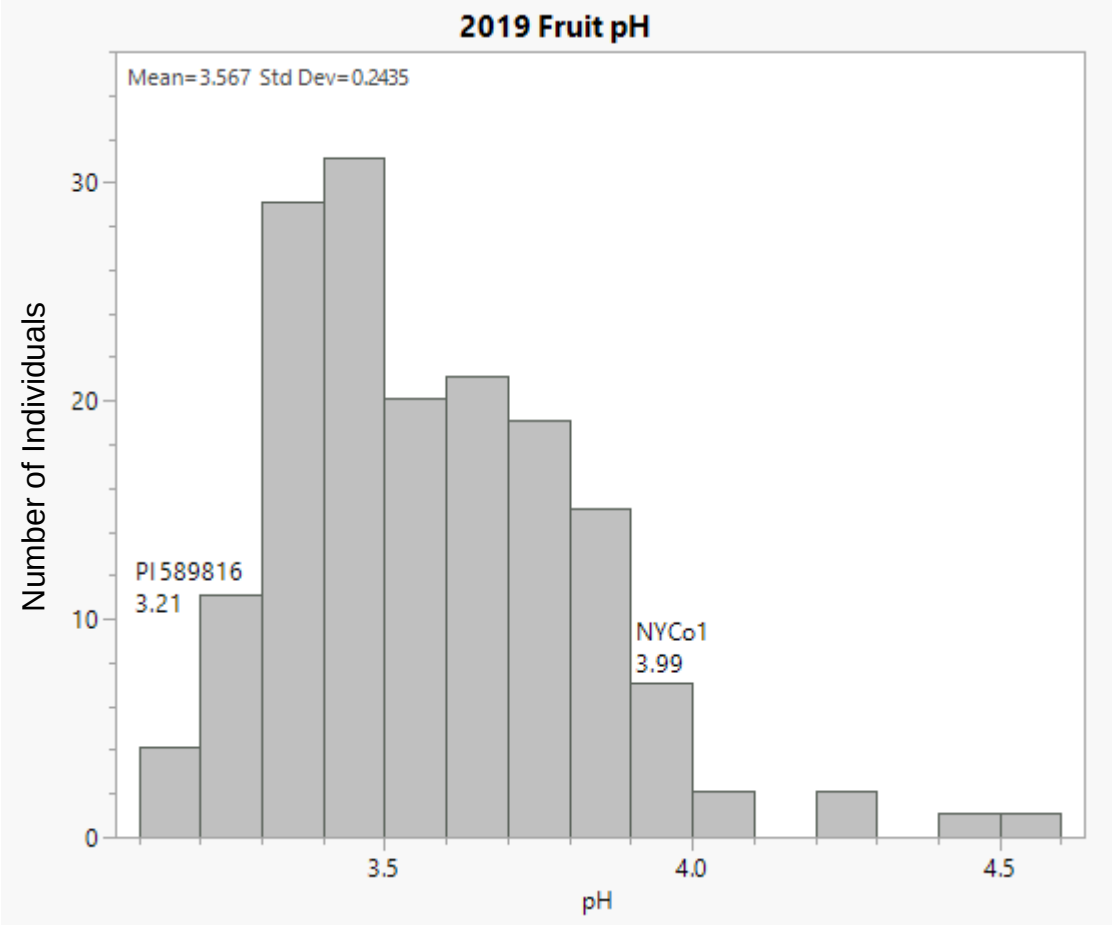
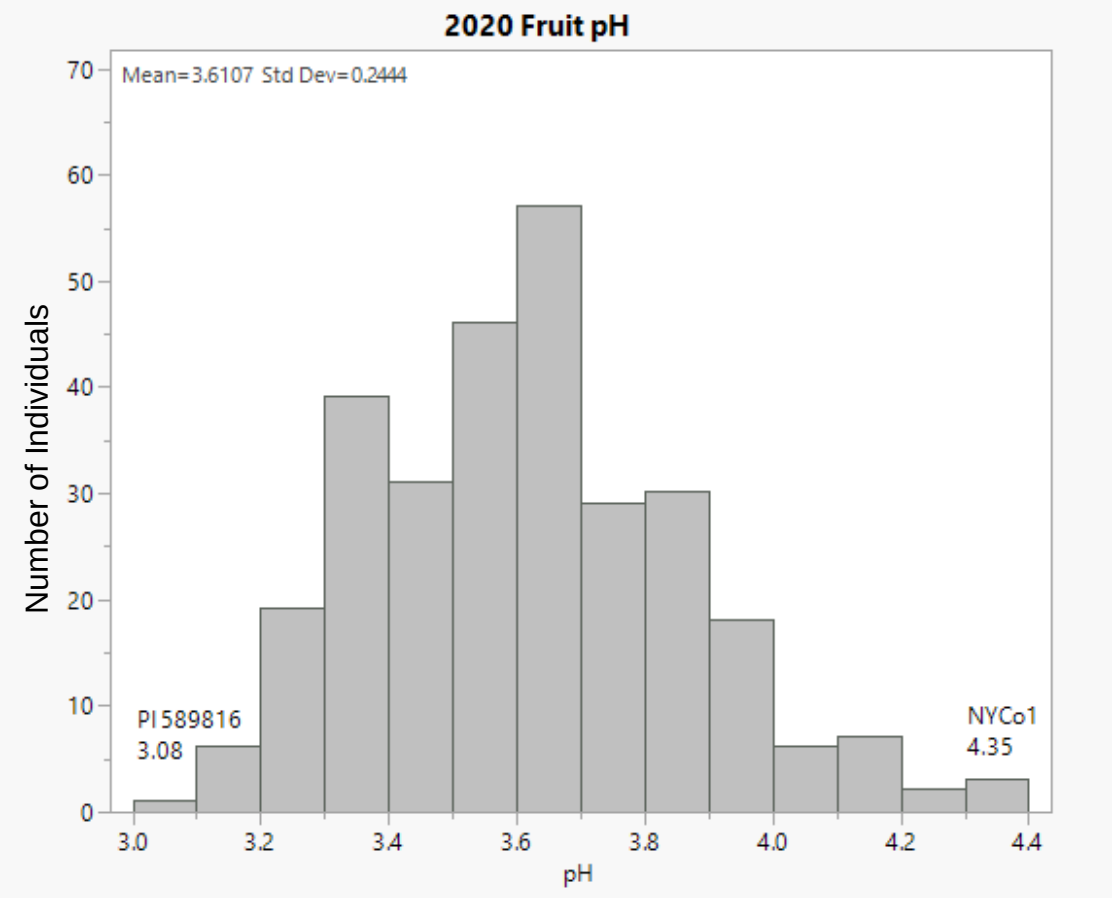


Figure 1.3 Distributions of Fruit Titratable Acidity in **A**. 2019, **B**. 2020, and **C**. 2021. NYCo1 and PI589816 TAs are marked.

A


2019 Fruit pH
Mean=3.567 Std Dev=0.2435
Number of Individuals
30
20
10
0
PI 589816
3.21
NYCo1
3.99
3.5
4.0
4.5
pH

B.


2020 Fruit pH
Mean=3.6107 Std Dev=0.2444
Number of Individuals
70
60
50
40
30
20
10
0
PI 589816
3.08
NYCo1
4.35
3.0
3.2
3.4
3.6
3.8
4.0
4.2
4.4
pH

C.

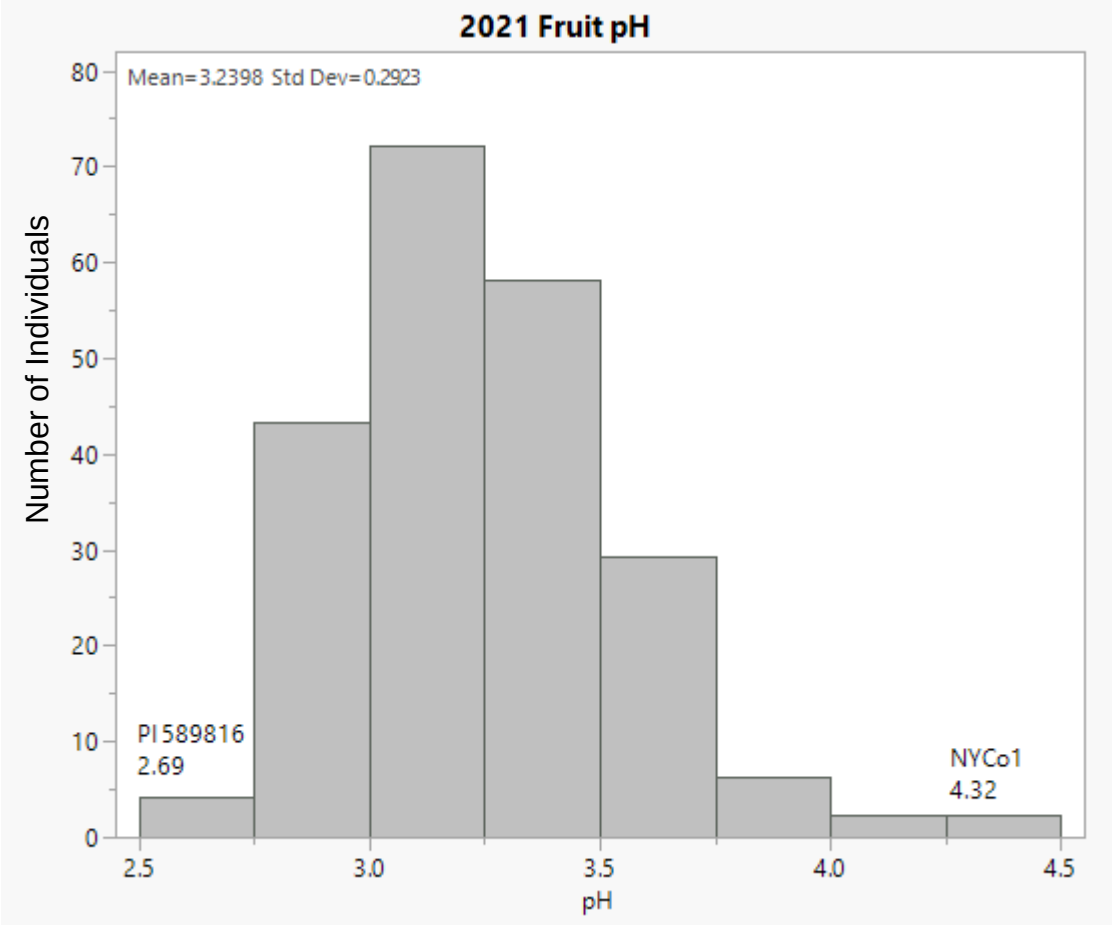


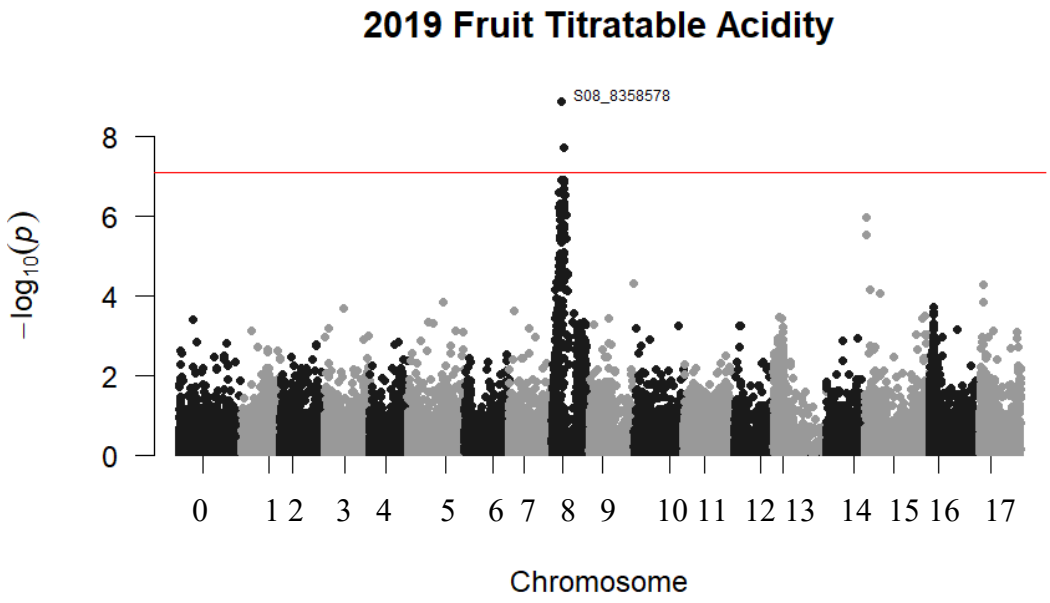
Figure 1.4 Distributions of Fruit pH in **A**. 2019, **B**. 2020, **C**. 2021. NYCo1 and PI589816 pH are marked.

Genetic Association

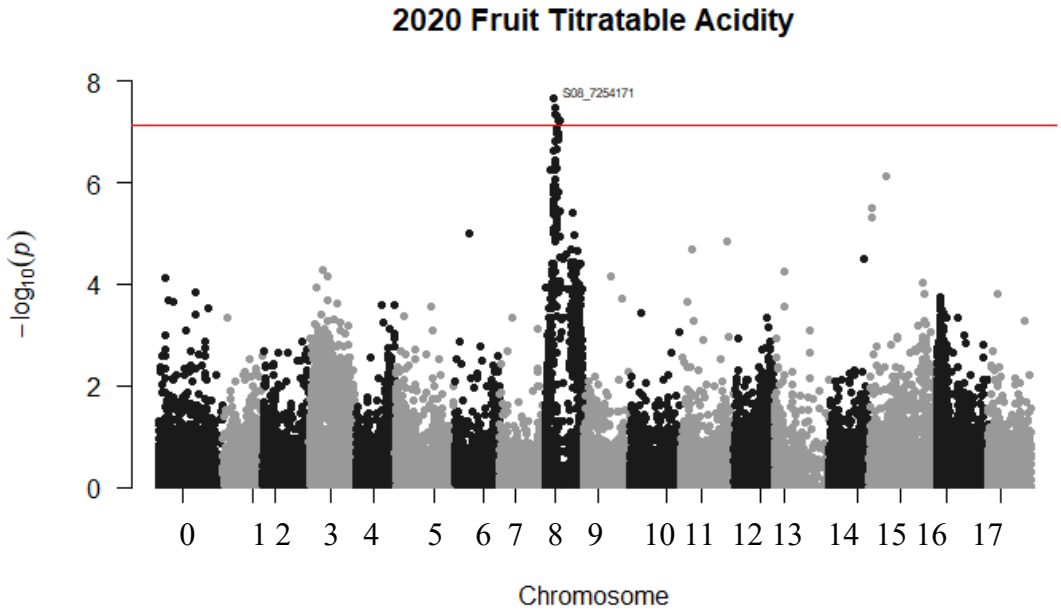
Kinship analysis and PCA identified twenty-two potentially outcrossed individuals which were removed from subsequent analyses. A single outcross was identified later during genotyping and was removed from the population.
Significant (LOD value > 7) SNP associations were identified on chromosomes 8 and 16 for pH, with peaks appearing on chromosome 8 each year (2019, 2020, 2021) however the peak on chromosome 16 only appeared associated with pH in two of the three years (2019, 2020) Manhattan plots (**Figure 1.6** and **Figure 1.7)**.
On Chromosome 8, in 2019 the SNP with strongest association was located at position 9617286, in 2020 at position 10291391, and in 2021 at position 8385007. Significant associations (LOD > 7) were also identified on chromosome 16 in 2019 and 2020 located at position 3136548 and 3132909 respectively. These genomic regions are near an Aluminum activated malate transporter family protein (MD16G1045200) with start position of 3176155 and stop position of 3178155 that has previously been identified as the causal gene behind the *Ma1* region (C. Li et al., 2020).

In all three years a SNP was identified associated with titratable acidity on chromosome 8. In 2019 the SNP was located at position 8358578, 2020 at position 10291391 and in 2021 at position 8933272. In proximity to this SNP are two candidate genes, *MYB73* and *MYB44* (MD08G1107400 with a start position of 9428392 and stop position at 9429258).

A.


2019 Fruit Titratable Acidity
S08_8358578
8
6
4
2
0
−log10(p)
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Chromosome

B.


2020 Fruit Titratable Acidity
S08_7254171
8
6
4
2
0
−log10(p)
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Chromosome

C.

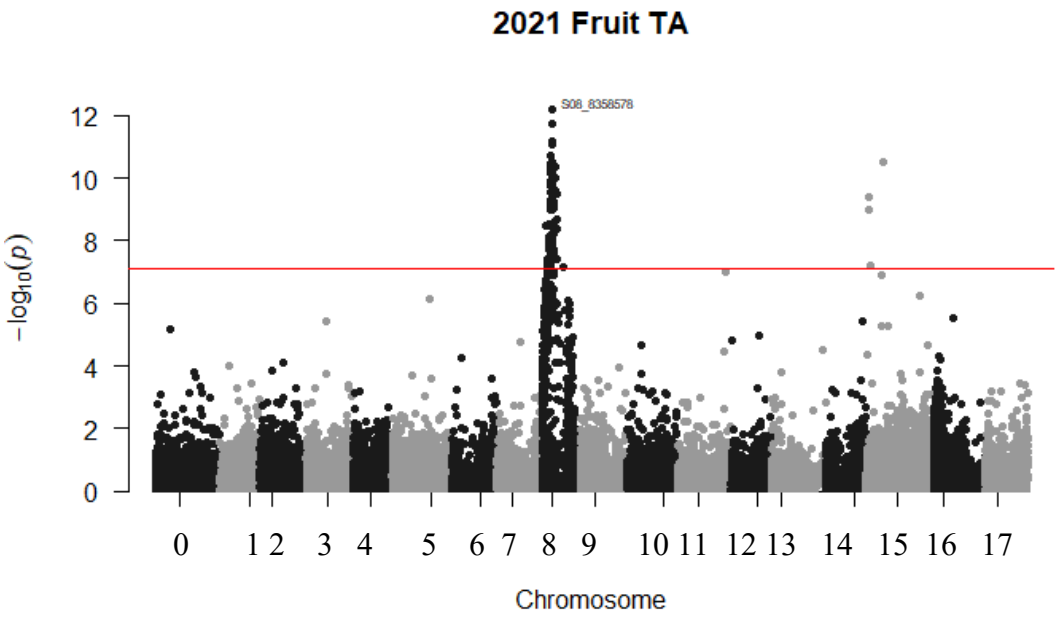
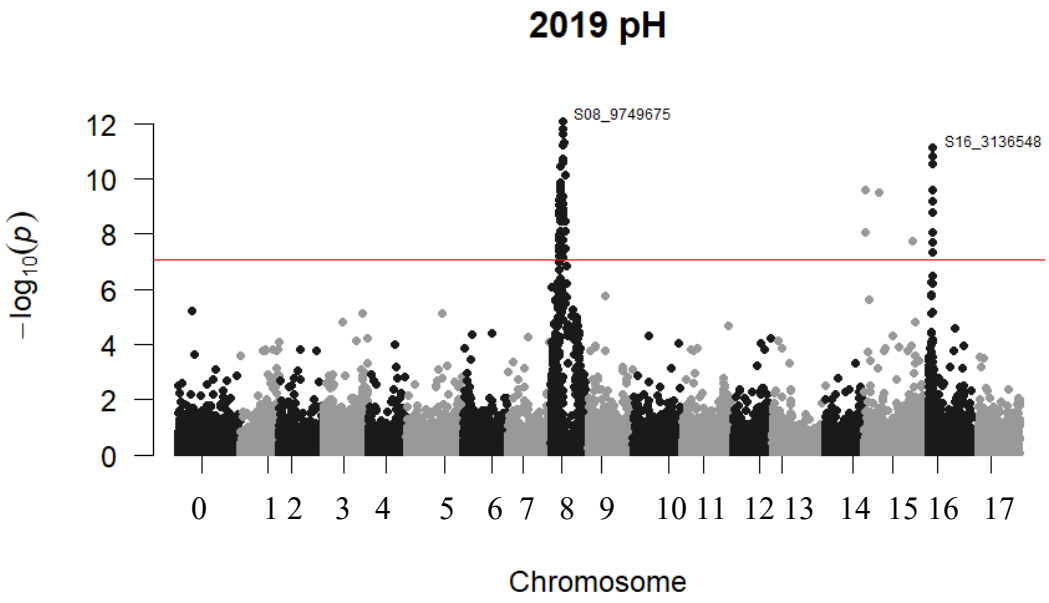
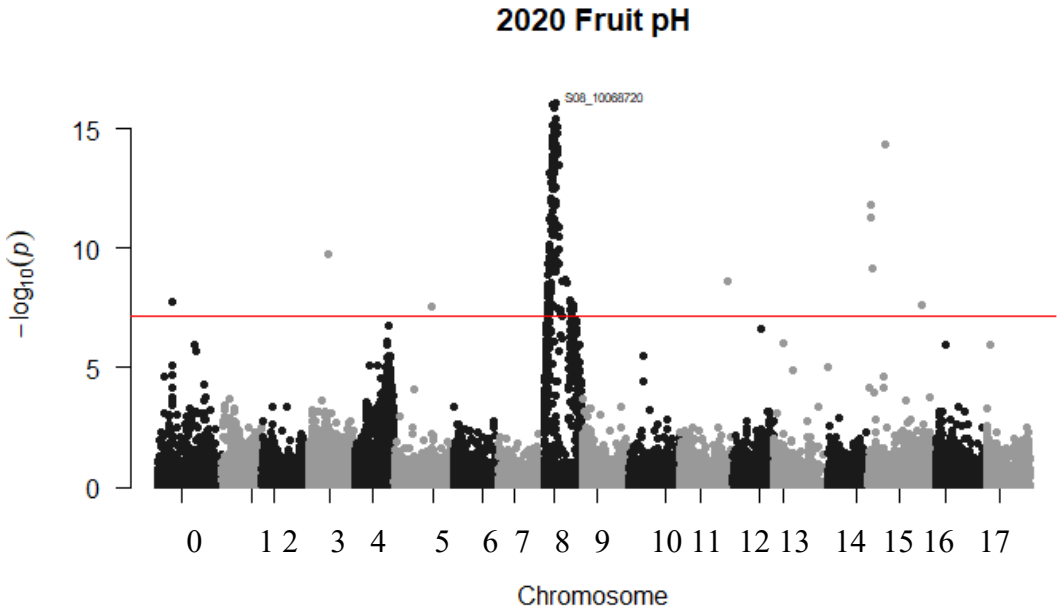


Figure 1.5 Manhattan Plots showing genetic associations determined by General Linear Model of Fruit TA in **A**. 2019, **B**. 2020, **C**. 2021. Significant LOD is marked at 7 with a red line.

A.


2019 pH
S08_9749675
S16_3136548
12
10
8
6
4
2
0
$-\log_{10}(p)$
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Chromosome

B.


2020 Fruit pH
S08_10068720
15
10
5
0
$-\log_{10}(p)$
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Chromosome

C.

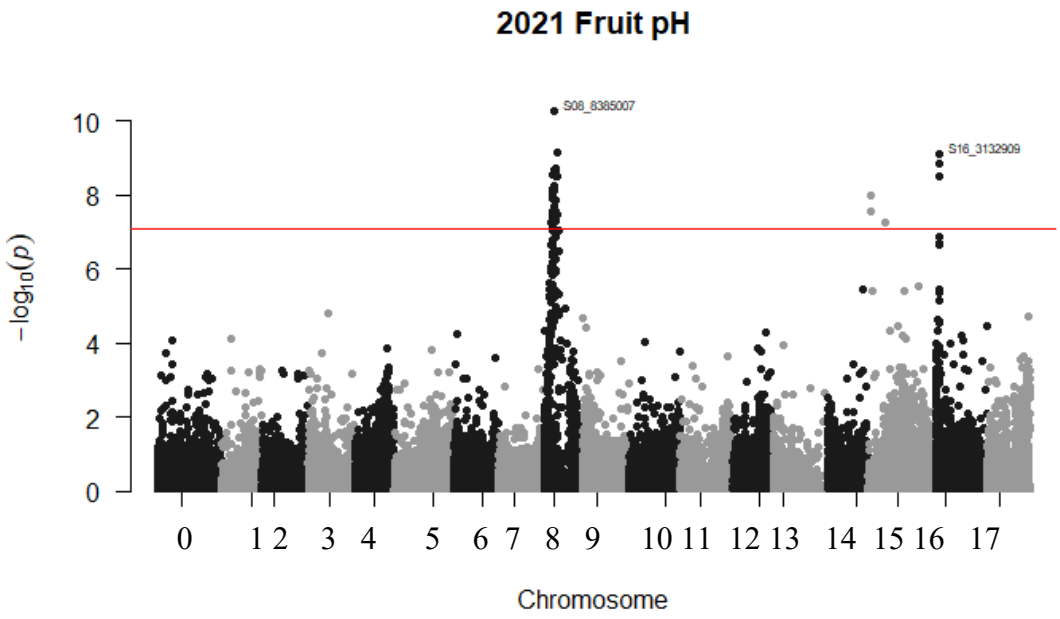


Figure 1.6 Manhattan Plots showing genetic associations determined by General Linear Model of Fruit TA in **A**. 2019, **B**. 2020, **C**. 2021. Significant LOD is marked at 7 with a red line.

Citric Acid

The analysis of citric acid content in juice samples using an enzymatic kit proved to be unreliable when measuring apple juice samples. Initial estimates of citrate in the samples ranged from 21 mg/L and 75 mg/L. However, attempts to validate measurements performed on twofold dilutions of the samples did not correlate with initial measurements at full strength indicating issues with this test. The samples were then tested using HPLC, but no identifiable peak associated with citric acid could be resolved.

HPLC

HPLC methods were evaluated further using a polar buffer of KNO3 and testing across multiple different concentrations and pressures. To verify that HPLC is an appropriate method, a "mock juice" that contained varying levels of sugars (sucrose, fructose, glucose) and acids (malic, citric, tartaric) was used. However, inconsistency in peak location and interactions between malic and citric acid were observed during these preliminary measurements and protocol establishment. Additionally, issues associated with the buffer solution precipitating while in storage and in between analysis tests could not be resolved and may have contributed to the inconsistency observed.

GCMS

The final method for citrate quantification utilized was gas chromatography-mass spectrometry (GCSMS). All samples tested in GCMS returned a peak corresponding to malic acid, with extensive variation in concentration between samples. However, there was no evidence of a peak corresponding to citric acid in any of the tested samples. The evaluated subsets of juice samples included samples collected from the two parental genotypes of this population as well as samples across the range of genotypes and phenotypes. As such, it can be assumed that citric acid is not the cause of the increased acidity in this population. The wide variation in malic acid concentration corroborates the TA measurements in the population, further documenting the differences in acidity in the population.

DISCUSSION

Apple fruit acidity is an important modern day breeding target as well as being a trait with a long history of selection during domestication. Some aspects of apple acidity are well understood, particularly the function of Ma1. However, most other loci associated with fruit acidity are less resolved. Wild species of apples often show high concentrations of acid, with the *M. prunifolia* parent in this population displaying acid levels far beyond what is seen in *M. domestica*. Wild species offer the opportunity to increase concentrations of minor acids like citric or ascorbic acid, as well as offer the chance to examine possible genes and their function that were selected against during domestication.

The impact of the Ma1 locus

The dominant allele of *Ma1* does not segregate in this population, with the progeny having possible *Ma1Ma1* or *Ma1ma1* genotypes, indicating that the progeny should all have medium to high acid phenotypes, however the population shows a wide range of acid concentrations. If *Ma1* is completely dominant, then we also would not expect to see any genetic association in its region, but we do see associations with pH in 2019 and 2021. Some of this discrepancy could be accounted for in part by the distorted segregation ratio for Ma1 genotypes observed. This can be in part attributed to NYCo1 and PI589816 being heterozygous for pale green lethal (*pgl*) which is also located on Chr 16 near *Ma1* (Orcheski et al., 2015). In the initial cross, 25% (150 seedlings) of the 600 seedlings perished at the cotyledon stage, which was the expected ratio considering the parents genotypes. Since both genes don't have large deletions causing their allelic forms, and both are caused by premature stop codons

without large frame shifts, there is not an expected significant difference in linkage between allelic types of *Ma1* and *pgl*.

If *pgl* hadn't caused lethality in this population there would have been an expected 1:1 ratio of *Ma1Ma1*: *Ma1ma1* progeny breaking down to an expected 300 individuals in each genotype class. *pgl*s lethality caused a reduction of approximately 136 homozygous dominant individuals and 41 heterozygous individuals in this population. We would expect an equal distribution of lethality if the allelic copies of Ma1 and *pgl* were similarly linked. This uneven distribution of lethality also suggests that one copy of dominant Ma1 is more closely linked to the PGL locus than the other. In this study we cannot identify the source of each allelic form of dominant *Ma1* in the progeny, so it is not feasible to definitively identify which parent donated the more closely linked allele without doing further experiments such as allele specific sequencing. It is known that genome sizes and chromosomal structure can vary between *Malus* species (Höfer & Meister, 2010; Z. Li et al., 2022). This difference in genome size and structure could account for a large proportion of lethality occurring in homozygotes (*Ma1Ma1*) in this population. Querying the *M. prunifolia* 'Fupingqiuzi' (www.rosaceae.org) reference genome for *ALMT* genes on chromosome 16 two mRNA transcripts can be found (MP16G1040400.1, MP16G1040600.1, MP16G1254500.1) with the start and end positions of 3357972-3361120 and 3369304-3372726. The *PHYLLO* gene responsible for *pgl* is located at 3860590 to 3870020 in the *M. prunifolia* reference genome. Compared with the *ALMT* transcripts from the GDDH13 reference genome located at start positions of 3166360-3168408, 3176155-3178155, and 3178157-3179520 and *PHYLLO* gene (MD16G1052600) at

3754626-3764152. These genes are in slightly different locations in their genomes, though still on the same chromosome, and that in the *M. prunifolia* reference genome *PHYLLO* and *Ma1* are closer by ~80-90 kb. If the dominant *Ma1* from *M. prunifolia* was linked to a recessive *pgl* it would result in a closer linkage between the two and could account for some of the distortion in *Ma1* mortality.

The expected phenotype of an individual that is homozygous dominant at *Ma1* would be high acid. The *M. domestica* parent, NYCo1, which is homozygous dominant for *Ma1* displays the lowest acidity profile of this entire population, contrasting with the expected phenotype. This may indicate that something is decreasing the expected phenotype produced by the *Ma1* locus (high acid), such as a modifier or non-functional gene that is outside the marker region for *Ma1*. PI589816 is heterozygous at *Ma1*, conferring a single low acid gene and a single high acid gene to the progeny. PI589816 is also the highest acid individual in the population, which also contrasts with the expected phenotype for a heterozygote.

The impact of the Ma3 locus

NYCo1 was found to be homozygous recessive at *Ma3* (*ma3ma3*) while *M. prunifolia* PI589816 was heterozygous (*Ma3ma3*). When simply contrasting the parental genotypes, it would be expected that NYCo1 would have a lower acidity than *M. prunifolia* PI589816, a result that was observed in this study. However, the observed acidity profile of PI589816 was much higher than the average *M. domestica* individual (PI589816 average TA = 21.33 g/L, average *M. domestica* TA 3–10 g/L) and was even observed to exceed acid levels of *M. domestica* individuals that contain high acid alleles at both the *Ma1* and *Ma3* loci (Kumar et al., 2021). This result

suggests that the genetic architecture of PI589816's malic acid concentration is divergent from *M. domestica* individuals potentially enhancing gene regulation or adding to the acidity profile, resulting in this incredibly high phenotype.

NYCo1's extremely low acid profile (TA < 1.0 g/L) does correlate with its Ma3 genotype of *ma3ma3*, however when compared to other apple cultivars that have been genotyped for *Ma1* and *Ma3*, NYCo1 still has lower TA and higher pH than similar genotypes of apple (Kumar et al., 2021).

The results of the genetic association analysis demonstrates that the *Ma3* region is the major driver of acidity phenotypes in this mapping population, however, the conflict between the parental genotypes and their expected and observed phenotype suggests that additional genes may be involved in the acidity of this mapping population. Despite clear associations of pH and TA in this population with both *Ma1* and especially the *Ma3* regions, it is difficult to identify genes that have a minor or additive effect on the phenotype due to limited linkage disequilibrium. There are two types of linkage disequilibrium that affects our ability to detect QTLs in this population, ancient and recent. The population was derived from an F1 cross of two highly heterozygous individuals. As a result, there has only been a single generation for recombination to occur, so highly detailed QTL analysis is a challenge (McVean, 2007)

Recently transcription factors have been a focus of identifying controls of commercially important traits (Allan & Espley, 2018). Several MYBs have been identified to affect acid concentration in *Malus* notably *MYB123* as a repressor, and *MYB44* and *MYB73* as enhancers that act on the *ALMT9* gene located at *Ma1*. *MYB44*

has been put forward as the initial start to a gene expression cascade associated with the region designated *Ma3*, so identifying this MYB transcription factor close to the SNPs associating with acidity on chromosome 8 in this population adds more evidence to its involvement in acidity control (Jia et al., 2021). A nearby transcription factor *MYB73* has also been associated with acidity in *Malus* by binding directly to the promotors of *ALMT9* including *MdVHA-A* (*vacuolar ATPase subunit A*) and *MdVHP1*(*vacuolar pyrophosphatase 1*) *(*Hu et al., 2017).

One explanation for the low acid phenotype in NYCo1 despite the homozygous dominant *Ma1* genotype is that the recessive *Ma3* genetics reduces expression of the *Ma1* gene through the gene cascade proposed by Jia et al., (2021) or another yet to be identified factor like *MYB123*, which binds to the promotor region but causes no transcription, resulting it the discrepancy between genotype and phenotype (Zheng et al., 2023a). This would result in the transcription factors located in the *Ma3* region to be responsible for the activation or increased expression of the *ALMT9* gene located at *Ma1*, attributing most of the variation in acidity to be associated with the region of chromosome 8.

M. prunifolia PI589816 is heterozygous at both loci, suggesting that it would have a medium or a moderate acidity profile, however it reliably is one of the highest or the highest TA in the population. This result may suggest that there are other factors that are conferring a higher acidity phenotype other than just the *Ma1* and *Ma3* loci. These factors might be unique to *M. prunifolia* or more present in wild species than in the breeding populations of *M. domestica.* Duan et al., (2017) identified two regions surrounding *Ma1* that had significantly reduced genetic diversity in *M.*

domestica cultivars when compared to the progenitor species *M. sieversii* suggesting that the area surrounding *Ma1* was intensely selected upon during domestication and further breeding. This greater genetic diversity near *Ma1* descending from *M. prunifolia* could also add some confounding factors to these results because of the possibility of genes that have been selected against in these higher diversity areas. Identifying minor gene acidity controls that are unique to *M. prunifolia* is outside the abilities of this association model, but these findings do suggest that minor genes are playing a role in high acid individuals. They also add more evidence to the proposed idea that the casual gene behind *Ma3*'s effect is a part of a gene cascade involving transcription factors such as *MYB44* and *MYB73*. It is possible in this population that NYCo1 contains two *myb123* transcription factors, causing the *ALMT9* gene located at *Ma1* to have reduced expression. Since NYCo1 is homozygous recessive at *Ma3* as well, those MYB transcription factors would not act on the *ALMT9* gene to restore expression, resulting in a low acid fruit, contrasting with its genotype. This would also help explain why a SNP was identified on chromosome 16 in only two of the trait × year combinations, if a transcription factor was acting on *ALMT9* at *Ma1* in this population, preventing its expression. This would explain NYCo1's low acidity as well along with the discrepancy with the expected genetic association on chromosome 16.

This work aims to further understand acidity genetics that originate from a wild species apple. Frequently in apple introgression studies it is stated that introgression results in linkage drag that confers small fruit, high acidity, and high tannins, all traits seen as negative in apple breeding. Using a large population number

and allowing most of the trees to develop fruit, this work shows that there is the possibility of developing apples that have acceptable ranges of acidity using a wild species as one of the parents. In this study progeny that fell into the acceptable acidity range were counted each year with 27 in 2019, 63 in 2020, and 73 in 2021, out of a total population of 423 individuals. There is still necessary work to be done to better understand the controls of acidity in *Malus*, and that will have to include transcription studies to identify transcription factors that act on known genes for acidity. By introgressing wild species, there is also the possibility of introducing higher concentrations of acids like citric and ascorbic acid, that would add more complexity to the acidity controls but would be unique and health benefiting traits that might be of value to consumers.

REFERENCES

Allan, A. C., & Espley, R. V. (2018). MYBs Drive Novel Consumer Traits in Fruits and Vegetables. *Trends in Plant Science*, *23*(8). https://doi.org/10.1016/j.tplants.2018.06.001

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. *Molecular Genetics and Genomics*, *287*(8), 663–678. https://doi.org/10.1007/s00438-012-0707-7

Ban, S., & Xu, K. (2020). Identification of two QTLs associated with high fruit acidity in apple using pooled genome sequencing analysis. *Horticulture Research*, *7*(171), 1–14. https://doi.org/10.1038/s41438-020-00393-y

Barbier-Brygoo, H., De Angeli, A., Filleur, S., Frachisse, J. M., Gambale, F., Thomine, S., & Wege, S. (2011). Anion channels/transporters in plants: From molecular bases to regulatory networks. *Annual Review of Plant Biology*, *62*, 25–51. https://doi.org/10.1146/annurev-arplant-042110-103741

Blanpied, G. D., & Silsby, K. J. (1992). Predicting Harvest Date Window for Apples. *CCE Publications*, 1–12.

Cummings, G. A., & Reeves, J. (1971). *Factors Influencing Chemical Characteristics of Peaches1*.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

D. Turner, S. (2018). qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *Journal of Open Source Software*, *3*(25), 731. https://doi.org/10.21105/joss.00731

Duan, N., Wang, N., Xiang, L., Mao, L., Legall, N., Mao, Z., Fazio, G., Wang, K., Sun, H., Wang, X., Wang, Y., Liu, J., Xu, Y., Jiao, C., Wang, D., Yin, C., Jiang, S., Fei, Z., Wan, S., … Chao, T. (2017). Genome re-sequencing reveals the history of apple and supports a two-stage model for fruit enlargement. *Nature Communications*, *8*(1). https://doi.org/10.1038/s41467-017-00336-7

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach

for high diversity species. *PLoS ONE*, *6*(5), 1–10. https://doi.org/10.1371/journal.pone.0019379

Etienne, A., Génard, M., Lobit, P., Mbeguié-A-Mbéguié, D., & Bugaud, C. (2013a). What controls fleshy fruit acidity? A review of malate and citrate accumulation in fruit cells. In *Journal of Experimental Botany* (Vol. 64, Issue 6, pp. 1451–1469). https://doi.org/10.1093/jxb/ert035

Etienne, A., Génard, M., Lobit, P., Mbeguié-A-Mbéguié, D., & Bugaud, C. (2013b). What controls fleshy fruit acidity? A review of malate and citrate accumulation in fruit cells. *Journal of Experimental Botany*, *64*(6), 1451–1469. https://doi.org/10.1093/jxb/ert035

Flores, P., Hellín, P., & Fenoll, J. (2012). Determination of organic acids in fruits and vegetables by liquid chromatography with tandem-mass spectrometry. *Food Chemistry*, *132*(2), 1049–1054. https://doi.org/10.1016/j.foodchem.2011.10.064

Génard, M., Bruchou, C., & Souty, M. (1991). *Variabilité de la croissance et de la qualité chez la pêche (Prunus persica L Batsch) et liaison entre croissance et qualité* (Vol. 11, Issue 10). https://hal.science/hal-00885337

Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., & Buckler, E. S. (2014). TASSEL-GBS: A High Capacity Genotyping by Sequencing Analysis Pipeline. *PLOS ONE*, *9*(2), e90346. https://doi.org/10.1371/JOURNAL.PONE.0090346

Gutierrez, B. L., Zhong, G. Y., & Brown, S. K. (2018). Genetic diversity of dihydrochalcone content in *Malus* germplasm. *Genetic Resources and Crop Evolution*, *65*(5), 1485–1502. https://doi.org/10.1007/s10722-018-0632-7

Hale, C. (1957). Synthesis of Organic Acids in the Fruit of the Grape. In *Z. vergl. Physiol* (Vol. 73, Issue 4). &I.

Höfer, M., & Meister, A. (2010). Genome Size Variation in Malus Species . *Journal of Botany*, *2010*, 1–8. https://doi.org/10.1155/2010/480873

Hu, D. G., Li, Y. Y., Zhang, Q. Y., Li, M., Sun, C. H., Yu, J. Q., & Hao, Y. J. (2017). The R2R3-MYB transcription factor MdMYB73 is involved in malate accumulation and vacuolar acidification in apple. *Plant Journal*, *91*(3), 443–454. https://doi.org/10.1111/tpj.13579

Hu, D. G., Sun, C. H., Ma, Q. J., You, C. X., Cheng, L., & Hao, Y. J. (2016). MdMYB1 regulates anthocyanin and malate accumulation by directly facilitating their transport into vacuoles in apples. *Plant Physiology*, *170*(3), 1315–1330. https://doi.org/10.1104/pp.15.01333

Jia, D., Shen, F., Wang, Y., Wu, T., Xu, X., Zhang, X., & Han, Z. (2018). Apple fruit acidity is genetically diversified by natural variations in three hierarchical epistatic genes: Mdsaur37, mdpp2ch and mdalmtii. *Plant Journal*, *95*(3), 427–443. https://doi.org/10.1111/tpj.13957

Jia, D., Wu, P., Shen, F., Li, W., Zheng, X., Wang, Y., Yuan, Y., Zhang, X., & Han, Z. (2021). Genetic variation in the promoter of an R2R3-MYB transcription factor determines fruit malate content in apple (Malus domestica Borkh). *Plant Physiology*, *186*(1), 549–568. https://doi.org/10.1093/plphys/kiab098

Kenis, K., Keulemans, J., & Davey, M. W. (2008). Identification and stability of QTLs for fruit quality traits in apple. *Tree Genetics & Genomes*, *4*(4), 647–661. https://doi.org/10.1007/s11295-008-0140-6

Khan, M. A., Olsen, K. M., Sovero, V., Kushad, M. M., & Korban, S. S. (2015). Fruit Quality Traits Have Played Critical Roles in Domestication of the Apple. *The Plant Genome*, *7*(3), 0. https://doi.org/10.3835/plantgenome2014.04.0018

Khan, S. A., Beekwilder, J., Schaart, J. G., Mumm, R., Soriano, J. M., Jacobsen, E., & Schouten, H. J. (2013). Differences in acidity of apples are probably mainly caused by a malic acid transporter gene on LG16. *Tree Genetics and Genomes*, *9*(2), 475–487. https://doi.org/10.1007/s11295-012-0571-y

Kovermann, P., Meyer, S., Hörtensteiner, S., Picco, C., Scholz-Starke, J., Ravera, S., Lee, Y., & Martinoia, E. (2007). The Arabidopsis vacuolar malate channel is a member of the ALMT family. *Plant Journal*, *52*(6), 1169–1180. https://doi.org/10.1111/j.1365-313X.2007.03367.x

Kumar, S. K., Wojtyna, N., Dougherty, L., Xu, K., & Peck, G. (2021). Classifying Cider Apple Germplasm Using Genetic Markers for Fruit Acidity. *Journal of the American Society for Horticultural Science*, 1–16. https://doi.org/10.21273/jashs05056-21

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods 2012 9:4*, *9*(4), 357–359. https://doi.org/10.1038/NMETH.1923

Li, C., Dougherty, L., Coluccio, A. E., Meng, D., El-Sharkawy, I., Borejsza-Wysocka, E., Liang, D., Piñeros, M. A., Xu, K., & Cheng, L. (2020). Apple ALMT9 requires a conserved C-terminal domain for malate transport underlying fruit acidity. *Plant Physiology*, *182*(2), 992–1006. https://doi.org/10.1104/pp.19.01300

Li, M., Li, D., Feng, F., Zhang, S., Ma, F., & Cheng, L. (2016). Proteomic analysis reveals dynamic regulation of fruit development and sugar and acid accumulation in apple. *Journal of Experimental Botany*, *67*(17), 5145–5157. https://doi.org/10.1093/jxb/erw277

Li, Z., Wang, L., He, J., Li, X., Hou, N., Guo, J., Niu, C., Li, C., Liu, S., Xu, J., Xie, Y., Zhang, D., Shen, X., Lu, L., Geng, D., Chen, P., Jiang, L., Wang, L., Li, H., … Guan, Q. (2022). Chromosome-scale reference genome provides insights into the genetic origin and grafting-mediated stress tolerance of Malus prunifolia. *Plant Biotechnology Journal*, *20*(6), 1015–1017. https://doi.org/10.1111/pbi.13817

Liebhard, R., Kellerhals, M., Pfammatter, W., Jertmini, M., & Gessler, C. (2003). Mapping quantitative physiological traits in apple (*Malus x domestica* Borkh.). *Plant Molecular Biology*, *52*(3), 511–526. https://doi.org/10.1023/A:1024886500979

Linlin, X., Xin, Q., Mingyue, Z., & Shaoling, Z. (2018). Genome-Wide analysis of aluminum-activated malate transporter family genes in six rosaceae species, and expression analysis and functional characterization on malate accumulation in Chinese white pear. *Plant Science*, *274*, 451–465. https://doi.org/10.1016/j.plantsci.2018.06.022

Lobit, P., Genard, M., Soing, P., & Habib, R. (2006). Modelling malic acid accumulation in fruits: Relationships with organic acids, potassium, and temperature. *Journal of Experimental Botany*, *57*(6), 1471–1483. https://doi.org/10.1093/jxb/erj128

Lobit, P., Soing, P., Génard, M., & Habib, R. (2002). Theoretical analysis of relationships between composition, pH, and titratable acidity of peach fruit. *Journal of Plant Nutrition*, *25*(12), 2775–2792. https://doi.org/10.1081/PLN-120015538

Ma, B., Chen, J., Zheng, H., Fang, T., Ogutu, C., Li, S., Han, Y., & Wu, B. (2015). Comparative assessment of sugar and malic acid composition in cultivated and wild apples. *Food Chemistry*, *172*, 86–91. https://doi.org/10.1016/J.FOODCHEM.2014.09.032

Ma, B., Ding, Y., Li, C., Li, M., Ma, F., & Yuan, Y. (2019). Comparative proteomic analysis reveals key proteins linked to the accumulation of soluble sugars and organic acids in the mature fruits of the wild Malus species. *Plants*, *8*(11), 8–10. https://doi.org/10.3390/plants8110488

Ma, B., Liao, L., Fang, T., Peng, Q., Ogutu, C., Zhou, H., Ma, F., & Han, Y. (2019). A Ma10 gene encoding P-type ATPase is involved in fruit organic acid accumulation in apple. *Plant Biotechnology Journal*, *17*(3), 674–686. https://doi.org/10.1111/pbi.13007

Ma, B., Liao, L., Zheng, H., Chen, J., Wu, B., Ogutu, C., Li, S., Korban, S. S., & Han, Y. (2015). Genes encoding aluminum-activated malate transporter II and their association with fruit acidity in apple. *Plant Genome*, *8*(3), 1–14. https://doi.org/10.3835/plantgenome2015.03.0016

Ma, B., Yuan, Y., Gao, M., Li, C., Ogutu, C., Li, M., & Ma, F. (2018). Determination of predominant organic acid components in malus species: Correlation with apple domestication. *Metabolites*, *8*(4). https://doi.org/10.3390/metabo8040074

Maliepaard, C., Alston, F. H., van Arkel, G., Brown, L. M., Chevreau, E., Dunemann, F., Evans, K. M., Gardiner, S., Guilford, P., van Heusden, A. W., Janse, J., Laurens, F., Lynn, J. R., Manganaris, A. G., den Nijs, A. P. M., Periam, N., Rikkerink, E., Roche, P., Ryder, C., … King, G. J. (1998). Aligning male and female linkage maps of apple (*Malus pumila* Mill.) using multi-allelic markers. *Theoretical and Applied Genetics*, *97*(1–2), 60–73. https://doi.org/10.1007/s001220050867

McVean, G. (2007). The structure of linkage disequilibrium around a selective sweep. *Genetics*, *175*(3), 1395–1406. https://doi.org/10.1534/genetics.106.062828

Migicovsky, Z., Gardner, K. M., Money, D., Sawler, J., Bloom, J. S., Moffett, P., Chao, C. T., Schwaninger, H., Fazio, G., Zhong, G.-Y., & Myles, S. (2016). Genome to Phenome Mapping in Apple Using Historical Data. *The Plant Genome*, *9*(2), plantgenome2015.11.0113. https://doi.org/10.3835/PLANTGENOME2015.11.0113

Nielsen, S. S. (2017). *Food Science Text Series Food Analysis Laboratory Manual*. Springer. www.springer.com/series/5999

Orcheski, B., Parker, R., & Brown, S. (2015). Pale green lethal disorder in apple (*Malus*) is caused by a mutation in the PHYLLO gene which is essential for phylloquinone (vitamin K 1) biosynthesis. *Tree Genetics & Genomes*, *11*(131). https://doi.org/10.1007/s11295-015-0956-9

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

Roelfsema, M. R. G., & Hedrich, R. (2005). In the light of stomatal opening: New insights into 'the Watergate'. In *New Phytologist* (Vol. 167, Issue 3, pp. 665–691). Blackwell Publishing Ltd. https://doi.org/10.1111/j.1469-8137.2005.01460.x

Rymenants, M., van de Weg, E., Auwerkerken, A., De Wit, I., Czech, A., Nijland, B., Heuven, H., De Storme, N., & Keulemans, W. (2020). Detection of QTL for apple fruit acidity and sweetness using sensorial evaluation in multiple pedigreed full-sib families. *Tree Genetics and Genomes*, *16*(5). https://doi.org/10.1007/s11295-020-01466-8

Sasaki, T., Yamamoto, Y., Ezaki, B., Katsuhara, M., Ahn, S. J., Ryan, P. R., Delhaize, E., & Matsumoto, H. (2004). A wheat gene encoding an aluminum-activated malate

transporter. *Plant Journal*, *37*(5), 645–653. https://doi.org/10.1111/j.1365-313X.2003.01991.x

Scherer, R., Rybka, A. C. P., Ballus, C. A., Meinhart, A. D., Filho, J. T., & Godoy, H. T. (2012). Validation of a HPLC method for simultaneous determination of main organic acids in fruits and juices. *Food Chemistry*, *135*(1), 150–154. https://doi.org/10.1016/j.foodchem.2012.03.111

Spengler, R. N. (2019). Origins of the apple: The role of megafaunal mutualism in the domestication of *Malus* and rosaceous trees. *Frontiers in Plant Science*, *10*(May), 1–18. https://doi.org/10.3389/fpls.2019.00617

Sweetman, C., Deluc, L. G., Cramer, G. R., Ford, C. M., & Soole, K. L. (2009). Regulation of malate metabolism in grape berry and other developing fruits. *Phytochemistry*, *70*(11–12), 1329–1344. https://doi.org/10.1016/j.phytochem.2009.08.006

Takos, A. M., Jaffé, F. W., Jacob, S. R., Bogs, J., Robinson, S. P., & Walker, A. R. (2006). Light-induced expression of a MYB gene regulates anthocyanin biosynthesis in red apples. *Plant Physiology*, *142*(3), 1216–1232. https://doi.org/10.1104/pp.106.088104

Tijero, V., Girardi, F., & Botton, A. (2021). Fruit development and primary metabolism in apple. In *Agronomy* (Vol. 11, Issue 6). MDPI AG. https://doi.org/10.3390/agronomy11061160

Velasco, R., Zharkikh, A., Affourtit, J., Dhingra, A., Cestaro, A., Kalyanaraman, A., Fontana, P., Bhatnagar, S. K., Troggio, M., Pruss, D., Salvi, S., Pindo, M., Baldi, P., Castelletti, S., Cavaiuolo, M., Coppola, G., Costa, F., Cova, V., Dal Ri, A., … Viola, R. (2010). The genome of the domesticated apple (*Malus* ×*domestica* Borkh.). *Nature Genetics*, *42*(10), 833–839. https://doi.org/10.1038/ng.654

Verma, S., Evans, K., Guan, Y., Luby, J. J., Rosyara, U. R., Howard, N. P., Bassil, N., Bink, M. C. A. M., van de Weg, W. E., & Peace, C. P. (2019). Two large-effect QTLs, Ma and Ma3, determine genetic potential for acidity in apple fruit: breeding insights from a multi-family study. *Tree Genetics and Genomes*, *15*(2). https://doi.org/10.1007/s11295-019-1324-y

Wang, H., Ma, F., & Cheng, L. (2010). Metabolism of organic acids, nitrogen and amino acids in chlorotic leaves of 'Honeycrisp' apple (Malus domestica Borkh) with excessive accumulation of carbohydrates. *Planta*, *232*(2), 511–522. https://doi.org/10.1007/s00425-010-1194-x

Xu, K., Wang, A., & Brown, S. (2011). Genetic characterization of the Ma locus with pH and titratable acidity in apple. *Molecular Breeding*, *30*(2), 899–912. https://doi.org/10.1007/s11032-011-9674-7

Yao, Y. X., Li, M., Liu, Z., You, C. X., Wang, D. M., Zhai, H., & Hao, Y. J. (2009). Molecular cloning of three malic acid related genes MdPEPC, MdVHA-A, MdcyME and their expression analysis in apple fruits. *Scientia Horticulturae*, *122*(3), 404–408. https://doi.org/10.1016/j.scienta.2009.05.033

Zhang, Y., Li, P., & Cheng, L. (2010). Developmental changes of carbohydrates, organic acids, amino acids, and phenolic compounds in 'Honeycrisp' apple flesh. *Food Chemistry*, *123*(4), 1013–1018. https://doi.org/10.1016/j.foodchem.2010.05.053

Zheng, L., Liao, L., Duan, C., Ma, W., Peng, Y., Yuan, Y., Han, Y., Ma, F., Li, M., & Ma, B. (2023a). Allelic variation of MdMYB123 controls malic acid content by regulating MdMa1 and MdMa11 expression in apple . *Plant Physiology*. https://doi.org/10.1093/plphys/kiad111

Zheng, L., Liao, L., Duan, C., Ma, W., Peng, Y., Yuan, Y., Han, Y., Ma, F., Li, M., & Ma, B. (2023b). *Allelic variation of MdMYB123 controls malic acid content by 3 regulating MdMa1 and MdMa11 expression in apple*. https://doi.org/10.1093/plphys/kiad111/7050031

CHAPTER 2

GENETIC ASSOCIATIONS OF FRUIT QUALITY TRAITS

INTRODUCTION

Fruit quality traits are major determinates in market competitiveness, including fruit size, color, flavor (sugar accumulation, acidity, phenolics), and russeting. Research into the history of apple domestication shows gene fixation surrounding genetic loci associated with fruit quality traits such as number of fruits per cluster, fruit size, fruit weight, taste/flavor, and texture (Cornille et al., 2014a). Peel color and appearance is one of the major criteria that is used by consumers to determine quality (Dar et al., 2019). To be successful at the market level new cultivars must be an improvement over what is currently on the market. Traits which increase market competitiveness include increases in disease resistance, distinctive qualities, new flavor profiles, enhanced nutrition, or superior processing ability (Brown and Maloney 2004). Since many of these traits were selected for during the process of domestication, *M. domestica* cultivars typically display lower genetic diversity surrounding selected traits e.g. acidity, fruit size, and sweetness (Duan et al., 2017). Wild species offer the opportunity to increase the genetic diversity of these traits as well as potentially novel traits, however, the use of wild species to add unique characteristics has been relatively limited (Pereira-Lorenzo et al. 2018).

Breeding approaches utilizing wild apple species offer the opportunity to increase the health benefits of apples as well as introgress novel traits. For example, in a study of 247 accessions of wild and cultivated apples examining phenolic compound diversity, wild accessions contained higher concentrations of phenolic compounds, as

well as more diversity of compounds, including the dihydrochalcone phloridzin, which has been shown to have anti-cancer and anti-diabetes properties (Farneti et al., 2015). Similarly, when comparing nine historic cultivars (‘Benito’, ‘Rossa di Corfono’, ‘Della Piastra’, Del Giappone’, ‘Lugliese Grisanti’, San Michele’, ‘Del Debbio’, ‘Del Sangue’, ‘Ruggine’) from a repository in Italy, higher levels of nutraceutical properties coming from polyphenols and metabolites were found and appear to have been selected against during domestication and breeding, resulting in a reduction in polyphenols in modern cultivars (Lo Piccolo et al., 2019)

Fruit size/shape

Fruit size is also a trait that appears to have been selected upon first by animals, as the seed dispersal agents changed from birds to large mammals, and then fixed by humans during domestication and subsequent breeding work (Duan et al., 2017; Juniper & Mabberley, 2006; Yao et al., 2015). Today, fruit shape and size are both economically important traits in apple, with consumers preferring larger apples (7.4 to 7.6 cm in diameter) over small, and rounder apples over oblong (Hampson et al., 2001; Kumar et al., 2014).

Identification of the genetic controls and development of markers for fruit size is an area needed to enable and speed marker assisted selection of apples (He et al., 2014; McCord, 2018; Platten et al., 2019). Genes associated with fruit shape and size have been identified in studies using genetic mapping and association studies (Busatto et al., 2019; Devoghalaere et al., 2012; Kenis et al., 2008; Kouassi et al., 2009; Kumar et al., 2012; Potts et al., 2014; Yao et al., 2015; C. Zhang & Hao, 2020), however mechanisms of definitive control of fruit size in apple remains elusive. A microRNA

gene (*miRNA172*) was identified which reduces fruit size in over expressing apple lines, as well as in "crab apple" type species that didn't undergo domestication and selection (Yao et al., 2015). Two fruit weight QTLs (*fw1* and *fw2)* that were identified previously were found inside regions of the genome with evidence of selective sweeps located on chromosome 8 and 15. These regions contained genes for cell division regulation and two β-galactosidase genes, candidate genes that have also been identified in strawberry and were seen to affect fruit size (Duan et al., 2017; Paniagua et al., 2016).

Sun et al., (2012) used a population of 'Jonathan' x 'Golden Delicious' (*Malus domestica* Borkh) and found a 30:1 segregation ratio of small: large fruit shape index (FSI, the ratio of height:diameter) with small fruit being the dominant phenotype, matching the expected segregation for five independently segregating genes. Using frequency distribution and molecular marker analysis, four genomic regions that were associated with FSI were identified on chromosome 10, descending from 'Golden Delicious', and LGs 11, 12, and 13 from 'Jonathan'. One of these regions on LG10 co-located with the findings of (Kenis et al. 2008) who used a population of 'Telamon' x 'Braeburn' to identify QTLs for FSI on LG5, LG10, and LG17, however these associations varied between years. Cao et al., (2015) used the same 'Jonathan' x 'Golden Delicious' population as Sun et al., (2012) with simple sequence repeat (SSR) markers to map 15 QTLs that included four for fruit length, one for diameter, and ten for FSI. One QTL for FSI on chromosome 11 co-locates with the region identified by Sun et al., (2012) and corresponded to a peptidoglycan-binding LYSM domain-containing protein gene, a gene involved in cell wall macromolecule metabolism that

has been associated with cell growth and development. Chang et al., (2014) identified 45 QTLs for apple size, length, diameter, and FSI using SSR markers and the same 'Jonathan' x 'Golden Delicious' population. In this study, QTLs for fruit length correlated with those for FSI and QTLs for length and diameter collocated to the same LG as those identified by Kenis et al., (2008); for length LG 15 and 17, and diameter on LG 2, 5, and 9.

Color

Apple skin color is largely determined by the content and concentration of primary and secondary metabolites, such as anthocyanins, carotenoids and chlorophyll. Consumers have shown a preference for red or blushed fruit (Csihon & Gonda, 2016; Steyn, 2020). Due to this strong market preference, apple breeding programs typically select fruit that presents with bright red fruit finish. Studies examining the genetic control of fruit color have identified several key genes contributing to a red fruit peel, primarily MYB transcription factors. Breeder selections targeting red colored fruit over time has thus led to a high prevalence of the *MYB1* transcription factor in popular cultivars (Migicovsky et al., 2021), a type of modern-day selective sweep within the domesticated germplasm. Since peel color is an economically important trait, research to develop high quality markers to identify red progeny in marker assisted breeding is ongoing. The primary mechanism for color determination by MYB transcription factors is through the accumulation and location of anthocyanin in apple. The transcription factor *MYB1* (*MYBA*) functions to regulate the initial steps of the anthocyanin pathway in the fruit peel and has repeatedly been mapped to chromosome 9 in apple (Chagné et al., 2007; Dar et al., 2019; Gardner et

al., 2014; Howard et al., 2019). Three *MYB1* alleles have been identified, *MYB1-1* exhibits strong expression and is in 'Fuji' while alleles *MYB1-2* and *MYB1-3* show weak expression and have been identified in 'Golden Delicious' (Kikuchi et al., 2017). Absence of the *MYB1* alleles for red skin color can result in a yellow or green skinned fruit (Cheng et al., 1996). Modulation of the *MYB1* influence has been observed based on a long terminal repeat (LTR) is inserted upstream of *MYB1*. Without this LTR activator present, *MYB1* will not be expressed (L. Zhang et al., 2019) and fruit do not present with the red skinned trait. *MYB1* can also be miss-expressed in fruit flesh and foliage, and then is called *MYB10*, resulting in pink or red flesh and red foliage (Allan & Espley, 2018).

Another important pigment that affects the color of apple fruit is due to the presence of carotenoids, which contribute to yellow, peach, and orange toned coloration (Delgado-Pelayo et al., 2014). The genetic and mechanistic control of carotenoid synthesis in apple is due to the action of *phytoene synthase* (*PSY*) genes. Twelve *PSY* genes were found on four chromosomes (3, 9, 11, and 17), and a positive correlation between chlorophyll content and carotenoid content in the fruit skin was found, demonstrating that carotenoids are synthesized in the chloroplasts along with chlorophyll (Ampomah-Dwamena et al., 2015).

Sugar

Another polygenic trait that has been selected during domestication and subsequent intentional breeding is the perception of sweetness in apple fruit, accumulation, and diversity in sugar types (fructose, glucose, sucrose and sorbitol) and total soluble solids. Soluble solid content in fruit crops is most often referred to as brix

(Jaywant et al., 2022). Despite the complex and interesting genetic control of sugar and total soluble solids in apple fruit, sweetness perception is not simply due to higher levels of sugar. In fact, sweetness perception has been found to be more closely associated with a reduction in acidity, rather than a high sugar content. Using a population of pedigreed full siblings and trained evaluators Rymenants et al., (2020) found QTLs associated with sweetness co-located at acidity loci of *Ma1* and *Ma3* (discussed in detail in Acidity). This showed that less acidic genotypes of *Ma1* and *Ma3* were perceived as sweeter than genotypes that had higher acid but similar sugar levels. This result was validated by Amyotte et al., (2017) who also used a trained panel of evaluators to test different cultivars and found that acidity and sweetness were inversely correlated, and that brix was poorly correlated with the perception of sweetness. After quantifying the concentration of individual sugars (sucrose, glucose, fructose, xylose) as well as the sugar alcohol sorbitol Aprea et al., (2017) found again that individual sugars and soluble solids were poor indicators of sweetness, sorbitol concentration however correlated best to the perception of sweetness, with several volatile compounds also impacting this perception.

Understanding the genetic control of sweetness in apple is complicated by the mix of different sugar types, including simple sugars like glucose, fructose, and sucrose, and complex sugars such as starch and sorbitol. Since specific sugars range in perception of sweetness with glucose being on the less-sweet end of the spectrum it is possible that sweeter tasting fruit were selected upon. By separating soluble solids content into individual sugars, multiple QTLs for sugar type and concentration have been identified. Using a complex mapping population of 233 seedlings derived from

crossing 8 cultivars and two advanced selections in the Washington State University apple breeding program, QTLs associating with fructose were found on chromosomes 1, 3, and 15, QTLs for glucose were found on chromosome 1, 2, 3, 15, 16, sucrose QTLs were detected on chromosome 1, 3, 4, 9, and 12, and sorbitol accumulation QTLs were found on chromosome 1, 3, 5, 9, 11, 13, and 15. The QTLs identified on LG1 for each sugar overlapped and remained consistent for sugar level through storage (Guan et al., 2015). The same region of the genome on LG 1 was mapped by Larsen et al., (2019) as a ~620 kb long region with associations for total sugar accumulation, fructose, and sucrose. Using progeny of a mapping population derived from 'Discovery' x 'Fiesta', five genomic regions were identified for brix on LG 2, 6, 8, 9, and 14 (Liebhard et al., 2003). Like many other traits linked to consumer preference, individual sugar concentration appears to have been selected upon during domestication, especially as it relates to the perception of sweetness. In a study of 364 wild and cultivated accessions evaluated for individual sugar and malic acid concentration; the average concentration of sugars between wild and cultivated accessions were similar. In fruit collected from wild accessions, fructose and glucose were the major types of sugar, in cultivated fruit fructose and sucrose were instead the most prevalent (Ma et al., 2015).

Total phenolics

Phenolic compounds are a diverse class of compounds produced by plants that have been seen to influence disease resistance, herbivory defense, and preservation (Mansoor et al., 2020; Mikulič Petkovšek et al., 2009; Pascoal et al., 2015; J. Sun et al., 2017). In *Malus* the most prevalent phenolic compounds are proanthocyanins, also

called condensed tannins, and contribute a bitter taste (Wojdyło et al., 2008). Phenolic compounds are ubiquitous across plant tissues, but in apple fruit the majority are found in the peel of the apple, with 2-4 times greater concentration in the peel compared to the flesh of the fruit (Tsao et al., 2003). Lower phenolic concentration has been a trait that apple breeders have selected for in the last 200 years of cultivar development, resulting in apples that are less bitter and astringent than traditional cultivars (Watts et al., 2021).

Genetic studies into the control of phenolic concentration are of increasing interest in modern apple breeding (Farneti et al., 2015), though from two contrasting goals. First, there is a desire to select for lower phenolic apples to result in fruit with reduced bitterness. However, phenolics have also been shown to convey health benefits on human health (Farneti et al., 2015). A major QTL for phenolic content was found on LG 16 (Chagné et al., 2012; S. A. Khan et al., 2012; Verdu et al., 2014) in a similar location to the *Ma1* locus, one of the primary loci controlling fruit acid levels. Interestingly, high levels of phenolic compounds were observed in fruit with both high and low acid levels, suggesting that the two loci are independent. A *leucoanthocyanidin reductase* (*MdLAR1*) gene located in this region is considered a strong candidate gene to control the concentration of phenolics in apple (Chagné et al., 2012, Khan et al., 2012). *LAR* enzymes are known to act on the flavanes leucocyanidin and cyanidin to produce catechins and epicatechins. This LAR1 gene has recently been associated with the accumulation of the specific phenolics catechin, epicatechin, and procyanidins B1, B2, and C1, indicating further evidence that this gene controls most of the phenolic concentration (McClure et al., 2019).

QTLs for other phenolic compounds have been identified as well, though these do not appear to control as much of the total phenolic makeup in *Malus* as *MdLAR1*. Chlorogenic acid is one such phenolic compound and a major QTL that contained a *hydroxy cinnamate/ quinate transferase* (*HCT/HQT*) gene associated with chlorogenic acid concentration was found near the bottom of LG 17 (Chagné et al., 2012). Dihydrochalcones are a major flavonoid subclass, that has recently been the focus of studies due to their association with human health and nutrition; phloridzin (phloretin 2'-O-glucoside) is the most common with a few species also producing phloretin derivatives, sieboldin (3-hydroxyphloretin-4′-O-glucoside), and trilobatin (phloretin-4′-O-glucoside). There have yet to be any major QTL associated with phloridzin concentration, but two QTLs that associated with sieboldin and trilobatin were identified on chromosome 7 and 8 (Gutierrez, Arro, et al., 2018)

Russet

Russeting is a developmental disorder where the fruit cuticle cracks and is replaced with a corky tissue layer. Russeting affects the phenolic composition of the apple, increasing certain phenolics such as phloridzin in the peel, russetted apples also synthesize and accumulate more suberin in the peel than un-russetted fruit (Busatto et al., 2019; Gutierrez, Zhong, et al., 2018b). The genetic architecture of russeting continues to be a focus of apple improvement. Using a population of 117 progeny derived from a cross between the russeted cultivar 'Renetta Grigia di Torriana' and the non-russeted cultivar 'Golden Delicious' a QTL that correlated with the occurrence and quantity of russeting was found on LG 12 parent 'Renetta Grigia di Torriana' with a putative plasma *membrane-localized ATP-binding cassette half-transporter*

(*ABCG11*) that is active in cuticular defects in *Arabidopsis* (Falginella et al., 2015). Other studies using a population of four breeding families (NZSelectionT153, NZSelectionT179, 'Sciros' and 'Fuji' and two red-fleshed pollen parents NZSelectionT31 and NZSelectionT51) identified a major QTL on LG 1 with minor effect QTLs on LG 9 and 16, suggesting that russet is polygenically controlled, despite identification of QTLs in this study, no mechanistic candidate genes were identified (Kumar et al., 2012). Like other traits MYB transcription factors have also been implicated with russeting; A MYB transcription factor, *MYB93*, has been identified as a regulator of russeting and suberin deposition in apple using RNAseq and genetic over expression in *Nicotiana benthamiana* (Legay et al., 2016).

Objectives

The objectives of this study were to identify loci that may be unique to wild species apples using a segregating population with the parents *Malus domestica* (hereafter NY Co1) in the Cornell Apple breeding program and *Malus prunifolia* PI 589816 (hereafter 589816). This study presents one of the few studies of fruit quality traits, and the only study the author is aware of explores fruit quality traits using *M. prunifolia* as a parent of the population.

MATERIALS AND METHODS

Plant Material

This research was completed using a bi-parental, full sibling population of *M. domestica* (NYCo1) x *M. prunifolia* (PI 589816) grown on their own roots at the Lucy Research Farm located at Cornell Agritech in Geneva, NY (42.868536, -77.044668).

The cross was made in 2013 with 600 seeds collected, stratified, and planted once the radical emerged. 150 seedlings were identified following germination carrying the trait of pale green lethal, which is expected in material heterozygous for the disorder, resulting in a population of 445 individuals (Orcheski et al., 2015). Surviving seedlings were grown in the greenhouse for 2 years before being planted outside. NYCo1 is heterozygous for *Co*, the columnar gene, and produces large (>80 mm) yellow fruit with a pink blush where exposed to the sun with minimal russet outside of the stem well. *M. prunifolia* PI 589516 has standard architecture and produced small (<15 mm) yellow fruit, primarily with a waxy cuticle, but occasionally displays partial to full russet.

Fruit were harvested in 2019, 2020, 2021, and 2022. Trees that appeared unhealthy or had insufficient fruit were excluded from collection. Fruit was harvested starting at the beginning of September and extending to the end of October each year. Ripeness was determined by evaluating the ease of abscission from the tree by hand as well as ground color change. At least 30 fruit were randomly harvested from each tree, harvesting from different regions of the tree canopy. The fruit was placed into paper bags to allow gas exchange and labeled with the row number and tree number; this number was then used to track the fruit through data collection. The fruit was stored at 4 C in a commercial cooler under ambient atmospheric gases for 1-3 weeks until processing. Processing was completed within two weeks following the date of harvest. A subsample of 10 fruit that were deemed to be representative of the fruit collected were juiced using a juicer [ACME Juicerator] using a Kim-wipe as a disposable filter

to reduce fruit flesh in the juice sample. After juicing the samples were aliquoted into 50-mL tubes and stored in at -20 C until juice chemistry was evaluated.

DNA Extraction

DNA was extracted from this population in a collaboration between the USDA-ARS PGRU and the Apple Diversity Group of Dalhousie University (Migicovsky et al., 2016). Briefly, single nucleotide polymorphisms (SNP) were identified using genotyping-by-sequencing (GBS). Libraries were constructed using *Ape*KI restriction enzyme and Illumina HiSeq 2000 (Illumina Inc., San Diego CA, USA) sequences were completed at the Genomic Diversity Facility, Cornell University, Ithaca NY, USA (Gutierrez, Zhong, et al., 2018a).

Individuals in the population were scored using genotyping-by-sequencing (Elshire et al., 2011) using the methods described by (Gutierrez, Zhong, et al., 2018a). Sequence tags were aligned to the *M. domestica* GDDH13 Whole Genome v1.1 (https://www.rosaceae.org/) using Bowtie 2 (Langmead & Salzberg, 2012) with parameters D, R, N, L and i set to 30, 5, 1, 15, and S, 1, 0.25 to reduce misalignment. The Tassel 5 (Glaubitz et al., 2014) GBS pipeline was used for SNP calling. Ten replicates of parents NYCo1 and PI 589816 were merged into single genotypes. VCFtools (Danecek et al., 2011) was used to filter data; sites were filtered for read depth of 8 and minor allele frequency of 0.20; and genotypes were filtered for 80% missing data. Genotypes were filtered based on Mendelian error calculated in PLINK 1.9 (Purcell et al., 2007) and set to missing. SNP markers were filtered based on chi-square tests for segregation distortion and > 0.95 similarity using JoinMap4.1.

Genetic Association

Marker-trait association analyses between 51442 Genotype-by-Sequencing markers and measured traits were performed using TASSEL v5 software (Bradbury et al. 2007). Using the Q method, a general lineal model (GLM) was done using numeric data joined to genotype data and principal component analysis (PCA) as population membership estimates, along with Ma1 and Ma3 genotypes, and tree architecture as covariates (*Ma1Ma1* = 1, *Ma1ma1* =0; *Ma3ma3* = 1, *ma3ma3* = 0; *Coco* = 1, *coco* =0). Data was exported from TASSEL and results were visualized using the 'qqman' R package (D. Turner, 2018). Significant SNPs were determined using the Bonferroni-corrected threshold for GWAS (Log of Odds > 7). Each year × trait was done independently of other years. Associations with diameter, height, mass, and FSI were made independently of each other, as well as compared across these fruit size related traits to identify broad fruit size associations.

Juice collection

A small number of trees fruited in 2018 but data wasn't analyzed. In 2019, 176 individuals fruited, in 2020, 288 individuals fruited, in 2021, 306 individuals fruited. Limited data was collected in 2022 with 308 fruiting trees but due to the time required only external measurements were taken. In 2021 5 fruit from each tree was verified for maturity using the starch pattern index (SPI) assay. Cortex SPI was determined by staining the stem side of a cross section taken at the equator of the apples with iodine solution (.22 w/v iodine, .88% w/v potassium iodine) and visual rating using the Cornell Starch Iodine rating where 1 indicated 100% staining and 8 indicated 0% staining, with higher levels of staining indicating more starch and conversely less ripe

(Blanpied & Silsby, 1992). Brix was measured from pure juice using a refractometer (ATAGO USA, Bellevue, WA)

Total Phenolic Measurement

Total polyphenol concentration was measured using the Folin-Ciocalteau method (Waterhouse, 2012). Standards for the total polyphenol concentration were generated using an eight-point standard curve with gallic acid from 0 – 3.0 $g \times L^{-1}$. The reaction was carried out in a Cellistar 96-well microplate (Greiner bio-one, Monroe, NC). The reaction mixture consists of 1.5 µL of standard or sample, 34.9 µL of water, and 90.9 19 µL of Folin-Ciocalteu reagent (Sigma-Aldrich, St. Louis, MO). Three minutes after the addition of the FC reagent, 72.7 µL of 7% (v/w) sodium carbonate (Sigma Aldritch, St Louis, MO) solution was added. The reaction mixture was incubated for 1 h at room temperature in the dark. Samples were measured at 765 nm and total polyphenol content was determined by linear regression from the standard curve plot and multiplying by the dilution factor (for accessions which needed to be diluted to fit within the linear range). Tannic acid standards (Tannic Acid Powder, ACS-Alfa Aesar, Ward Hill, MA) were then analyzed via the Folin-Ciocalteu assay. A conversion factor between the Löwenthal Permanganate Titration method (0.005 M solution titrated against 1 mL of juice sample) and the Folin-Ciocalteau method was developed by creating a standard curve plotting the concentration of gallic acid standards from the Folin-Ciocalteau method compared to the tannic acid equivalents values measured against the same standards via the Löwenthal Permanganate Titration method to identify the linear relationship between the assays. Ten juice samples were

measured for total polyphenol concentration via Löwenthal Permanganate Titration method and the Folin-Ciocalteau method to verify the linear relationship.

External Fruit traits

Ten fruit from each accession were randomly selected as a representative sample for external fruit traits. In 2019, 2020, 2021 and 2022, diameter and height measurements were taken. Diameter was measured at the widest point between the stem and calyx using a caliper (S_Cal EVO Smart, Sylvac SA, Switzerland) while height was taken at the widest point crossing both the stem and the calyx with the same caliper. Mass was taken using a scale (PB3002-S, Mettler-Toledo) in 2020, 2021, and 2022. In 2020 and 2021, fruit color was taken using a Nix Colorimeter. Analysis of Variance (ANOVA) using JMP Pro 16 was used to assess uniformity across years for fruit size, individuals were included if they had fruited for 3 years. A section of the peel that showed consistent color and didn't have lenticels, marks or blush coloration was measured on each of the 10 fruit and recorded as RGB values. If fruit had a secondary color (blush), similar criteria were used to determine region to measure. The RGB values were also transformed into CIE Lab* color space using R 4.2.3.

RESULTS

Genetic Filtering

Twenty-six individuals were removed for low mean depth; sites were filtered for read depth of 8 and minor allele frequency of 0.20; and genotypes were filtered for 80% missing data. Kinship analysis and PCA done in TASSEL identified twenty-two

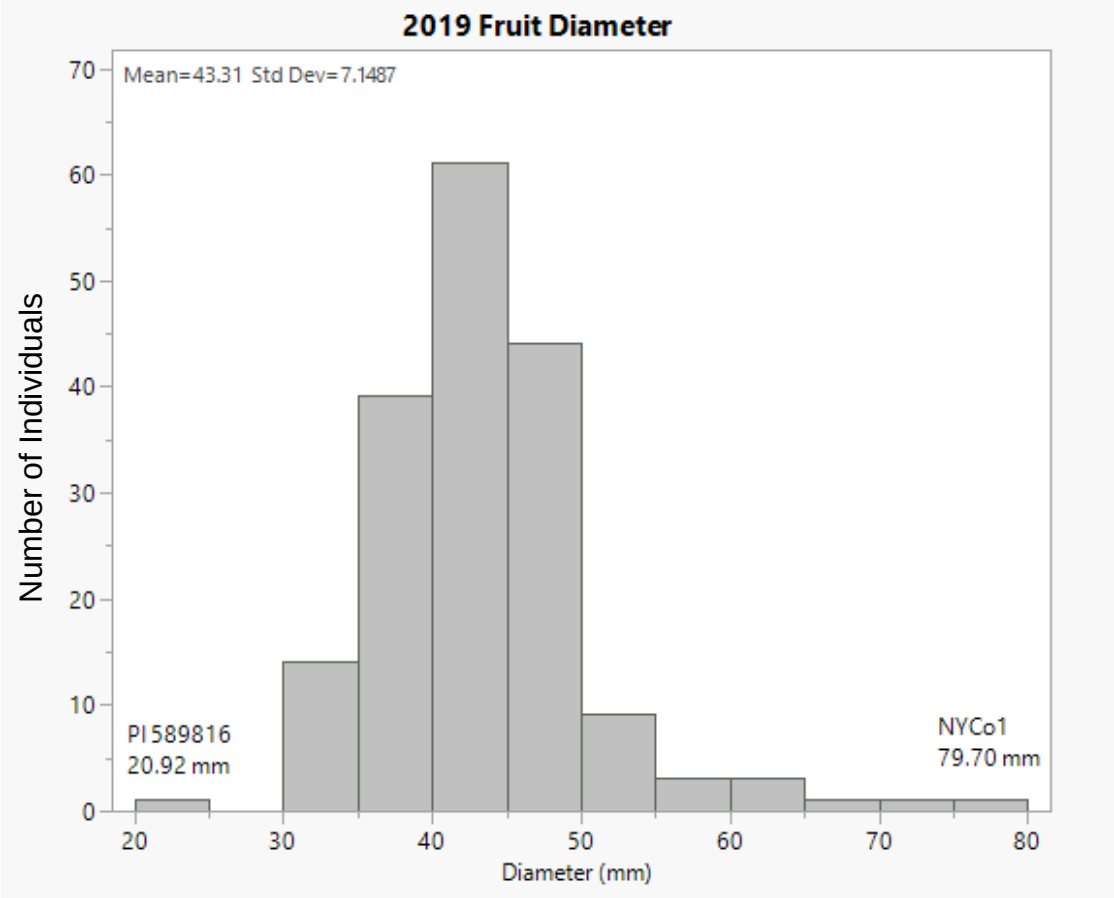
outliers. which were removed from subsequent analyses. A single outcross was identified later during genotyping and was removed from the population.

Fruit Size

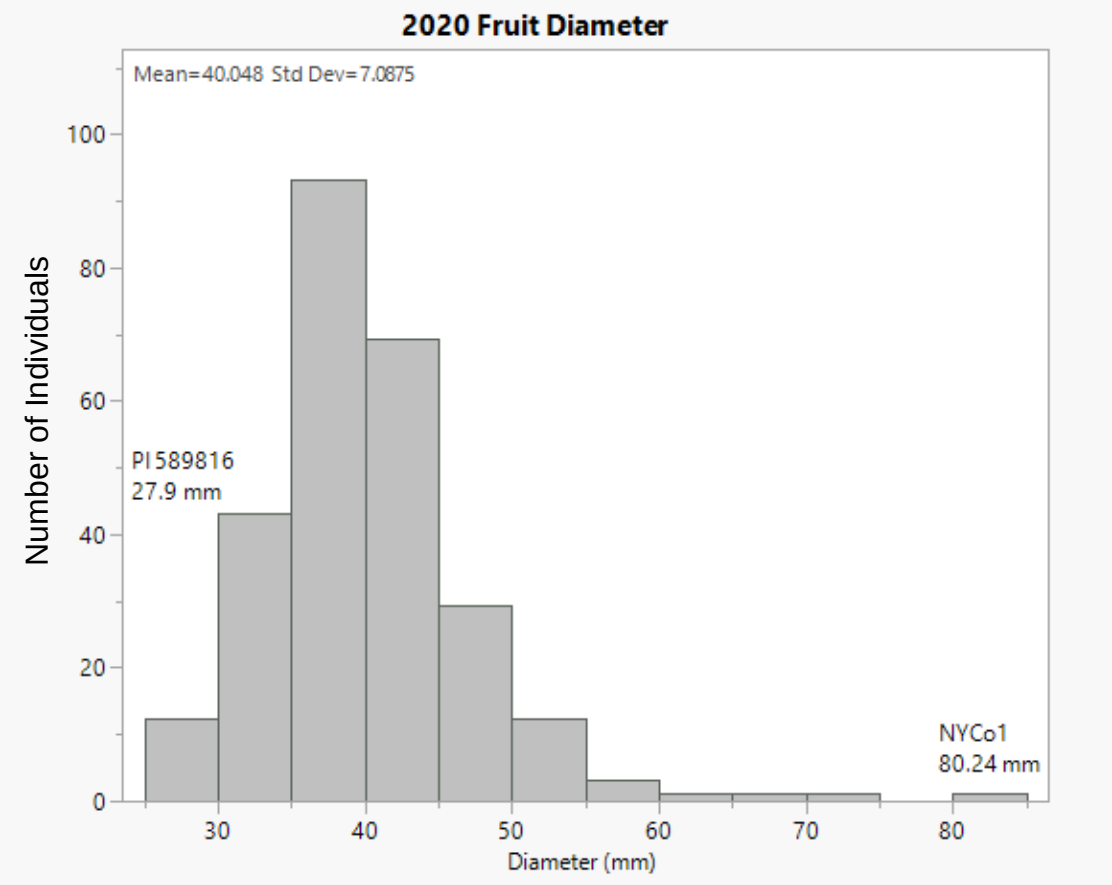
Diameter

Fruit Diameter was measured four years. In 2019, 176 individuals were measured and had a range of 20.92 mm - 79.70 mm and an average of 43.31 mm. in 2020, 265 individuals fruited with a range of 26.07 mm – 80.24 mm, and an average of 40.05 mm. in 2021, 307 individuals were measure with a range of 27.50 mm – 83.57 mm and an average of 43.88 mm. In 2022, 302 individuals were measured with a range of 23.12 mm – 82.01 mm and an average of 37.60 mm (**Figure 2.1**). Fruit diameter was normally distributed each year between the two parents. PI589816 was one of the smallest diameter fruits in the population each year (average diameter = 26.00 mm), while NYCo1 was consistently one of the largest diameter fruits (average diameter = 81.24 mm).

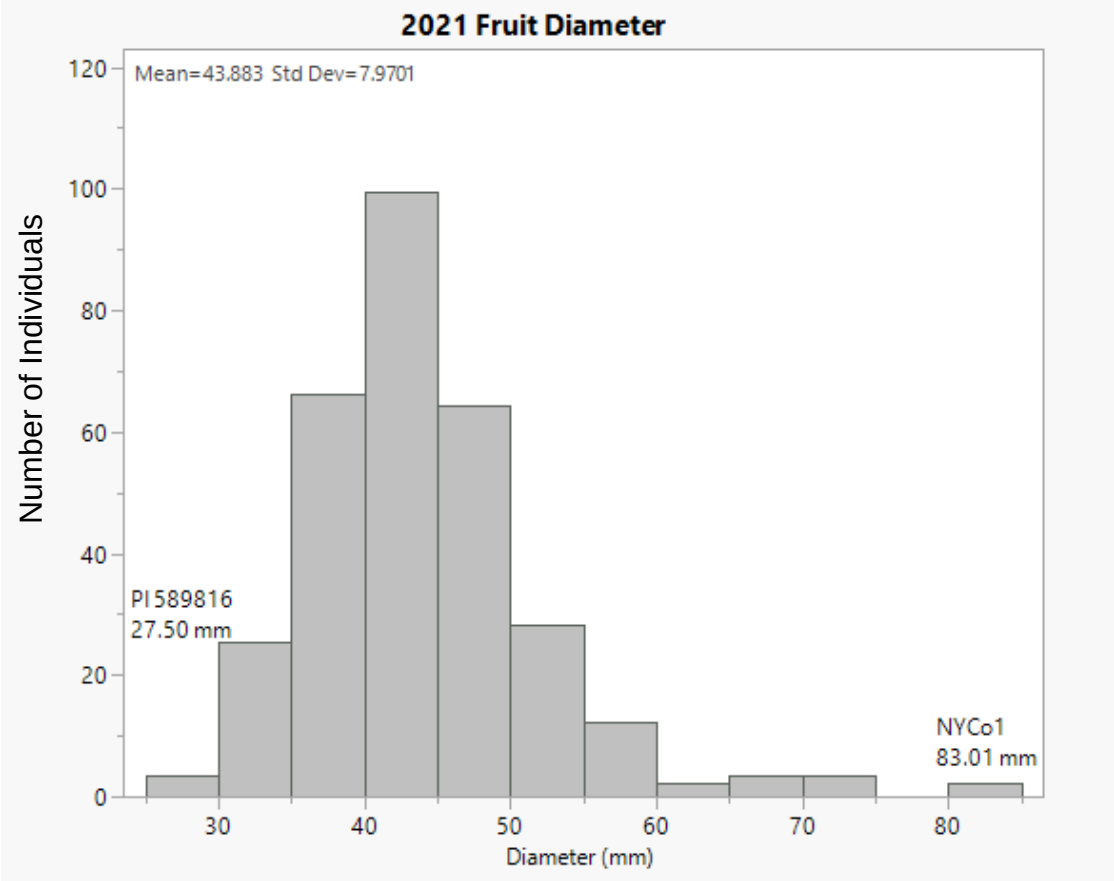
A


2019 Fruit Diameter
Mean=43.31 Std Dev=7.1487
Number of Individuals
70
60
50
40
30
20
10
0
PI 589816
20.92 mm
NYCo1
79.70 mm
20
30
40
50
60
70
80
Diameter (mm)

B


2020 Fruit Diameter
Mean=40.048 Std Dev=7.0875
Number of Individuals
100
80
60
40
20
0
PI 589816
27.9 mm
NYCo1
80.24 mm
30
40
50
60
70
80
Diameter (mm)

C


2021 Fruit Diameter
Mean=43.883 Std Dev=7.9701
Number of Individuals
120
100
80
60
40
20
0
PI 589816
27.50 mm
NYCo1
83.01 mm
30
40
50
60
70
80
Diameter (mm)

D

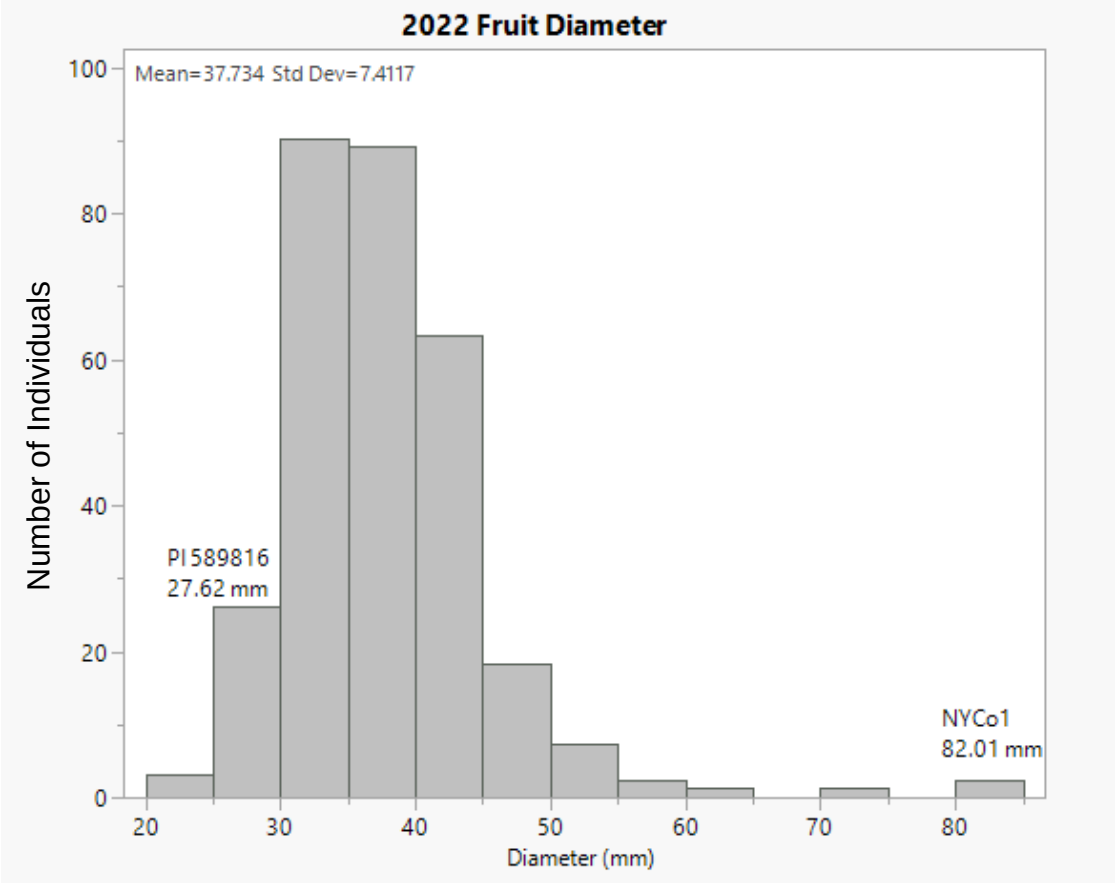
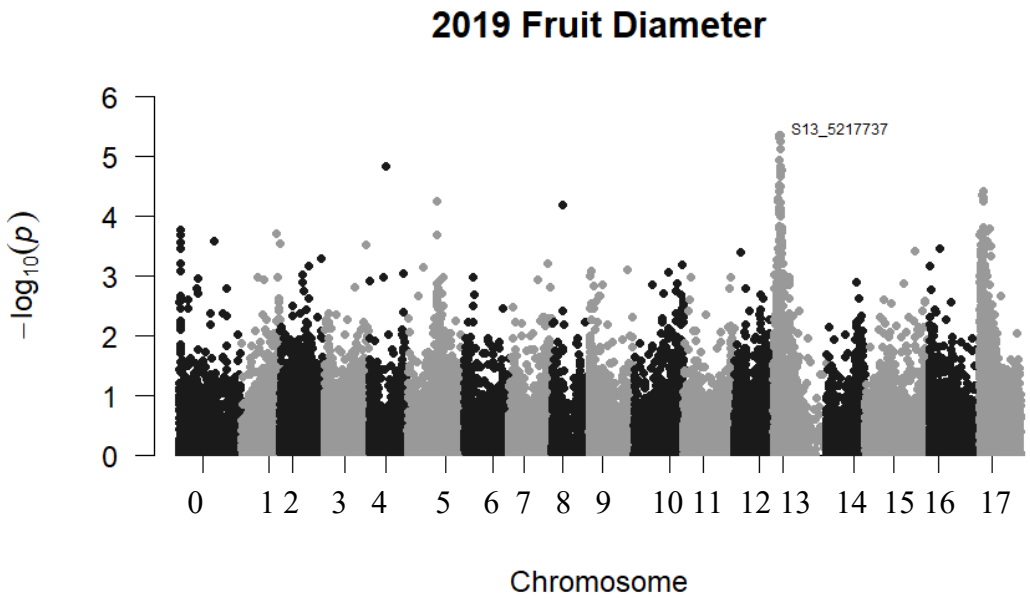


Figure 2.1 Distribution charts of fruit diameter in A. 2019, B. 2020, C. 2021, and D. 2022. Diameters of PI589815 and NYCo1 are marked in each chart.

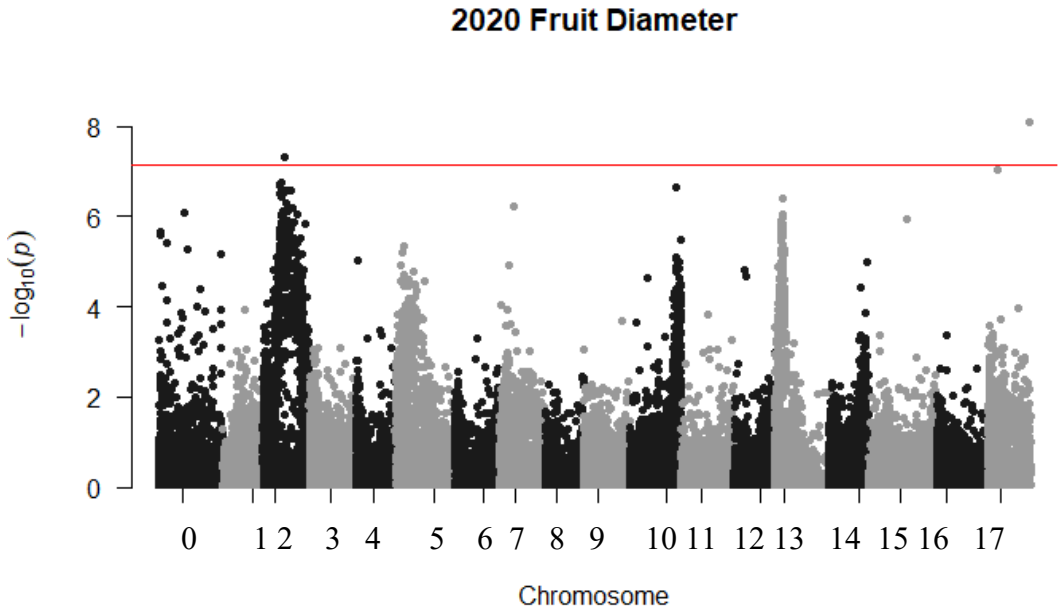
Significant SNPs were identified on Chromosome 2 in 2020, 2021, and 2022 (**Figure 2.2**) at 19848917, 30325743 and 18713115 respectively, there were no annotated genes within 20kb on either side of these SNPs located near each other in 2020 and 2022. In 2019 and 2021 a significant SNP was identified on Chromosome 13 at 5029123 and 176952 respectively. No significant SNPs were identified in 2019.

Figure 2.2A

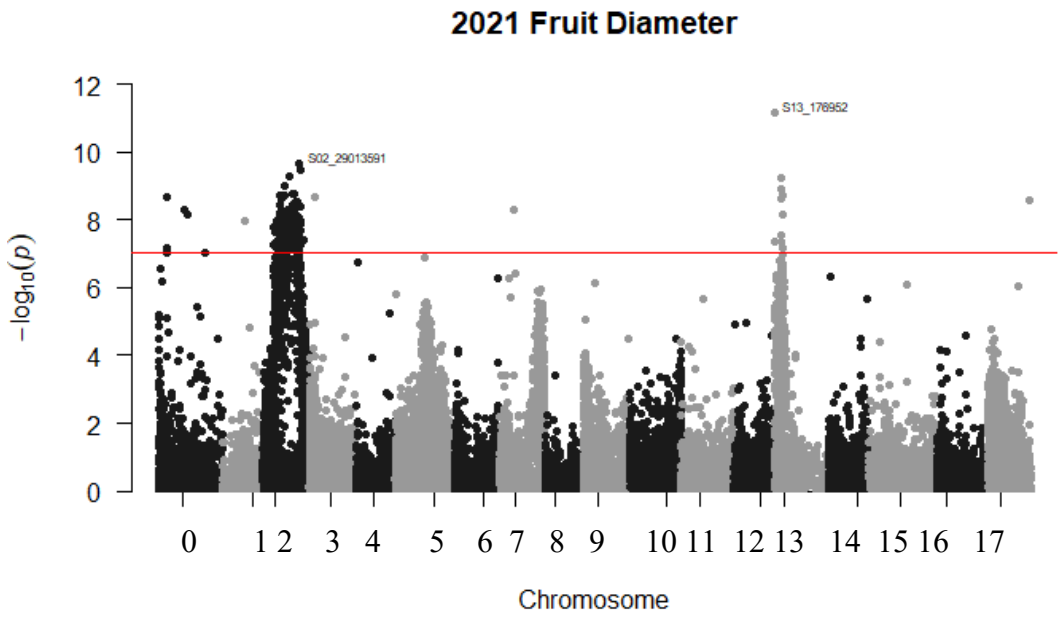
A


2019 Fruit Diameter
S13_5217737
−log10(p)
0
1
2
3
4
5
6
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Chromosome

B


2020 Fruit Diameter
−log10(p)
0
2
4
6
8
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Chromosome

C



D

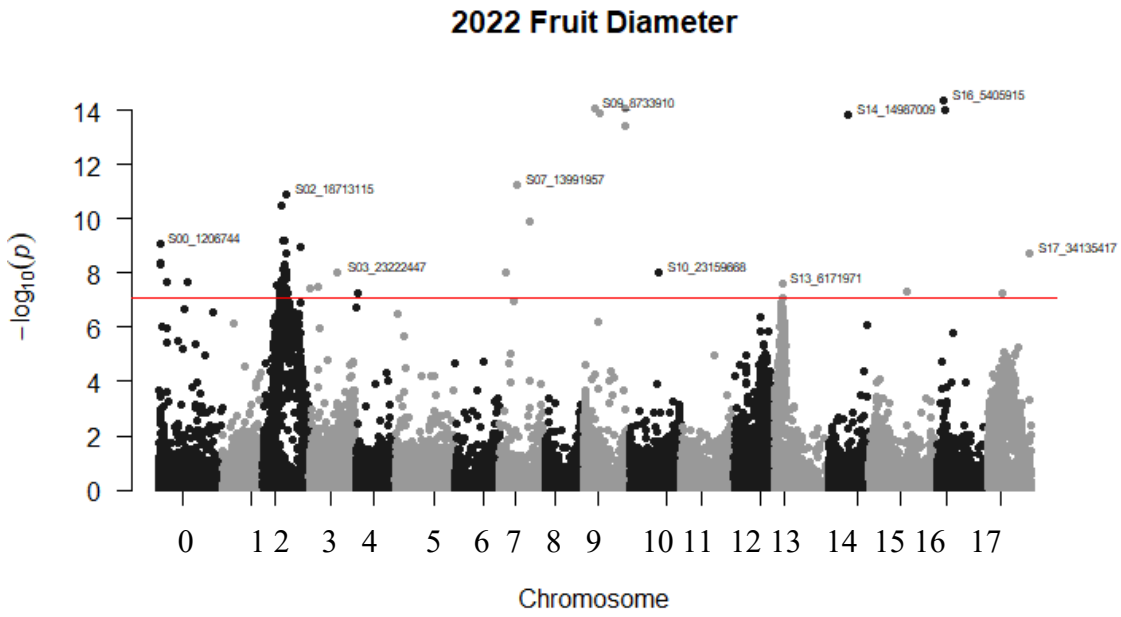


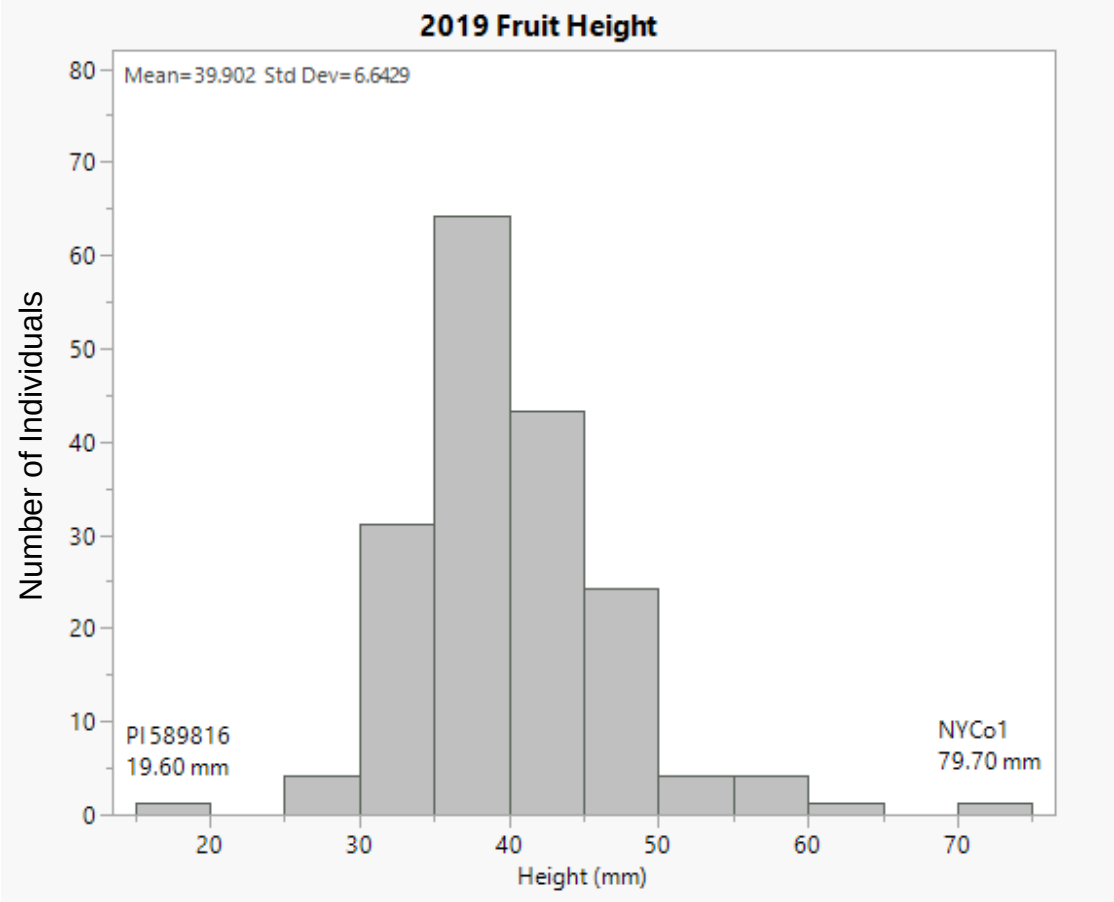
Figure 2.2 Manhattan Plots showing genetic associations determined by General Linear Model of Fruit Diameter in **A**. 2019, **B**. 2020, **C**. 2021 **D.** 2022. Significant LOD is marked at

7 with a red line.

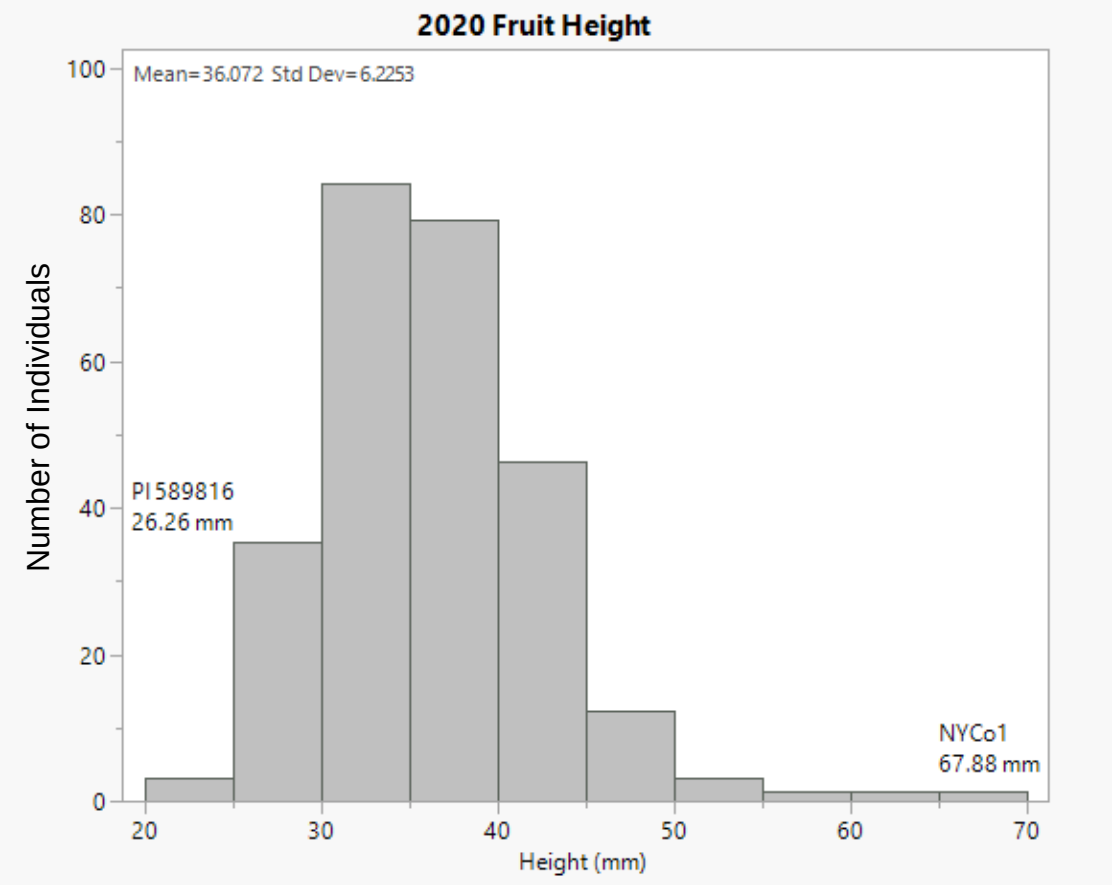
Height

Fruit height was measured in four years. In 2019, 177 individuals were measured and had a range of 19.60 mm - 79.17 mm and an average of 39.90 mm. in 2020, 265 individuals fruited with a range of 22.10 mm – 67.88 mm, and an average of 36.07 mm. in 2021, 307 individuals were measure with a range of 27.36 mm – 73.40 mm and an average of 40.37 mm. In 2022, 302 individuals were measured with a range of 20.00 mm – 70.12 mm and an average of 33.75 mm (**Figure 2.3**). Fruit height was normally distributed between the two parents. PI589816 was one of the smallest fruit heights in the population each year (average height = 25.15 mm), while NYCo1 was consistently one of the largest fruit heights (average height = 71.78 mm).

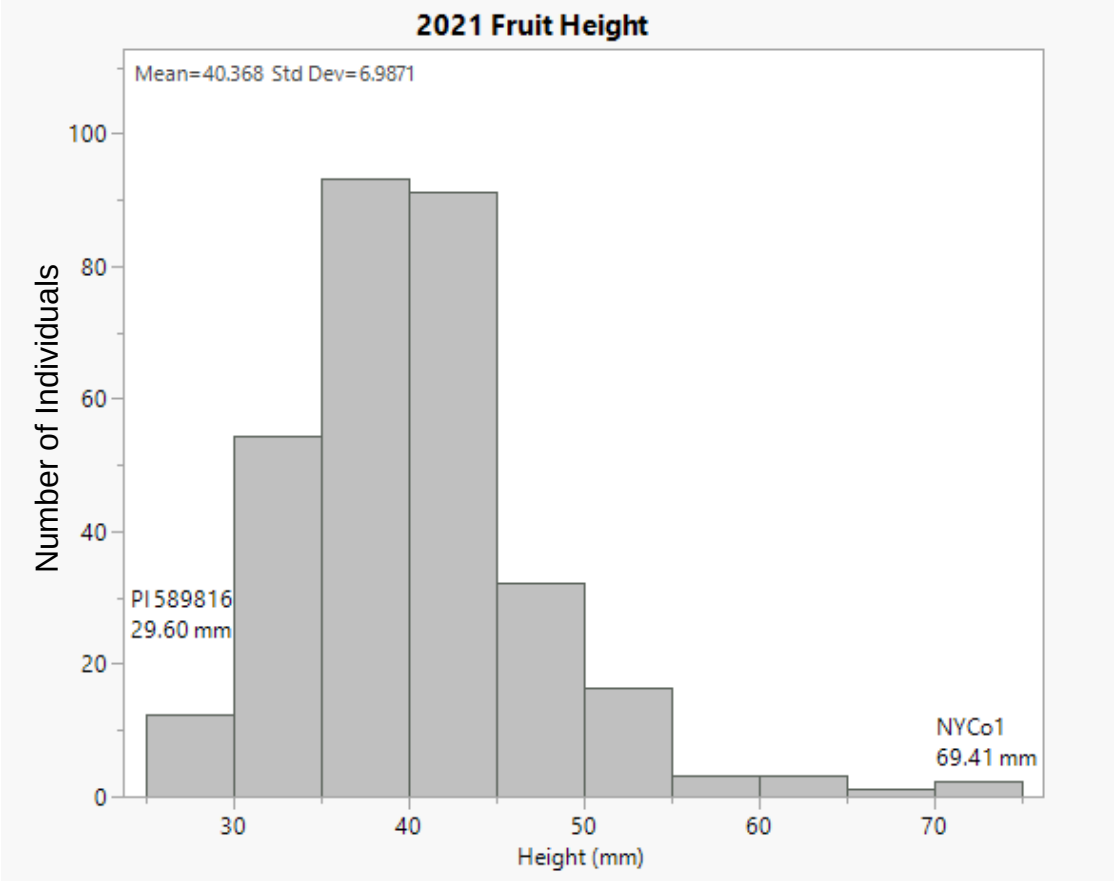
A


2019 Fruit Height
Mean=39.902 Std Dev=6.6429
Number of Individuals
80
70
60
50
40
30
20
10
0
PI 589816
19.60 mm
NYCo1
79.70 mm
20
30
40
50
60
70
Height (mm)

B


2020 Fruit Height
Mean=36.072 Std Dev=6.2253
Number of Individuals
100
80
60
40
20
0
PI 589816
26.26 mm
NYCo1
67.88 mm
20
30
40
50
60
70
Height (mm)

C


2021 Fruit Height
Mean=40.368 Std Dev=6.9871
Number of Individuals
100
80
60
40
20
0
PI 589816
29.60 mm
NYCo1
69.41 mm
30
40
50
60
70
Height (mm)

D.

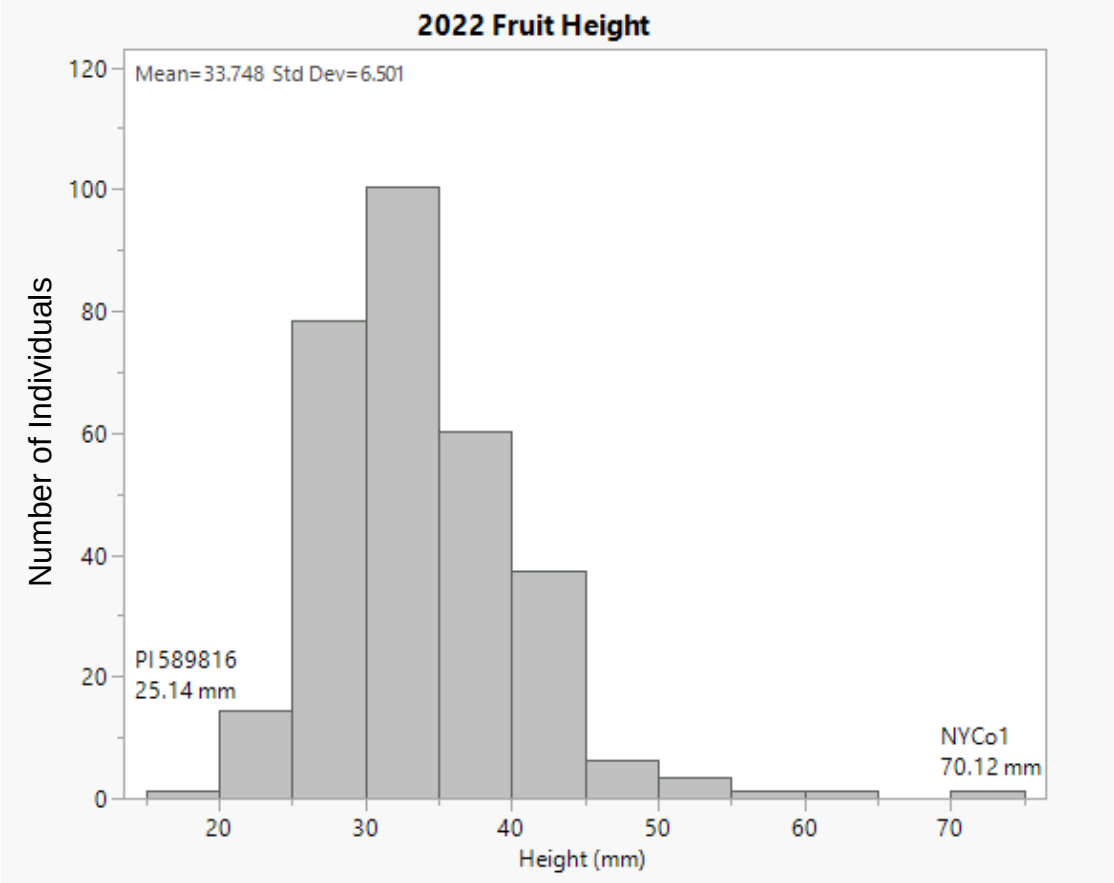
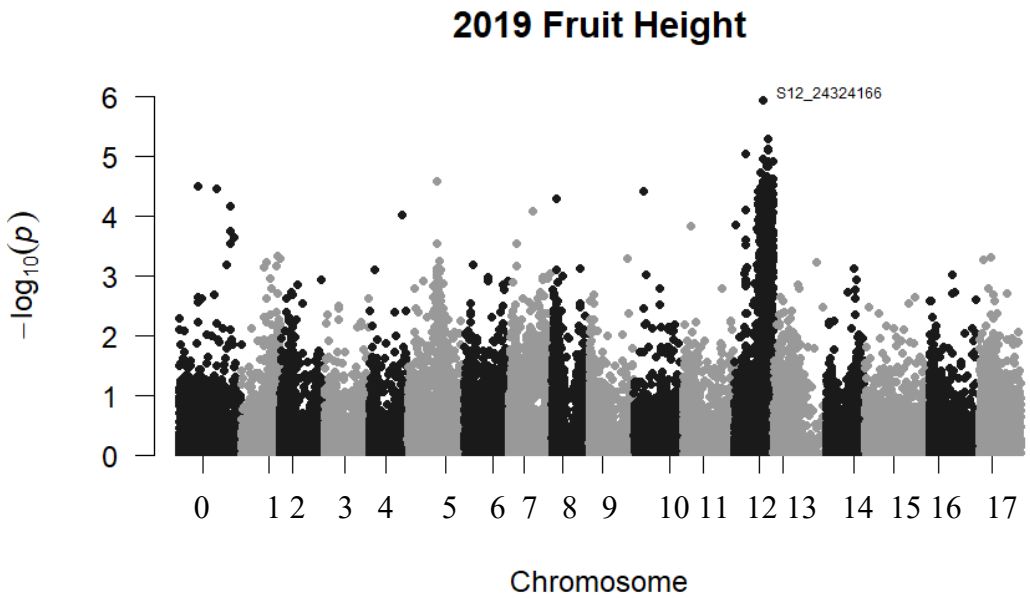


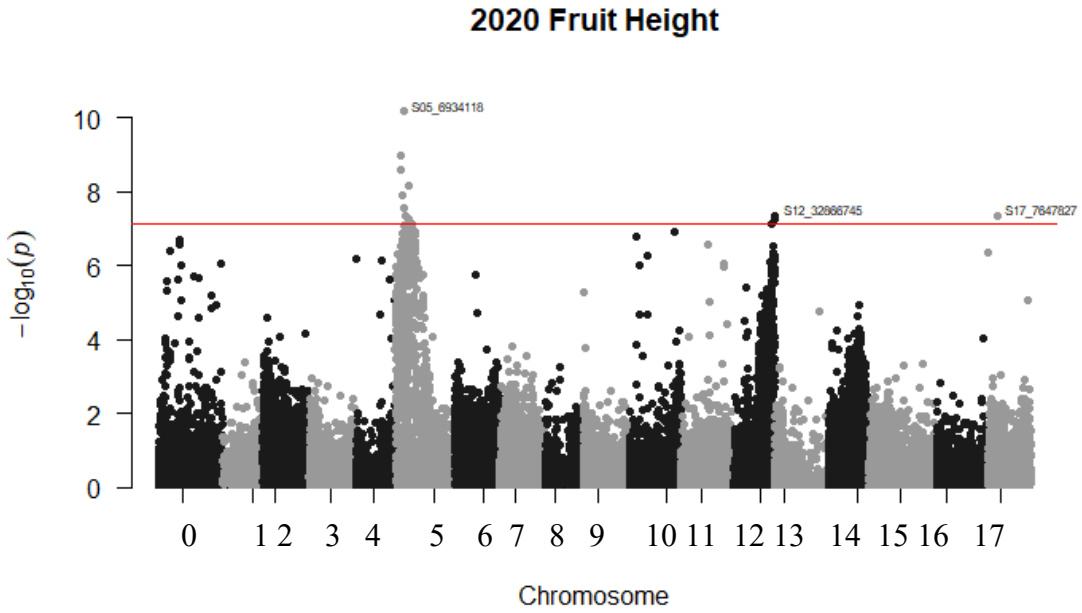
Figure 2.3 Distribution charts for fruit height for **A.** 2019, **B.** 2020, **C.** 2021, **D.** 2022. Fruit heights of PI589815 and NYCo1 are marked in each chart.

In three of the four years (2019, 2020, 2022) height was evaluated there were significant SNPs identified on chromosome 12 (**Figure 2.4)** located at positions 22935082, 24324166, and 32866745 respectively. Located near that region is a FKBP-type peptidyl-prolyl cis-trans isomerase family protein gene (MD12G1155600) at start position of 23621350 and end position 23630206 that is involved in cell differentiation.

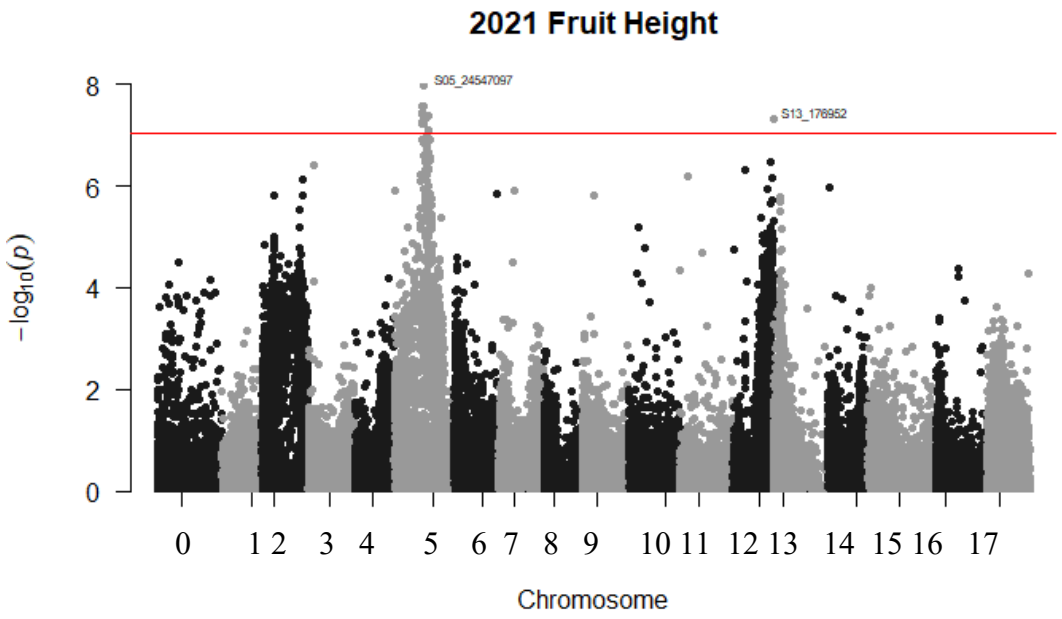
A


2019 Fruit Height
S12_24324166
−log10(p)
0
1
2
3
4
5
6
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Chromosome

B


2020 Fruit Height
S05_6934118
S12_32986745
S17_7647827
−log10(p)
0
2
4
6
8
10
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Chromosome

C



D

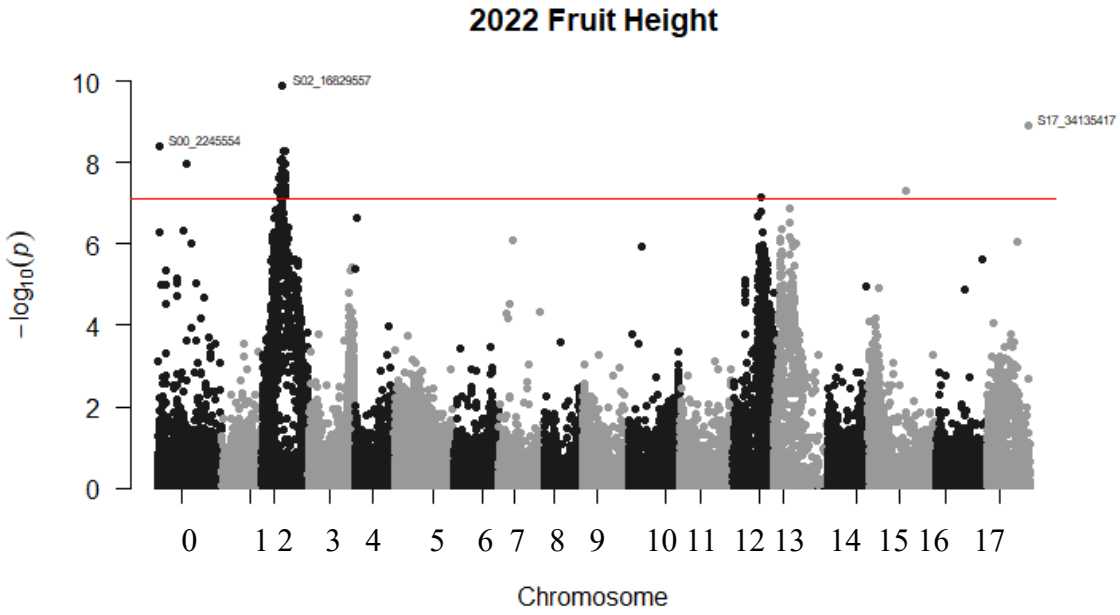


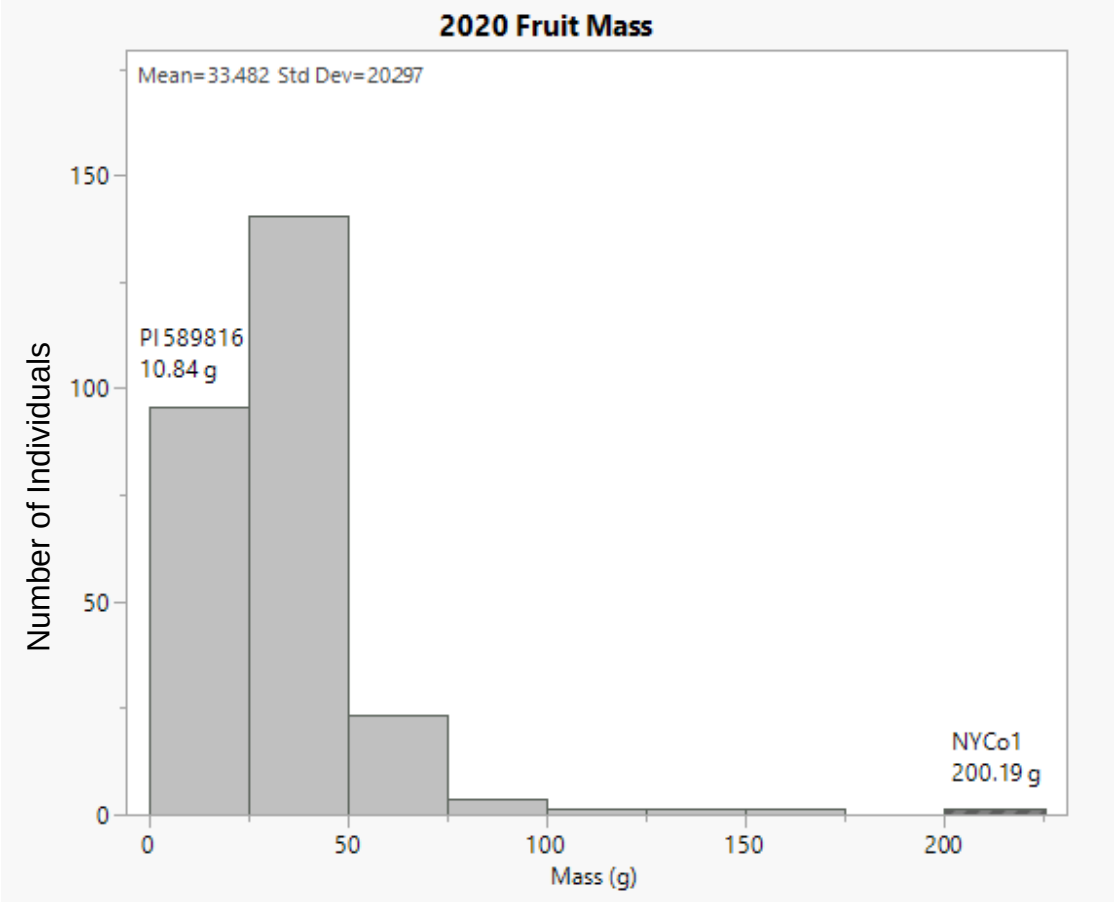
Figure 2.4 Manhattan Plots for Fruit Height associations in **A.** 2019 **B.** 2020 **C.** 2021 **D.** 2022. Significant LOD = 7 is marked with a red line.

Mass

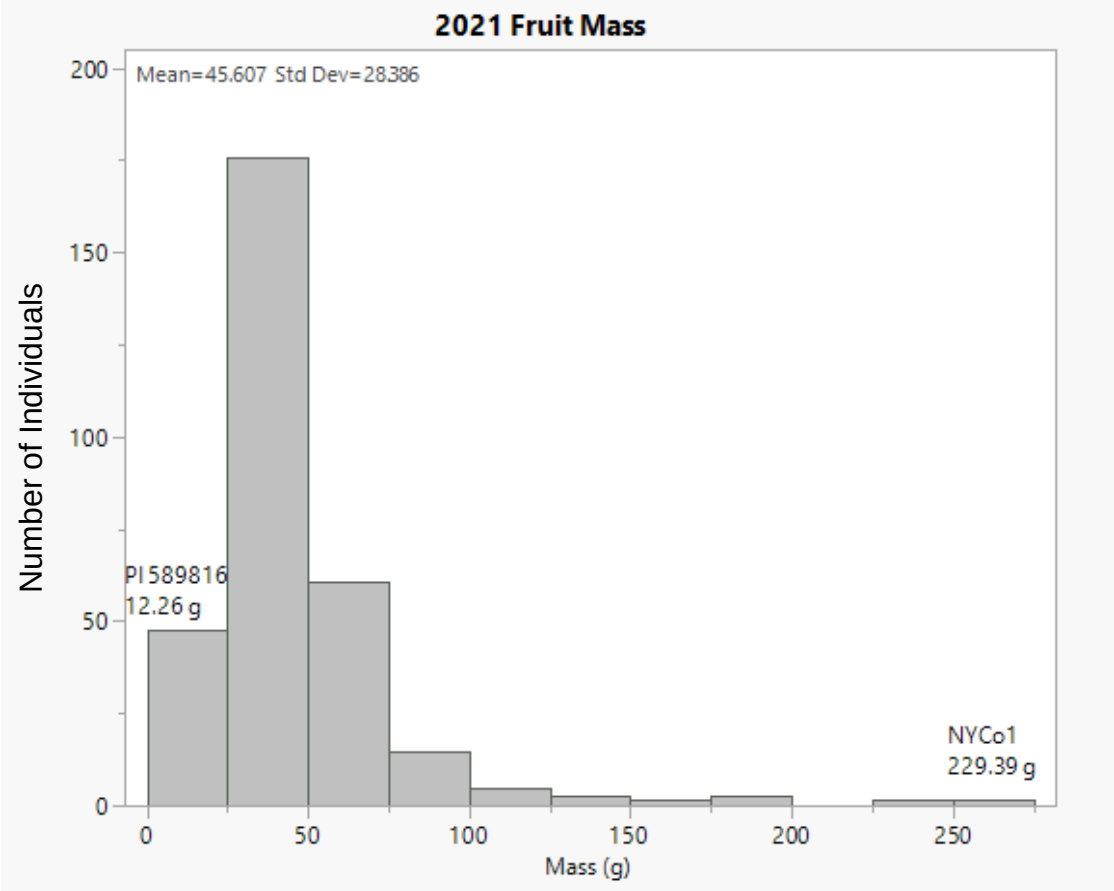
Fruit mass was measured in three years. In 2020, 265 individuals fruited with a range of 8.44 g – 200.19 g, and an average of 33.48 g. In 2021, 307 individuals were measured with a range of 12.16 g – 261.37 g and an average of 45.61 g. In 2022, 302 individuals were measured with a range of 6.91 g – 200.56 g and an average of 28.78 g (**Figure 2.5**). PI589816 was one of the smallest fruits masses in the population each year (average mass = 11.38 g), while NYCo1 was consistently one of the largest fruit masses (average mass = 210.05 g). Fruit mass was normally distributed with a strong skew favoring smaller mass.

There were no significant SNPs identified multiple years that associated with Mass. There were SNPs located in chromosomes 5 and 17 in 2020, chromosomes 2, 3, 5, 6, 7, 9, 11, 12, 13, 14 in 2021 and chromosomes 2, 4, 7, 13, 15 and 17 in 2022. The SNPs located on chromosomes 2 in 2021 and 2020 were at 30325743 and 16829557, on chromosomes 5 in 2020 and 2021 at 4847460 and 960700, on chromosome 7 in 2021 and 2022 on chromosome 7 at 13109953 and 10873972, in 2021 and 2022 on chromosome 13 at 5510629 and 6171996, and in 2020 and 2022 on chromosome 17 at 7647827 and 34135417 respectively. **Figure 2.6**

A


2020 Fruit Mass
Mean=33.482 Std Dev=20.297
Number of Individuals
150
100
50
0
PI 589816
10.84 g
NYCo1
200.19 g
0
50
100
150
200
Mass (g)

B


2021 Fruit Mass
Mean=45.607 Std Dev=28.386
Number of Individuals
200
150
100
50
0
PI 589816
12.26 g
NYCo1
229.39 g
0
50
100
150
200
250
Mass (g)

C

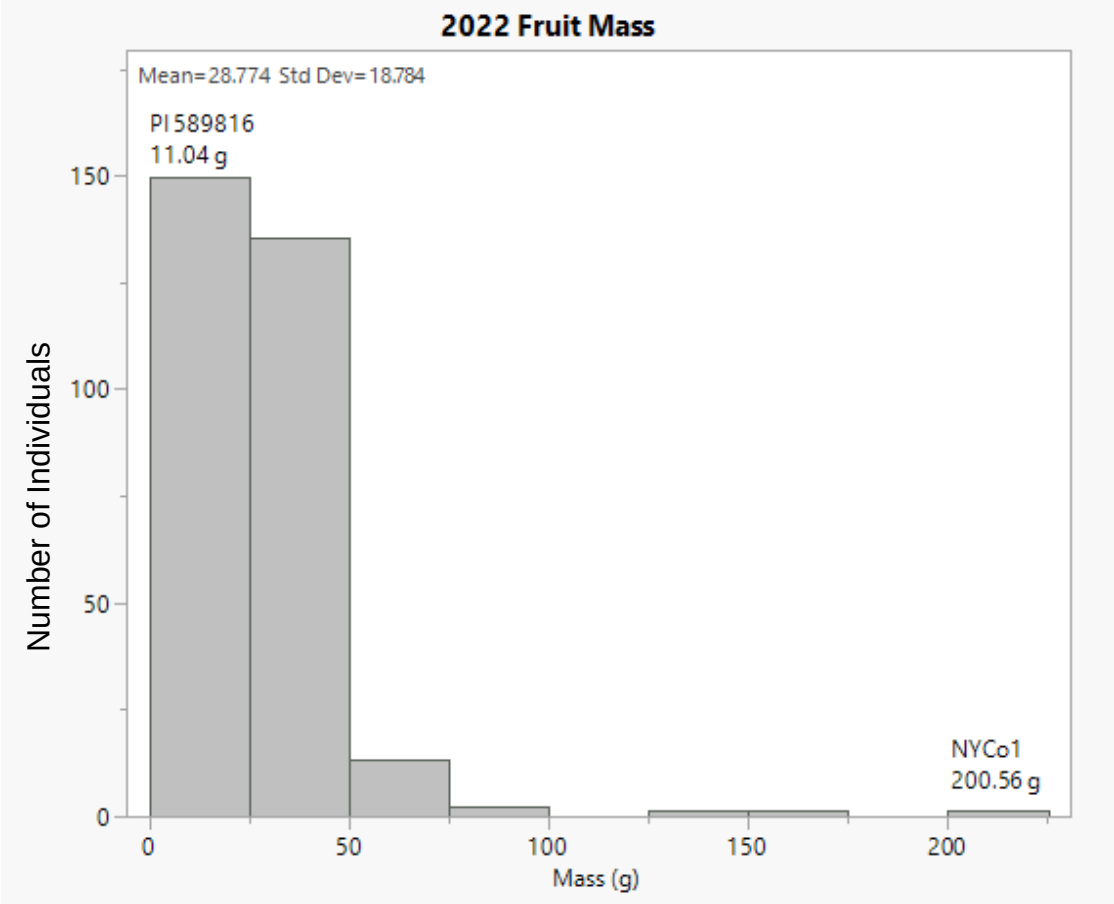
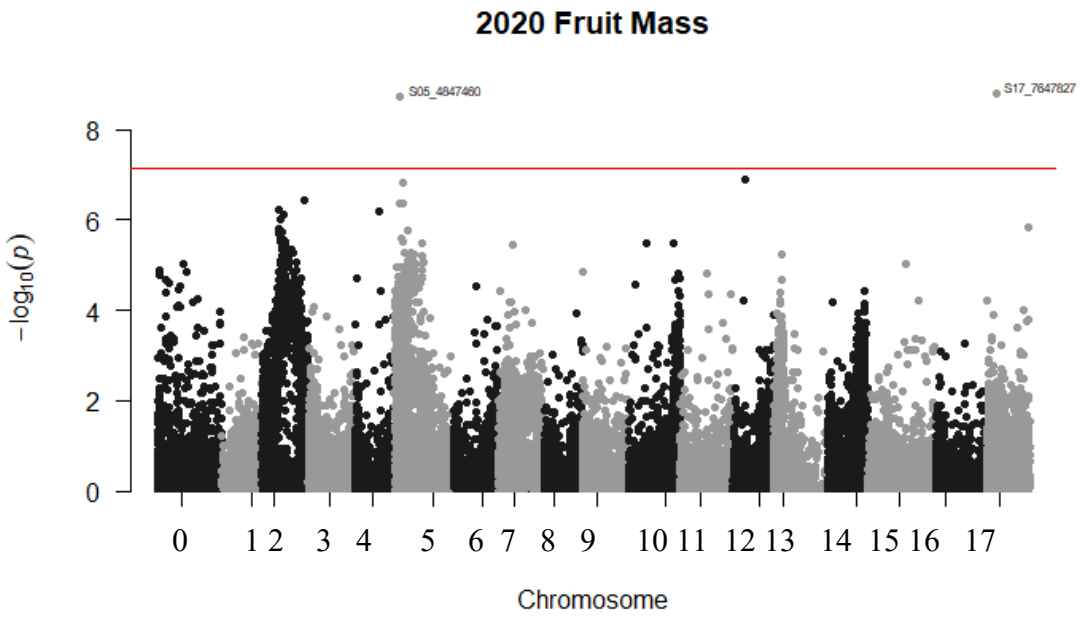
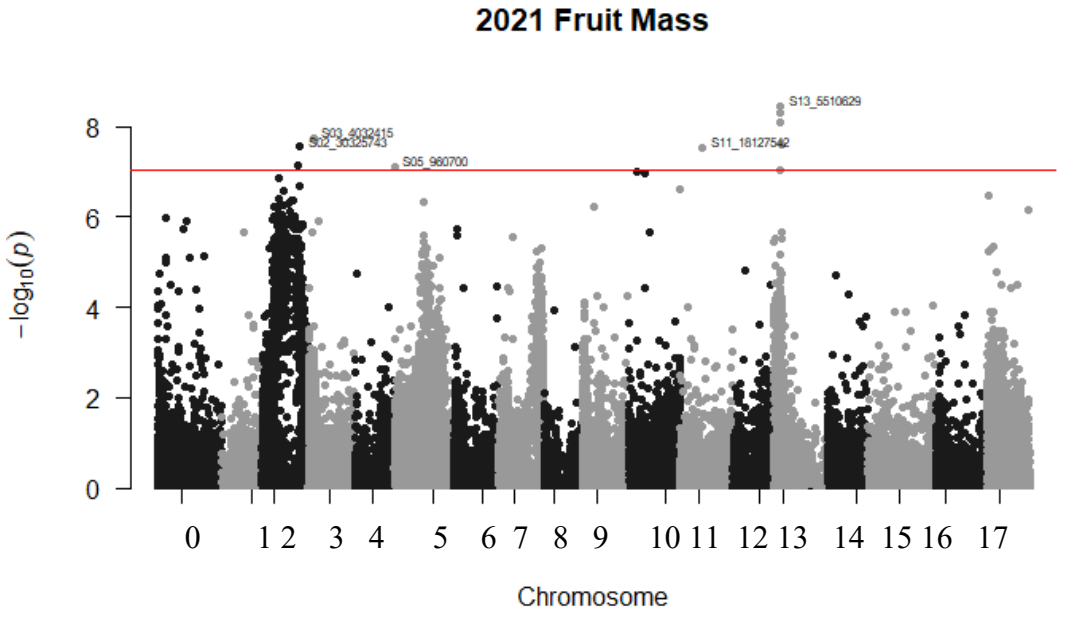


Figure 2.5 Distribution charts for fruit mass for **A.** 2020, **B.** 2021, **C.** 2022. Fruit masses of PI589815 and NYCo1 are marked in each chart.

A


2020 Fruit Mass
S05_4847460
S17_7647827
8
6
4
2
0
−log10(p)
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Chromosome

B


2021 Fruit Mass
S13_5510629
S03_4032415
S02_30325743
S11_18127542
S05_960700
8
6
4
2
0
−log10(p)
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Chromosome

C

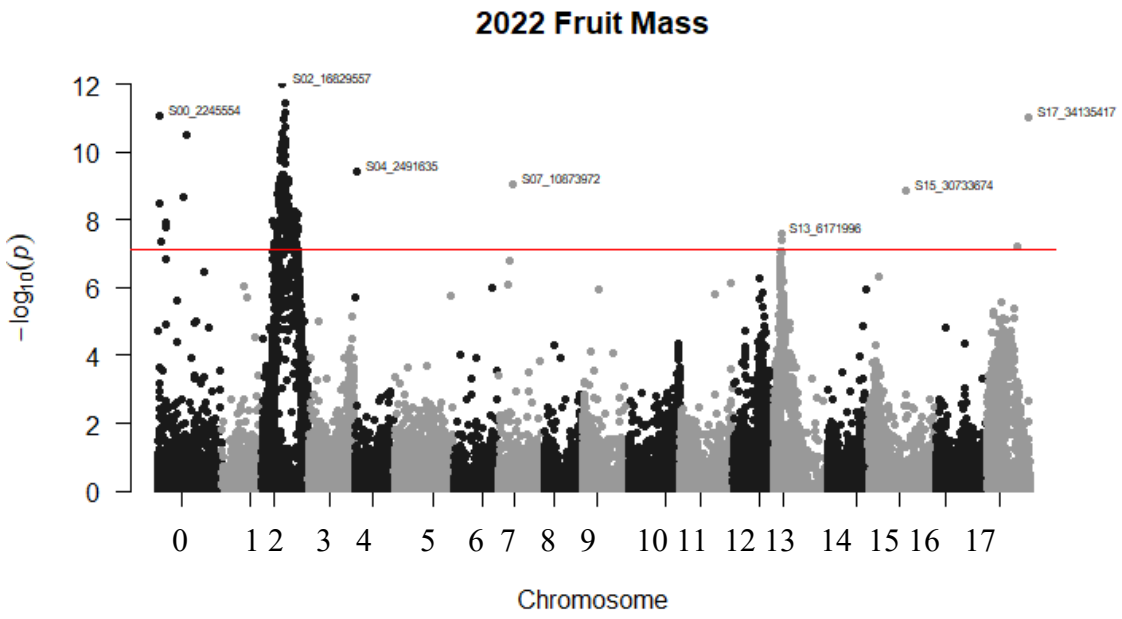
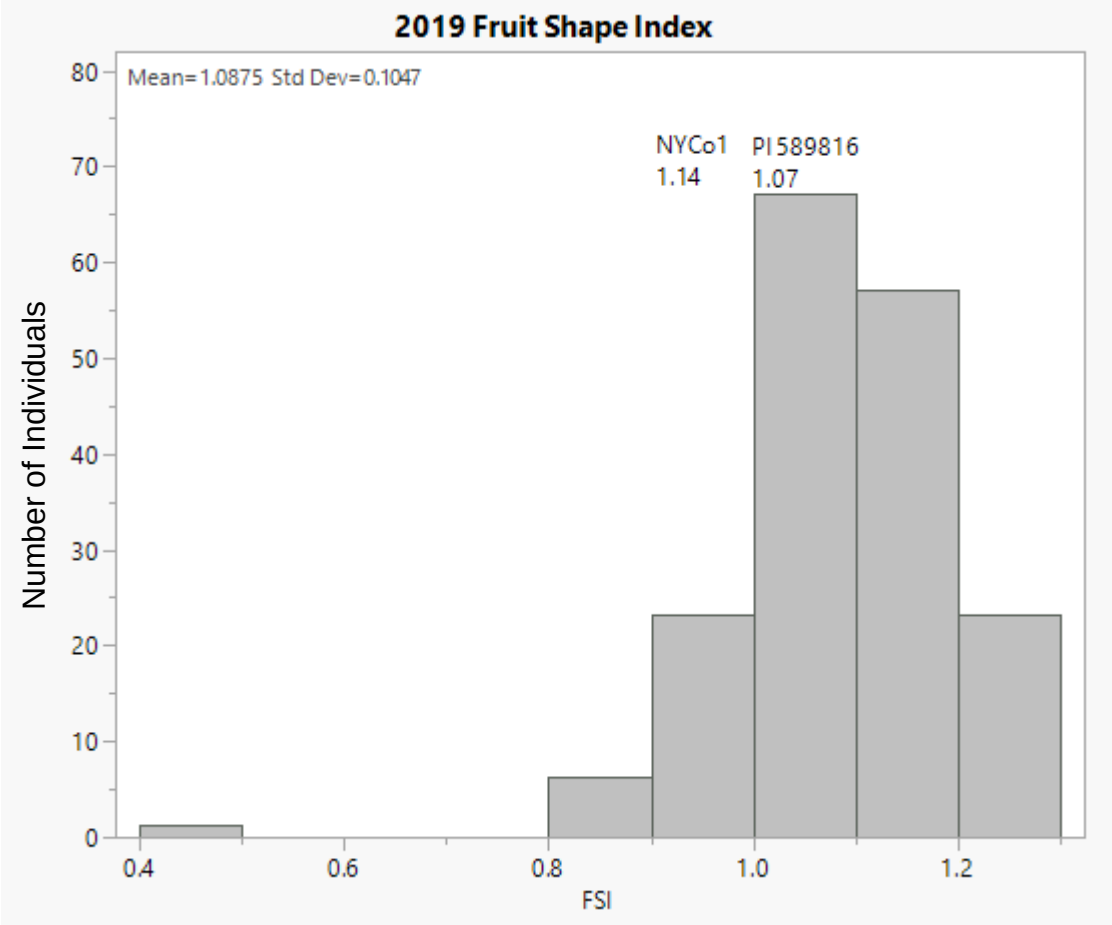


Figure 2.6 Manhattan Plots for Fruit Mass associations in **A.** 2020 **B.** 2021 **C.** 2022. Significant LOD = 7 is marked with a red line.

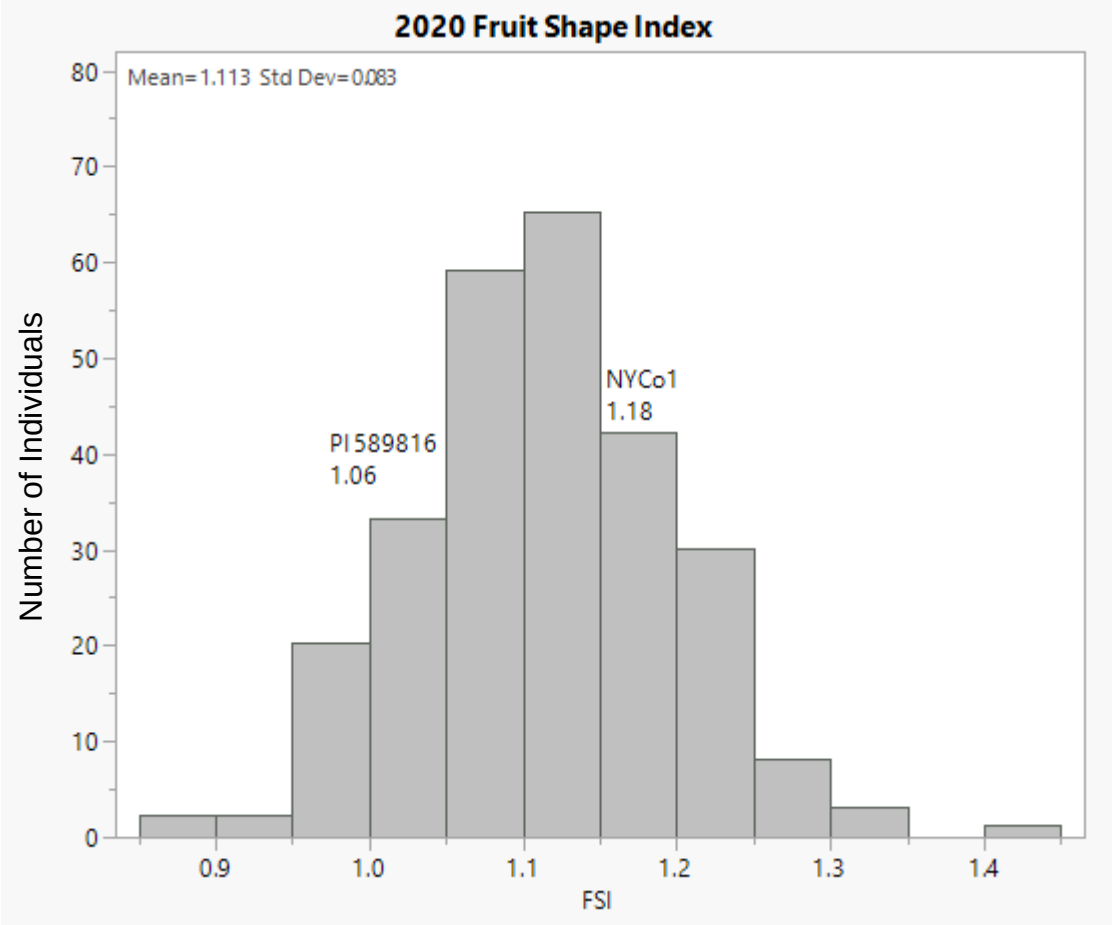
Fruit Shape Index

Fruit Shape index (FSI) was calculated from diameter and height measurements in four years. In 2019, 177 individuals were measured and had a range of 0.84 – 1.29 and an average of FSI of 1.09. In 2020, 265 individuals were measured with a range of 0.88 – 1.40 and an average of FSI of 1.11. in 2021, 307 individuals were measured with a range of 0.74 – 1.17 and an average FSI of 0.92. In 2022, 302 individuals were measured with a range of 0.85 – 1.35 and an average FSI of 1.12 (**Figure 2.7**). PI589816 and NYCo1 FSIs showed variability across years in the distribution with PI589816 having FSIs of 1.07, 1.06, 1.08 and 1.10 and NYCo1 having FSIs of 1.14, 1.18, 0.84, and 1.17 respectively each year.

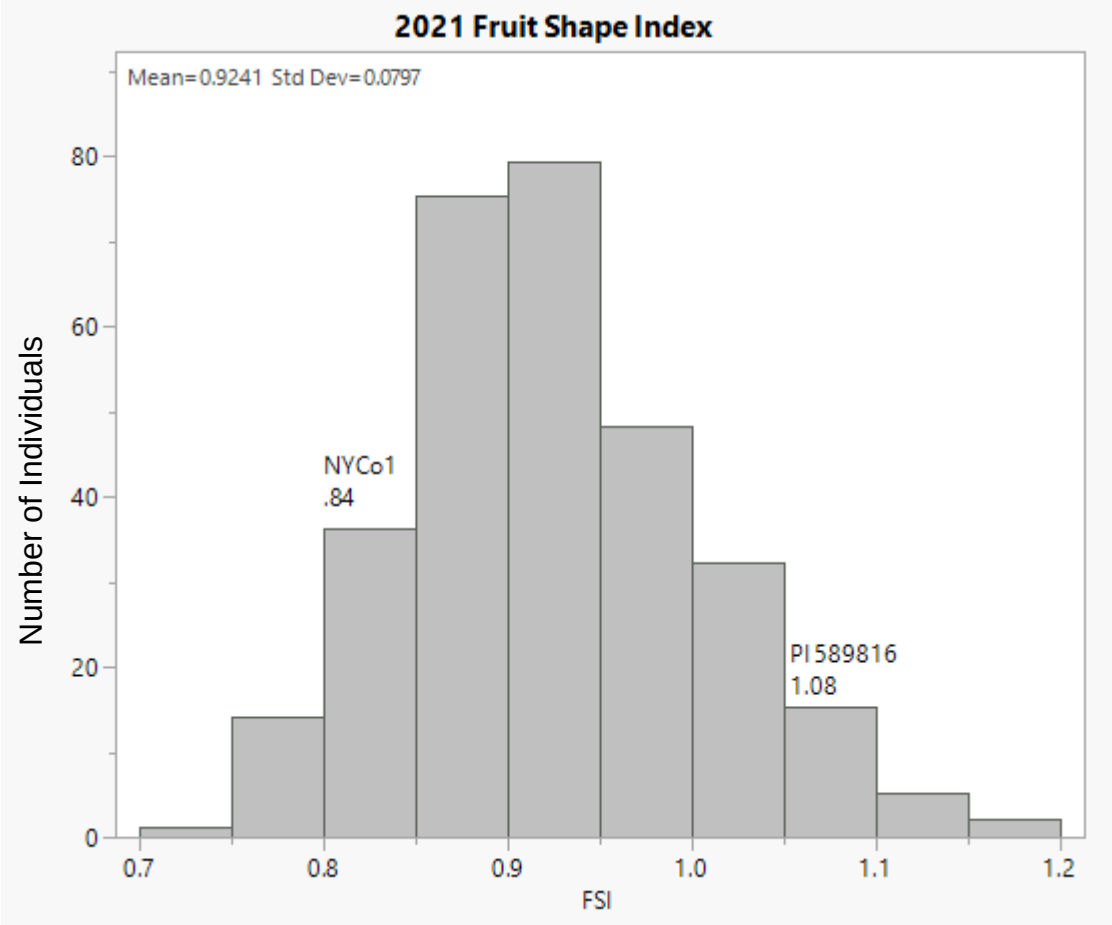
A


2019 Fruit Shape Index
Mean=1.0875 Std Dev=0.1047
NYCo1
1.14
PI 589816
1.07
Number of Individuals
80
70
60
50
40
30
20
10
0
0.4
0.6
0.8
1.0
1.2
FSI

B


2020 Fruit Shape Index
Mean=1.113 Std Dev=0.083
Number of Individuals
80
70
60
50
40
30
20
10
0
PI 589816
1.06
NYCo1
1.18
0.9
1.0
1.1
1.2
1.3
1.4
FSI

C


2021 Fruit Shape Index
Mean=0.9241 Std Dev=0.0797
Number of Individuals
80
60
40
20
0
NYCo1
.84
PI 589816
1.08
0.7
0.8
0.9
1.0
1.1
1.2
FSI

D

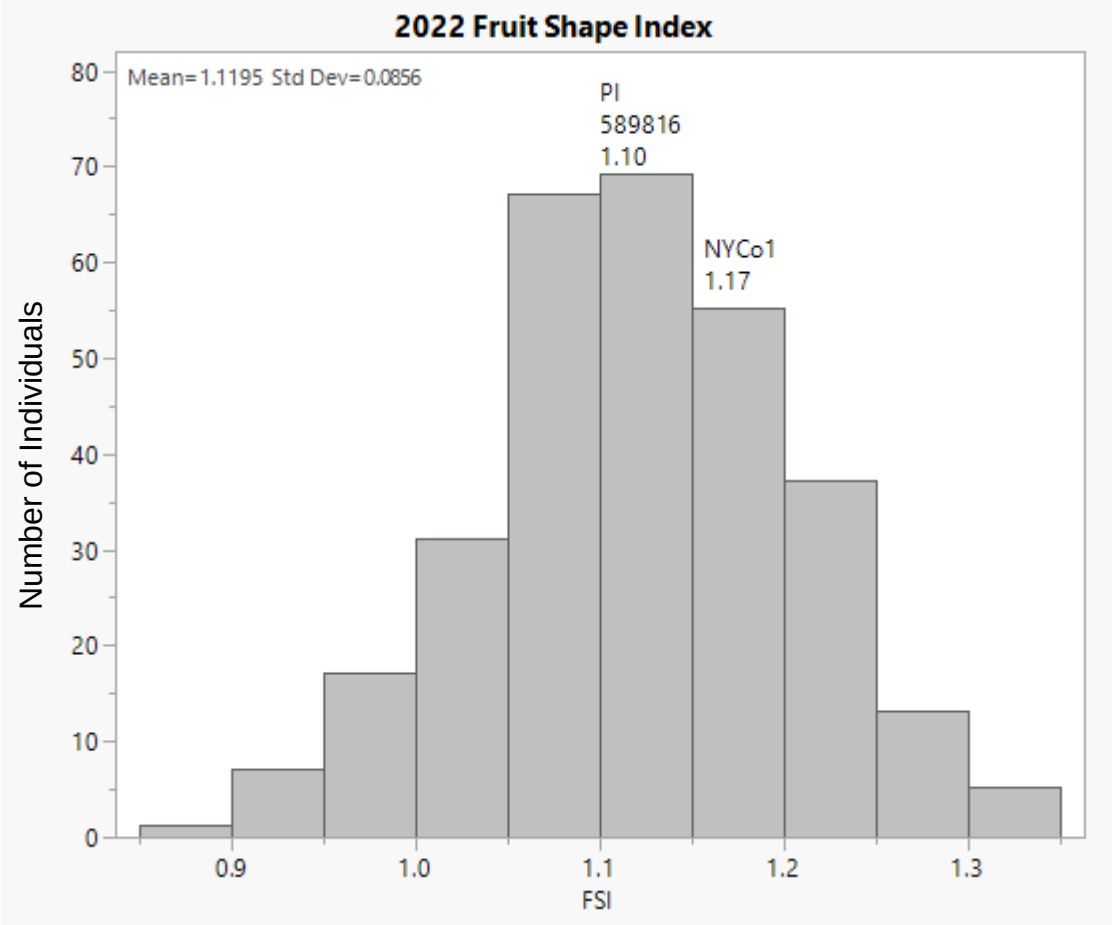
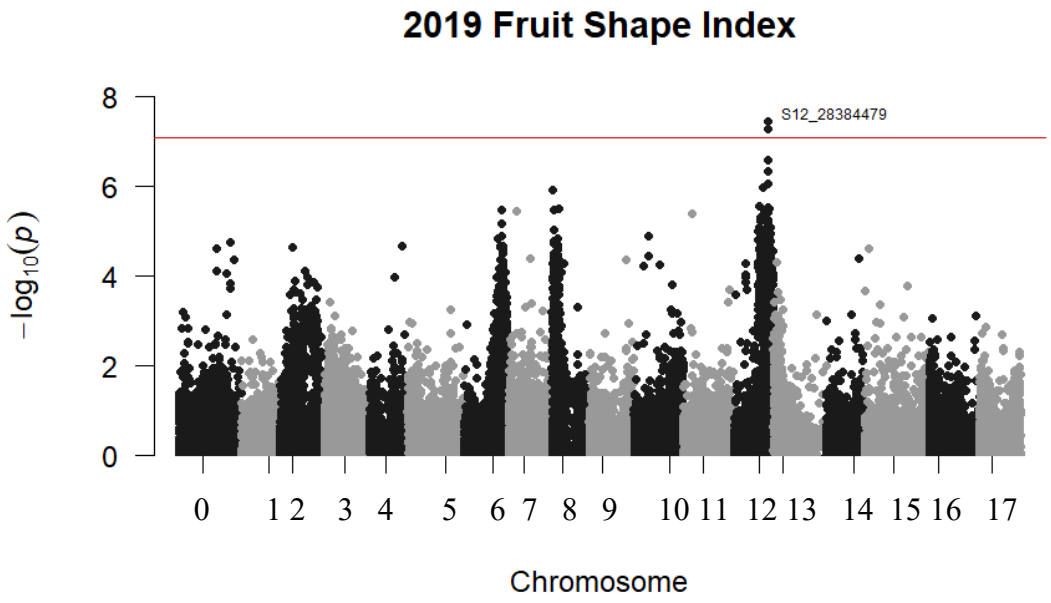


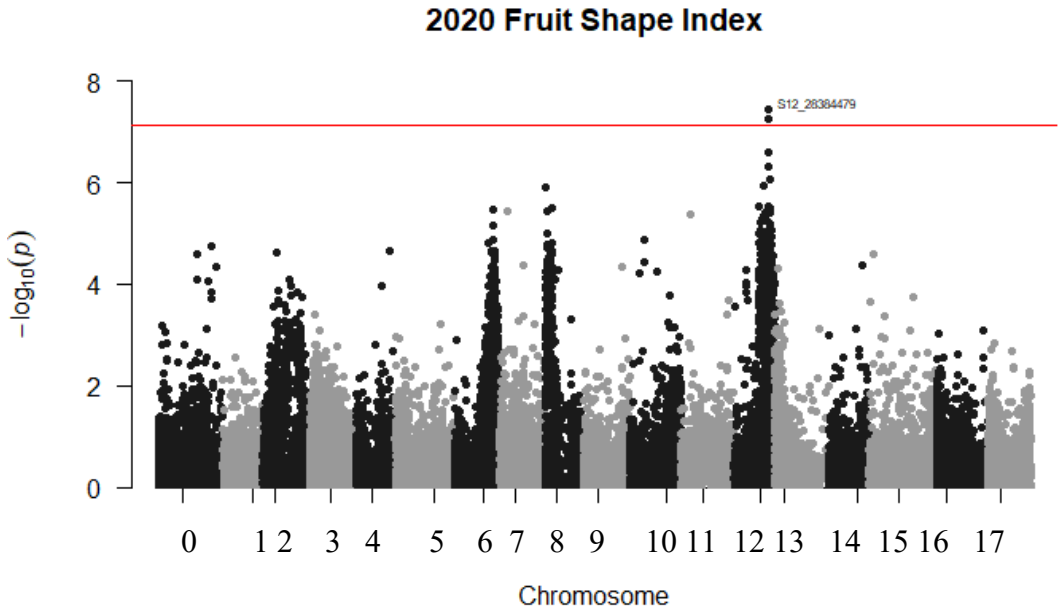
Figure 2.7 Distribution charts of fruit shape index (FSI) for **A.** 2019, **B.** 2020, **C.** 2021, **D.** 2022. FSI for PI589816 and NYCo1 are marked.

In 2019 and 2020 a SNP was identified on Chromosome 12 associated with Fruit Shape Index, the SNPs were located at position 28384479 and 28566637 respectively and there were no annotated genes within 20 kb on either side of the SNPS. In 2020 a SNP was located on Chromosome 13 at 7866484 and in 2021 a SNP was located on Chromosome 6 at 30288720. **Figure 2.8**

A


2019 Fruit Shape Index
S12_28384479
8
6
4
2
0
$-\log_{10}(p)$
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Chromosome

B


2020 Fruit Shape Index
S12_28384479
8
6
4
2
0
$-\log_{10}(p)$
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Chromosome

C

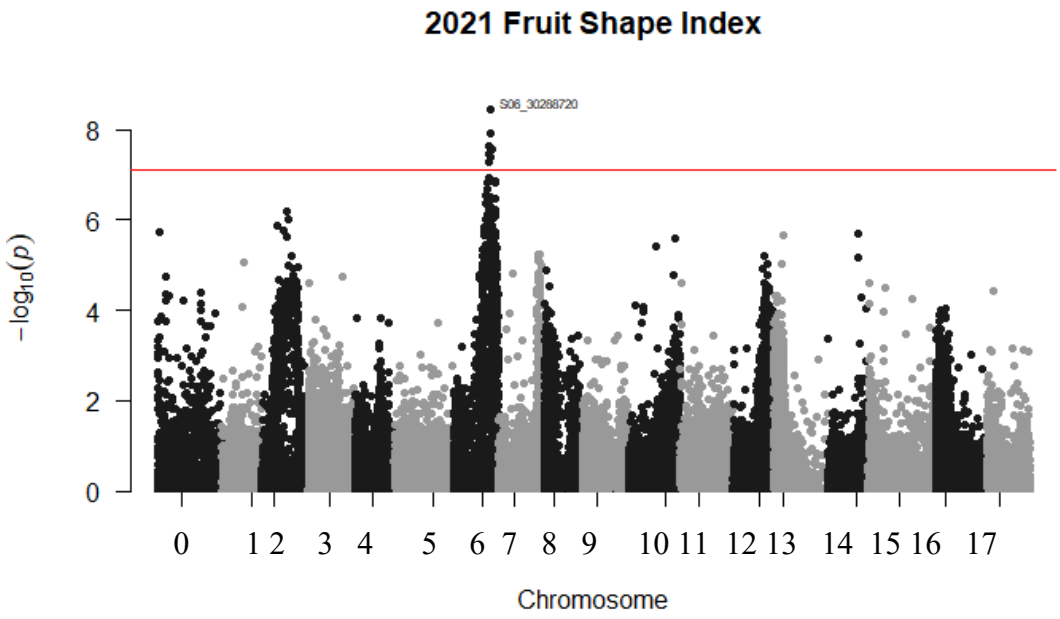


Figure 2.8 Manhattan Plots showing genetic associations determined by General Linear Model of Fruit Shape Index in **A**. 2019, **B**. 2020, **C**. 2021. Significant LOD is marked at 7 with a red line.

Across all fruit size measurements

Comparisons were done with significant SNP locations combining all trait classes and years for fruit size. Significant regions that appeared across several classes and years were identified. In 2020 a SNP for Height and Mass was identified on Chromosome 2 at position 16829557. Searching 10 kb on each side in the GDDH13 reference genome found a single Phototropic-responsive NPH3 family protein gene (MD02G1185300) with start position 16826500 and end position 16831141 involved in protein binding.

Also, on chromosome 2 in 2021 a SNP that associated with Diameter and Mass was identified at 30325743 searching the 10 kb flanking regions around this SNP

resulted in 3 genes, one of unknown function (MD02G1251600), a microtubule-associated proteins 70-1 in *Arabidopsis* (MD02G1251700) and a PR5-like receptor kinase (MD02G1251800) in *Arabidopsis* that is predicted to be a rust resistance kinase Lr10-like in apple.

In 2021 two SNPs were located on Chromosome 6 for FSI and Mass at positions 30288720 and 357270233 respectively. Within 10 kb on each side 3 genes were identified, one of unknown function (MD06G1196200), one *Arabidopsis* NAD(P)-binding Rossmann-fold superfamily protein (MD06G1196000), and a predicted NAC domain containing protein 100 (MD06G1196000) involved in DNA binding and templated DNA transcription.

On chromosome 7, SNPs for fruit mass in 2021 at position 13109953 and in 2022 for diameter at position 13991957 were found. Between these SNPS is an *Arabidopsis* C2H2-type zinc finger family protein gene (MD07G1113900) starting at position 13550825 and ending at position 13552303 that is involved in nucleic acid binding.

On Chromosome 17 a SNP for both height and mass was identified in 2020 at position 7647827, there were no genes located within 20 kb either side of this SNP Also, on chromosome 17 a SNP was identified at position 34135417 for 2020 diameter and for 2022 height and mass. A gene for an uncharacterized protein (MD17G1282100) is located nearby with the start position of 34135324 and stop position of 34136582.

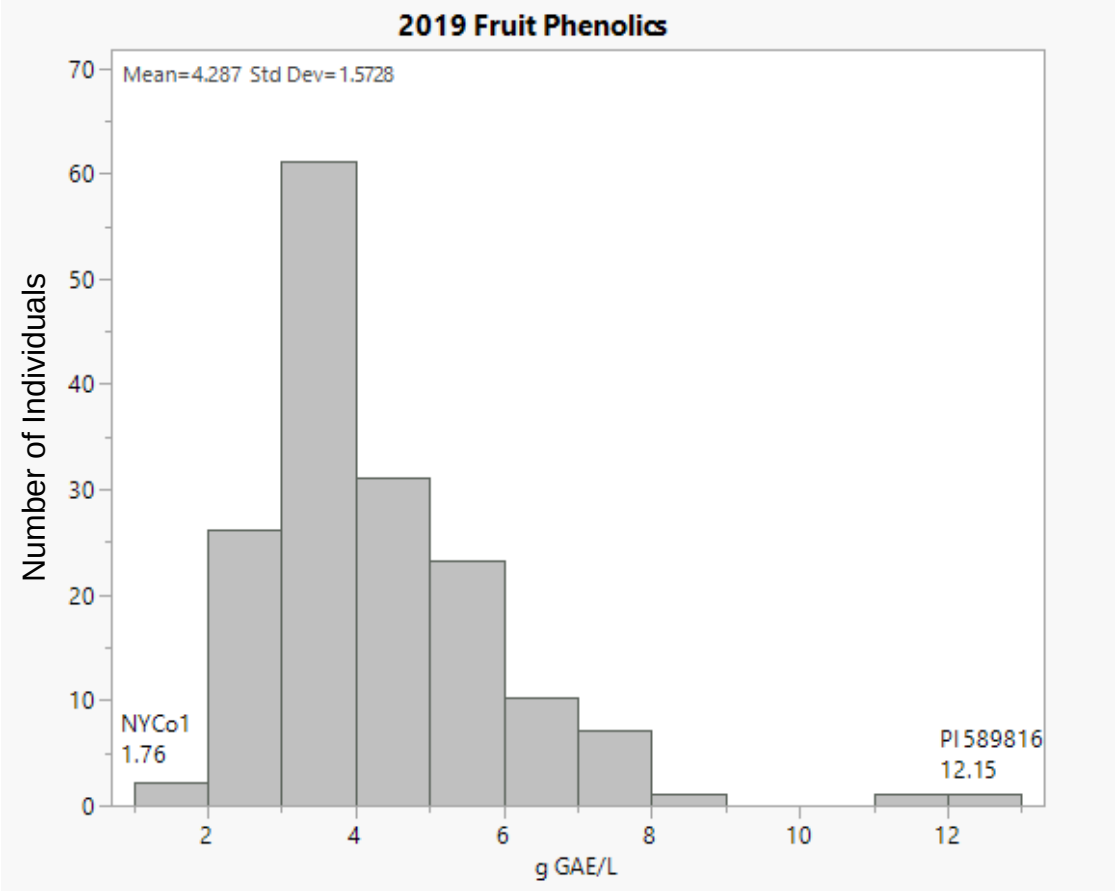
In 2021 a SNP was identified on chromosome 13 associated with diameter, height, and mass located at position 176952. This SNP is located near a UDP-N-

acetylmuramoyl-tripeptide--D-alanyl-D-alanine ligase-like gene (MD13G1002800) with start position of 171727 and stop position of 173804.

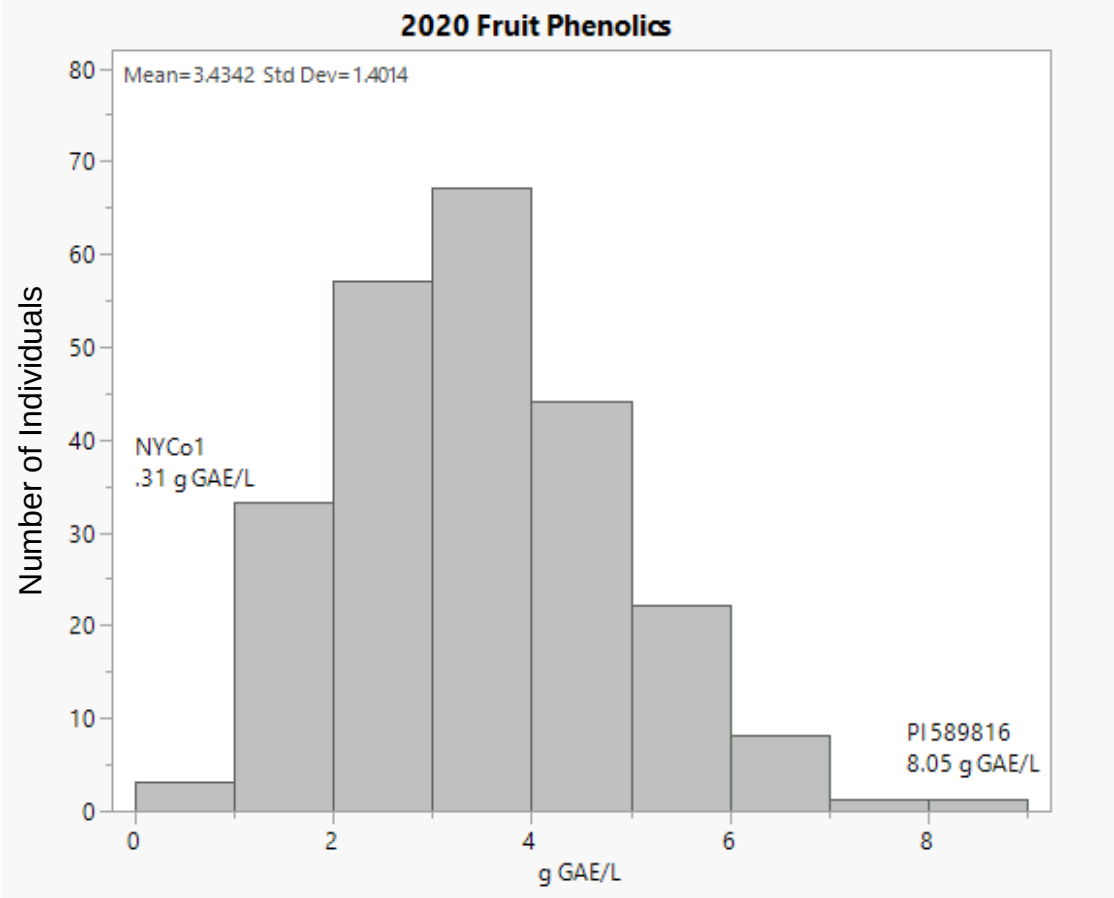
Phenolic Concentration

Phenolic concentration was measured in three years; 2019, 2020, and 2021. In 2019, 163 individuals were measured for phenolic concentration. In 2019 163 individuals were measured, concentrations varied from 1.76 g GAE/L – 12.15 g GAE/L with an average of 4.29 g GAE/L the median was 3.91 GAE/L, mean 4.29 GAE/L, standard deviation 1.57 and standard error mean .12. In 2020, 236 individuals were measured for phenolic concentration. Concentrations ranged from 0.31 g GAE/L – 8.05 g GAE/L with a median of 3.33 GAE/L, mean of 3.43 GAE/L, standard deviation of 1.40 and standard error mean of .09. In 2021, 290 individuals were measured for phenolic concentration. Concentration ranged from 0.18 g GAE/L – 6.11 g GAE/L with a median of 1.48 GAE/L, average concentration of 1.61 g GAE/L. standard deviation of .80 and standard error mean of .05. Phenolic concentration was normally distributed each year. PI589816 displayed one of the highest concentrations of phenolics each year (average concentration = 8.77 g GAE/L), while NYCo1 was consistently one of the lowest concentrations of phenolics (average concentration = 0.75 g GAE/L).

A


2019 Fruit Phenolics
Mean=4.287 Std Dev=1.5728
Number of Individuals
70
60
50
40
30
20
10
0
NYCo1
1.76
PI 589816
12.15
2
4
6
8
10
12
g GAE/L

B


2020 Fruit Phenolics
Mean=3.4342 Std Dev=1.4014
Number of Individuals
80
70
60
50
40
30
20
10
0
NYCo1
.31 g GAE/L
PI 589816
8.05 g GAE/L
0
2
4
6
8
g GAE/L

C

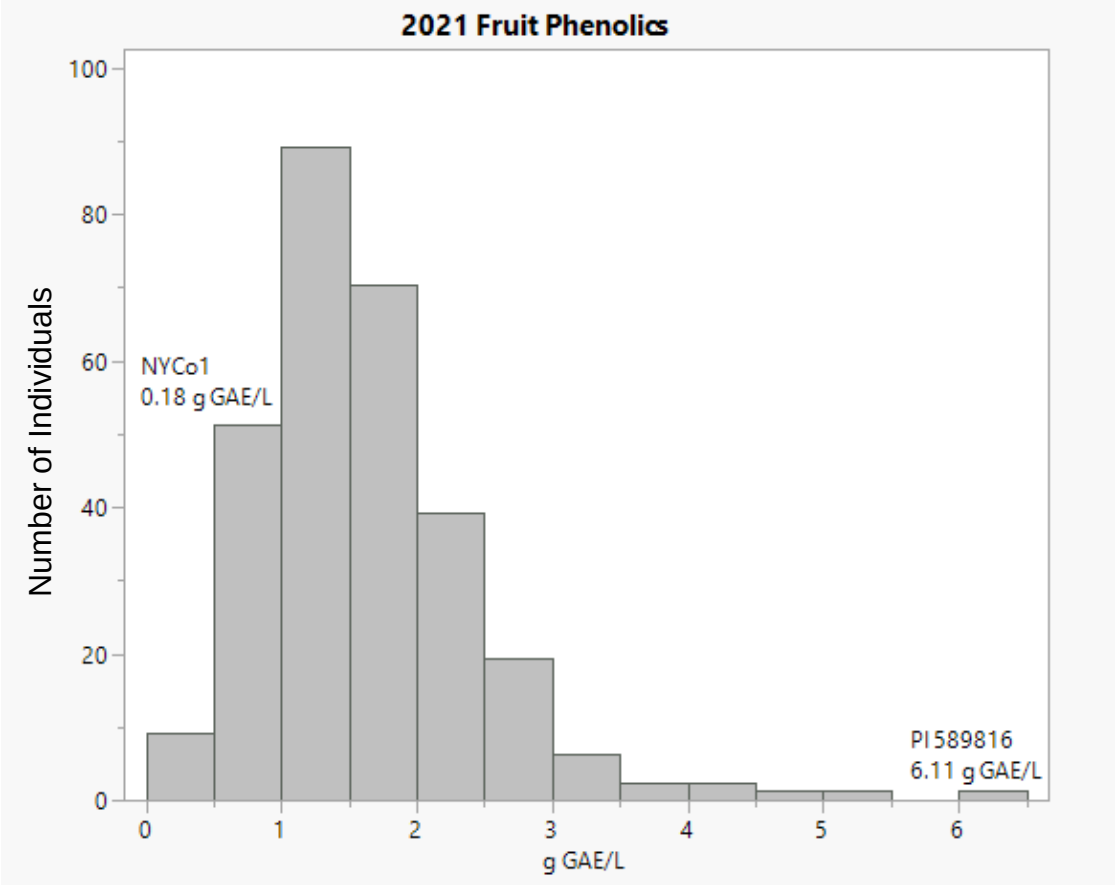


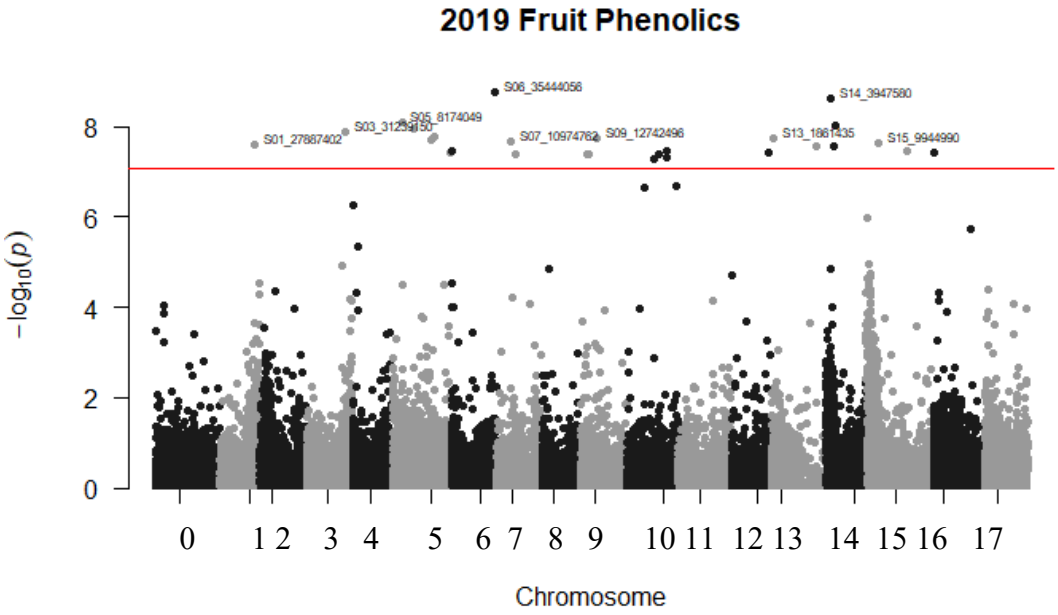
Figure 2.9 Distribution charts of fruit phenolic concentration for **A.** 2019, **B.** 2020, **C.** 2021. Phenolic concentrations for PI589816 and NYCo1 are marked.

In two of the three years (2020 and 2021) phenolic concentration was measured, and a significant SNP was identified on chromosome 16. This SNP was located at position 3438953 in both years. A significant SNP was identified on chromosome 16 in 2019 at position 119652. Located near the SNP identified in 2020 and 2021 is a Leucine-rich repeat protein kinase family protein (MD16G1048500). In 2019 and 2020 SNPs were identified on chromosome 15 at position 5323099 and 9944990 respectively. In 2019 SNPs were also identified on chromosomes 1, 3, 5, 6,

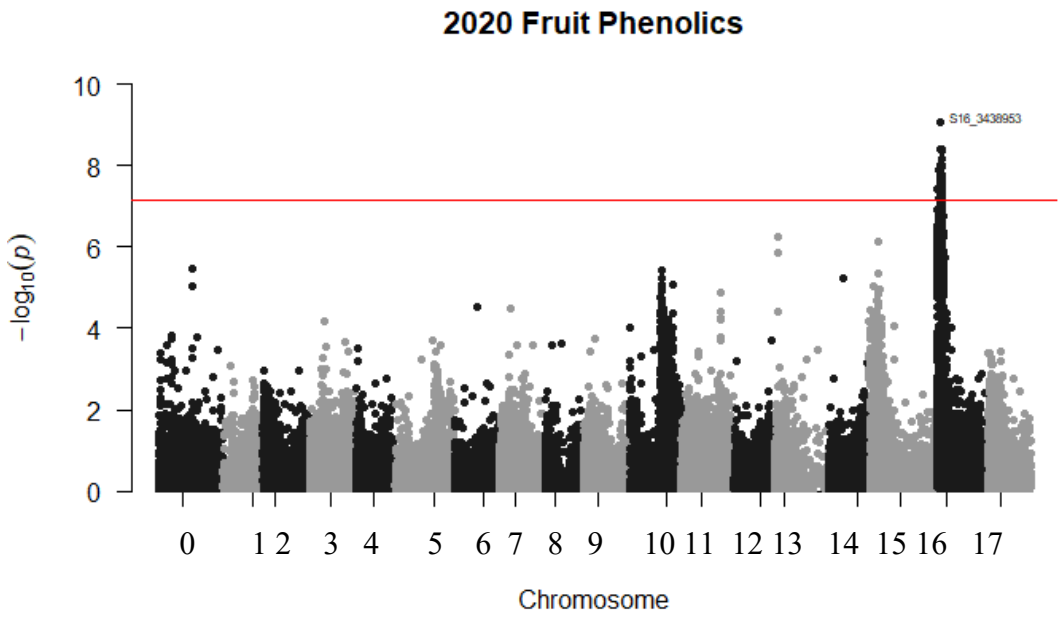
7, 9, 10, 12, 13, and 14 but these SNP associations were not identified in other years.

Figure 2.10

A



B


2020 Fruit Phenolics
S16_3438953
−log10(p)
10
8
6
4
2
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Chromosome

C

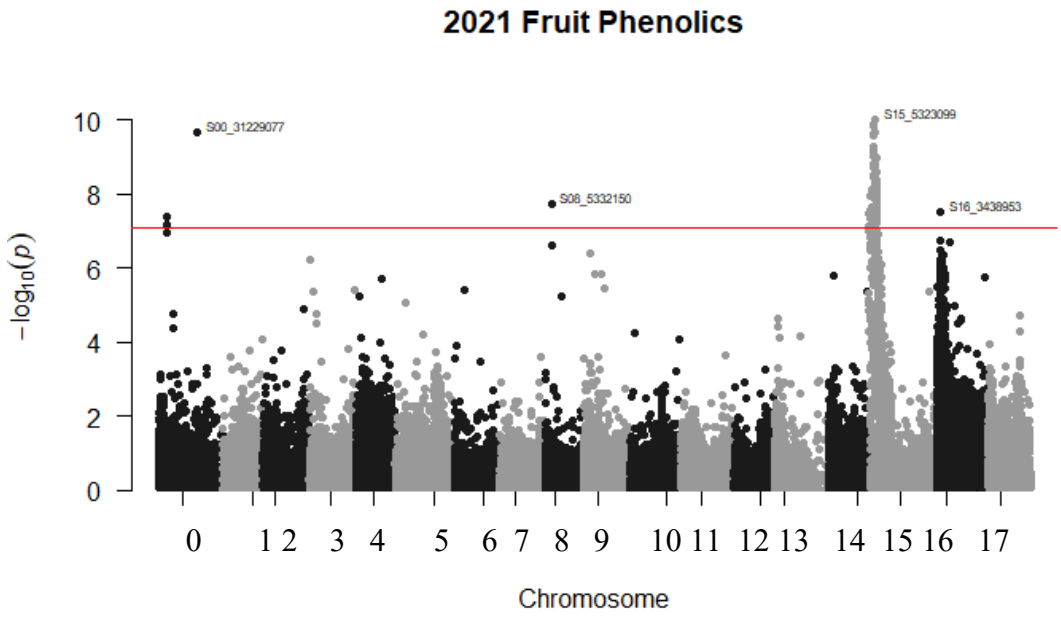
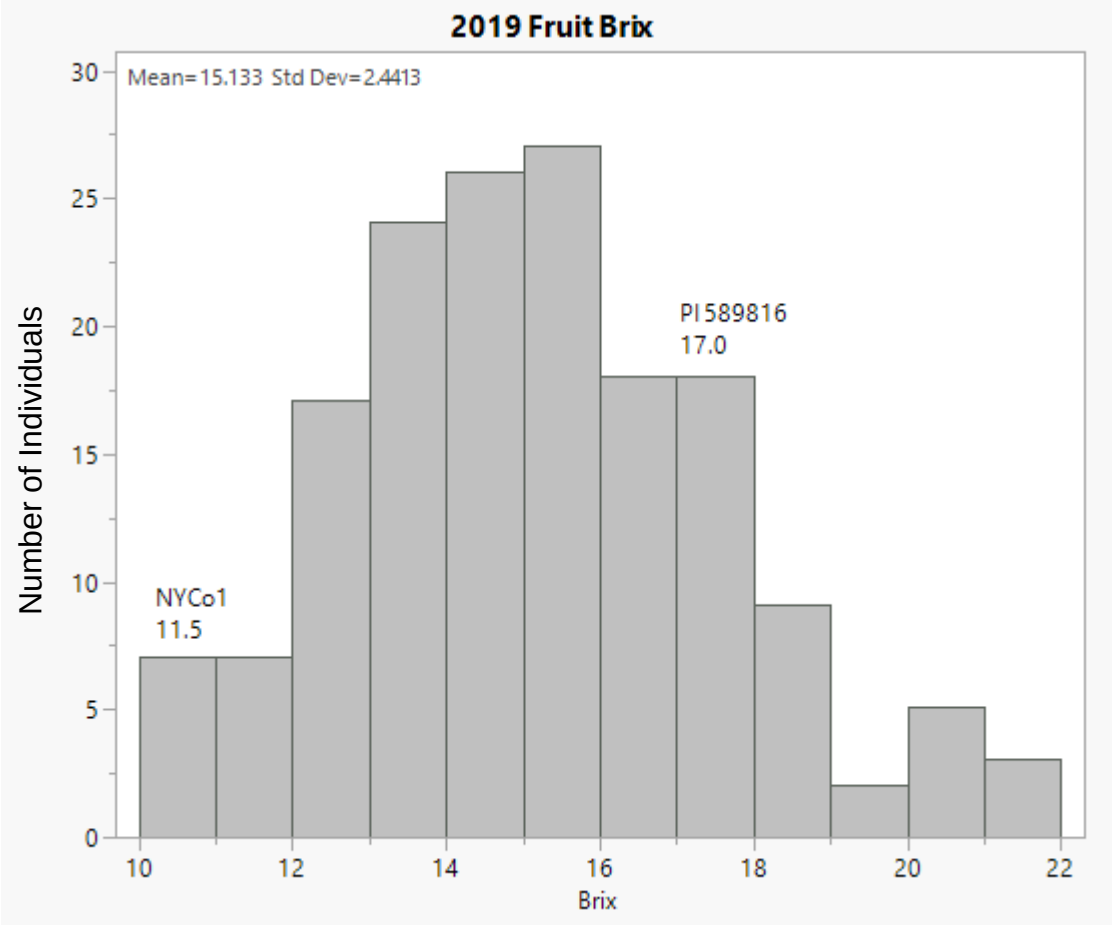

2021 Fruit Phenolics
S00_31229077
S08_5332150
S15_5323099
S16_3438953
−log10(p)
10
8
6
4
2
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Chromosome

Figure 2.10 Manhattan Plots showing genetic associations determined by General Linear Model of Fruit Phenolic Concentration in **A**. 2019, **B**. 2020, **C**. 2021. Significant LOD is marked at 7 with a red line.

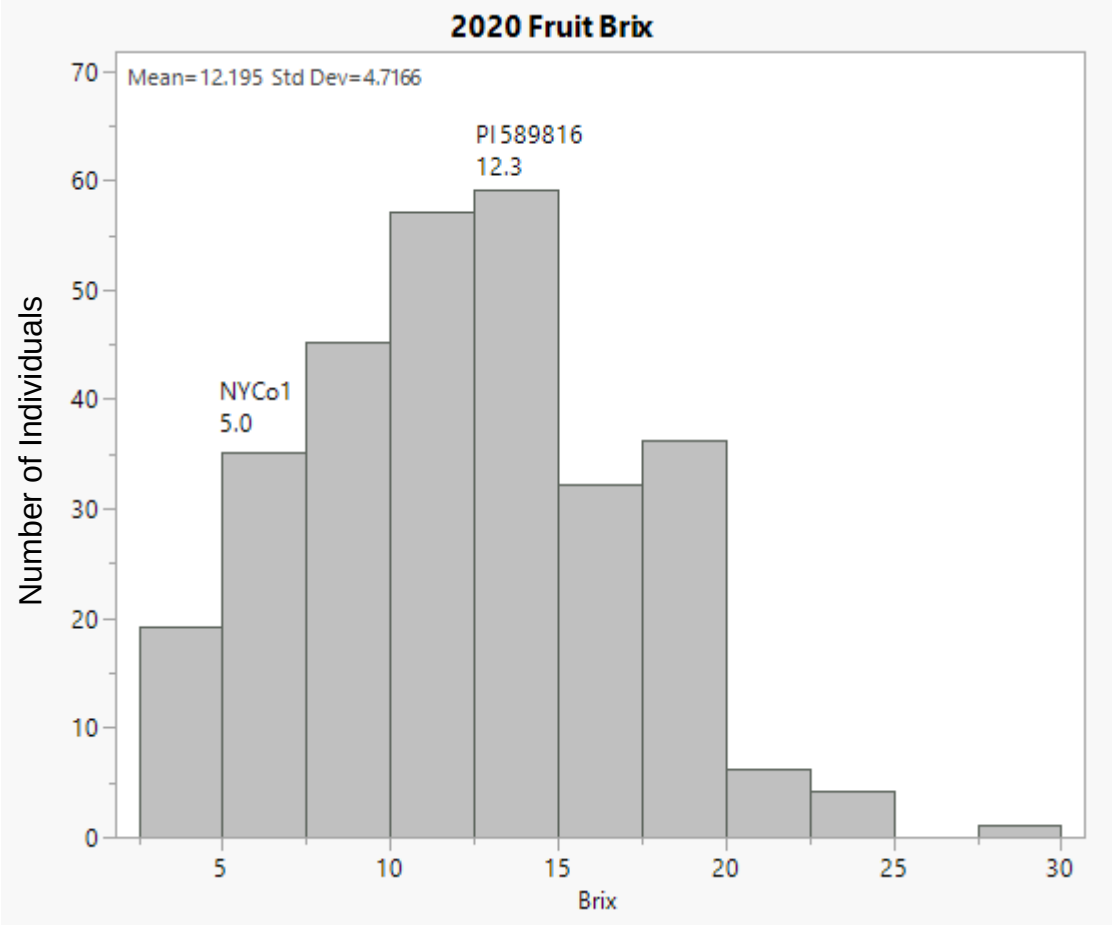
Brix

Brix was measured in three years; 2019, 2020, and 2021. In 2019. 163 individuals were measured for Brix and had a range from 10.2 – 21.5, with an average of 15.1. In 2020, 294 individuals were measured for Brix. There was a range from 2.5 – 28.0, with an average of 12.2. in 2021, 309 individuals were measured for Brix, with a range of 5.6 – 23.5 and an average of 15.0. Data was normally distributed each year. In all three years that Brix was measured NYCo1 consistently had lower Brix than PI589816, (PI589816 average Brix = 15.8, NYCo1 average Brix = 9.7) though their specific location in the distribution shifted between the years. **Figure 2.11**
In all the years that Brix was measured no significant SNPs were identified (LOD > 7).

A


2019 Fruit Brix
Mean=15.133 Std Dev=2.4413
Number of Individuals
30
25
20
15
10
5
0
NYCo1
11.5
PI 589816
17.0
10
12
14
16
18
20
22
Brix

B


2020 Fruit Brix
Mean=12.195 Std Dev=4.7166
PI 589816
12.3
NYCo1
5.0
Number of Individuals
70
60
50
40
30
20
10
0
5
10
15
20
25
30
Brix

C

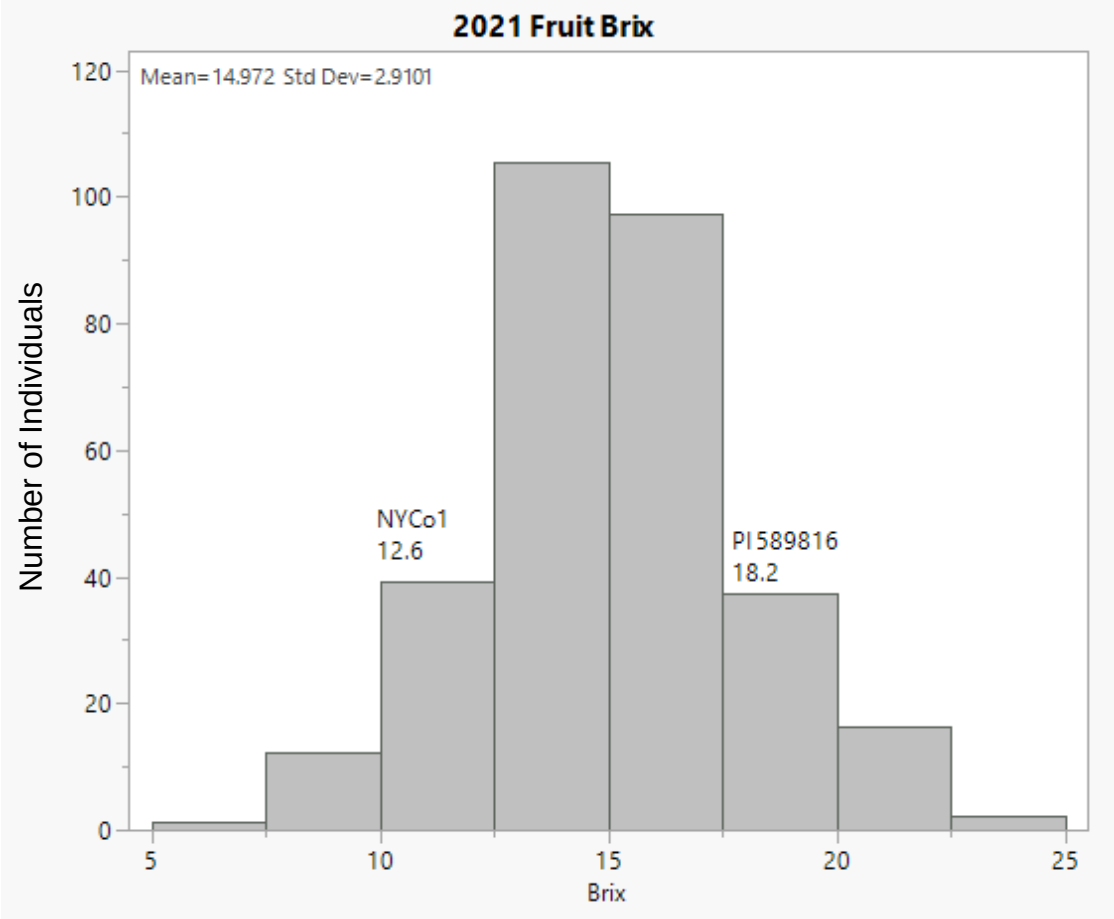


Figure 2.11 Distributions of Brix measurements in **A.** 2019 **B.** 2020 **C.** 2021. NYCo1 and PI589816 Brix measurements are marked.

Fruit Color

Fruit color as the presence or absence of blush was recorded in two years; 262 accessions were evaluated in 2020 and 293 accessions were evaluated in 2021. In 2020, 75 were determined to have blush coloration while 90 were determined to have blush in 2021, 32 accessions displayed blush both years.

Fruit Color was recorded as RBG values and transformed into CIE L*a*b* values for genetic association. No association with increased red (a/b) or with the R value in the

RGB data was found.

DISCUSSION

Fruit quality is complicated, an amalgamation of multiple different traits and specific consumer preferences combined in a hard to define measure of fruit quality. Presently, fruit quality is defined by traits like flavor, acidity, sweetness, phenolic type concentration, size, color, appearance, and peel texture. The cultivars consumers used to determine their preferences have also undergone domestication and selection over time, influencing what traits to which consumers are exposed to. In this study a wild species apple that hadn't been selected for fruit quality, was used as one parent and an advanced selection of *Malus domestica* as the other parent. The large contrast in traits between the two parents of the population, with the domestica parent being larger, with low phenolic concentration, low sugar (comparatively), and the prunifolia parent having a high concentration of phenolics, small fruit size, as well as high sugar made these traits good candidates of study and allowing a unique look into the effects of domestication and selection in the apple genome.

Fruit size is complicated and often difficult to measure trait of study. Fruit were measured by size, mass, and shape using height and diameter, mass, and the fruit shape index for shape. Comparing significant SNPs across these classes as well as years allows a greater possibility to identify regions that might control expression of fruit size. In the present study, fruit size did not appear to be controlled by a single gene, or even a small number of genes, but instead was a polygenic trait. Despite the identification of associated SNPs , most genomic regions contained no annotated genes . It is clear that advancements and upgrades to the reference genome will be

needed to further identify candidate genes associated with fruit size in this study. The gene located on chromosome 12 that associated with height in three years, a FKBP-type peptidyl-prolyl cis-trans isomerase family protein gene, is a part of a family of proteins that act as ubiquitous folding catalysts that have function in a number of cellular processes, showing activity as modulator of an intracellular calcium release channel, and to interact with virus infection in *Nicotiana* (H. H. Chang et al., 2022; Göthel & Marahiel, 1999). Further studies involving gene characterization would need to be conducted to determine if this gene is responsible for fruit heigh differences. The SNP identified on chromosome 13 in 2021 but associating with mass, diameter, and height has been shown to be active in the synthesis of the chloroplast membrane in plants, possibly relating it to fruit size generally through photosynthesis however further study would be needed to confirm this function (Dowson et al., 2022).

In the years included in this study, 2019 had the lowest number of trees fruiting (n=176) while this is still higher than other studies using similar genetic association methods, it does indicate that there might not be high enough power to detect loci that are in linkage disequilibrium with genetic markers or are controlled by several loci with weak effects. The lack of association for fruit diameter in 2019 fruit diameter might be attributed to this lower power, while in all the following years SNPs that were strongly associated were identified. Similar genetic association methods have been used in *Malus* to identify strongly linked traits such as MYB1 for fruit color with similar numbers of accessions and markers, however traits with minor effects have a lower probability of identification (McClure et al., 2018).

The SNP and nearby Leucine-rich repeat protein has previously been

associated with phenolic concentration in apple on chromosome 16 (Chagné et al., 2012; S. A. Khan et al., 2012; Verdu et al., 2014). This is the first study conducted in a wild species cross that suggests that the primary candidate gene associated with phenolic concentration is this LLR gene that has previously been identified in *M. domestica*. This result suggests that in *M. prunifolia* this LLR gene also controls phenolic concentration and that selection for lower phenolic concentration has primarily occurred on this gene through breeding and selection. Considering other genes have been identified that associate with specific phenolics like chlorogenic acid and the various dihydrochalcones, there is still opportunity to select for specific phenolics that add different health benefits or flavors. 2019 showed different association results compared to 2020 and 2021, this is likely due to a change in measurement method from a colorimetric assay to one using HPLC.

Identifying a gene that associates with total soluble solids, measured as brix, is a challenge due to the many different compounds that make up total soluble solids. In this study the range in brix levels segregated uniformly and as a result, no candidate gene regions were significantly associated with this trait. However, previous studies have been successful in associating genetic regions with specific simple or complex sugars which in combination, contribute to the trait of brix (Guan et al., 2015; Larsen et al., 2019; Li et al., 2012, 2016; Peng et al., 2020; Wei et al., 2014). Characterizing those sugars separately in this population may contribute to the foundation of understanding around the genetic control of brix but is outside the scope of study for this dissertation.

In the two years that fruit color was recorded using a colorimeter there weren't

any genetic association identified. This result is interesting considering the strong association with MYB1 discussed above. NYCo1 is the source of blush coloration in this population, however the blush is not intense, possibly attributing the low prevalence of blush in the progeny. Russet was found in this population; however, it did not segregate in large enough numbers for tests of association. PI 589816 has shown propensity to russet from field observations, but not uniformly and not consistently.

Together the traits studied assess several important fruit quality traits that have been selected upon during the initial domestication and then subsequent apple breeding (Cornille et al., 2014b, 2019; M. A. Khan et al., 2015; Migicovsky et al., 2021). This study also represents the first study focusing on fruit quality traits that uses a *M. prunifolia* accession as one of the mapping parents. *M. prunifolia* offers an opportunity to introgress important disease resistance traits that will become more important as the climate warms (Xu et al., 2022), such as resistance to the fungal pathogen of apple scab [*Venturia inaequalis* (Cooke)G)], bacterial pathogen fire blight [*Erwinia amylovora* (Burrill)], and claims of drought tolerance (Fu & Ma, 2012; Tan & Wang, 2020; Wang et al., 2011). Accessions of *M. prunifolia* have been used in rootstock breeding for their purported ability to tolerate drought and NaCl and other salts in the rootzone (Tan & Wang, 2020; Wang et al., 2011). Often introgression in apple breeding is avoided due to its long juvenility period and linkage drag from the wild species used, resulting in smaller, tannic, and low colored fruit after a long period of maturation (Le Roux et al., 2012). This study provides information and context on fruit quality traits and the potential impacts of introgression using this *M. prunifolia*

accession, that could be a valuable source of easily introgressed health compounds and disease resistance genes.

REFERENCES

Allan, A. C., & Espley, R. V. (2018). MYBs Drive Novel Consumer Traits in Fruits and Vegetables. *Trends in Plant Science*, *23*(8). https://doi.org/10.1016/j.tplants.2018.06.001

Ampomah-Dwamena, C., Driedonks, N., Lewis, D., Shumskaya, M., Chen, X., Wurtzel, E. T., Espley, R. V., & Allan, A. C. (2015). The Phytoene synthase gene family of apple (*Malus x domestica*) and its role in controlling fruit carotenoid content. *BMC Plant Biology*, *15*(1), 1–14. https://doi.org/10.1186/s12870-015-0573-7

Amyotte, B., Bowen, A. J., Banks, T., Rajcan, I., & Somers, D. J. (2017). Mapping the sensory perception of apple using descriptive sensory evaluation in a genome wide association study. *PLoS ONE*, *12*(2). https://doi.org/10.1371/journal.pone.0171710

Aprea, E., Charles, M., Endrizzi, I., Corollaro, M. L., Betta, E., Biasioli, F., & Gasperi, F. (2017). *Sweet taste in apple: the role of sorbitol, individual sugars, organic acids and volatile compounds*. https://doi.org/10.1038/srep44950

Blanpied, G. D., & Silsby, K. J. (1992). Predicting Harvest Date Window for Apples. *CCE Publications*, 1–12.

Brown, S. K., & Maloney, K. E. (2004). BIOTECHNOLOGY IN AGRICULTURE SERIES. In *Biotechnology In Agriculture Series* (pp. 475–511).

Busatto, N., Matsumoto, D., Tadiello, A., Vrhovsek, U., & Costaid, F. (2019). Multifaceted analyses disclose the role of fruit size and skin-russeting in the accumulation pattern of phenolic compounds in apple. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0219354

Cao, K., Chang, Y., Sun, R., Shen, F., Wu, T., Wang, Y., Zhang, X., & Han, Z. (2015). Candidate gene prediction via quantitative trait locus analysis of fruit shape index traits in apple. *Euphytica*, *206*(2), 381–391. https://doi.org/10.1007/s10681-015-1488-y

Chagné, D., Carlisle, C. M., Blond, C., Volz, R. K., Whitworth, C. J., Oraguzie, N. C., Crowhurst, R. N., Allan, A. C., Espley, R. V, Hellens, R. P., & Gardiner, S. E. (2007). *Mapping a candidate gene (MdMYB10) for red flesh and foliage colour in apple*. https://doi.org/10.1186/1471-2164-8-212

Chagné, D., Krieger, C., Rassam, M., Sullivan, M., Fraser, J., André, C., Pindo, M., Troggio, M., Gardiner, S. E., Henry, R. A., Allan, A. C., McGhie, T. K., & Laing, W. A. (2012). QTL and candidate gene mapping for polyphenolic composition in apple fruit. *BMC Plant Biology*, *12*(1), 12. https://doi.org/10.1186/1471-2229-12-12

Chang, H. H., Lee, C. H., Chang, C. J., & Jan, F. J. (2022). FKBP-type peptidyl-prolyl cis-trans isomerase interacts with the movement protein of tomato leaf curl New Delhi virus and impacts viral replication in Nicotiana benthamiana. *Molecular Plant Pathology*, *23*(4), 561–575. https://doi.org/10.1111/mpp.13181

Chang, Y., Sun, R., Sun, H., Zhao, Y., Han, Y., Chen, D., Wang, Y., Zhang, X., & Han, Z. (2014). Mapping of quantitative trait loci corroborates independent genetic control of apple size and shape. *Scientia Horticulturae*, *174*(1), 126–132. https://doi.org/10.1016/j.scienta.2014.05.019

Cheng, F. S., Weeden, N. F., & Brown, S. K. (1996). Identification of co-dominant RAPD markers tightly linked to fruit skin color in apple. *Theoretical and Applied Genetics*, *93*(1–2), 222–227. https://doi.org/10.1007/BF00225749

Cornille, A., Antolín, F., Garcia, E., Vernesi, C., Fietta, A., Brinkkemper, O., Kirleis, W., Schlumbaum, A., & Roldán-Ruiz, I. (2019). A Multifaceted Overview of Apple Tree Domestication. In *Trends in Plant Science* (Vol. 24, Issue 8, pp. 770–782). Elsevier Ltd. https://doi.org/10.1016/j.tplants.2019.05.007

Cornille, A., Giraud, T., Smulders, M. J. M., Roldán-Ruiz, I., & Gladieux, P. (2014a). The domestication and evolutionary ecology of apples. *Trends in Genetics*, *30*(2), 57–65. https://doi.org/10.1016/j.tig.2013.10.002

Cornille, A., Giraud, T., Smulders, M. J. M., Roldán-Ruiz, I., & Gladieux, P. (2014b). The domestication and evolutionary ecology of apples. *Trends in Genetics*, *30*(2), 57–65. https://doi.org/10.1016/j.tig.2013.10.002

Csihon, A., & Gonda, I. (2016). Fruit coloration of apple cultivars. *International Journal of Horticultural Science*, *22*(1–2), 11–14.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Dar, J. A., Wani, A. A., Ahmed, M., Nazir, R., Zargar, S. M., & Javaid, K. (2019). Peel colour in apple (*Malus* × *domestica* Borkh.): An economic quality parameter in fruit market. *Scientia Horticulturae*, *244*(July 2018), 50–60. https://doi.org/10.1016/j.scienta.2018.09.029

Delgado-Pelayo, R., Gallardo-Guerrero, L., & Hornero-Méndez, D. (2014). Chlorophyll and carotenoid pigments in the peel and flesh of commercial apple fruit varieties. *Food Research International*, *65*(PB), 272–281. https://doi.org/10.1016/j.foodres.2014.03.025

Devoghalaere, F., Doucen, T., Guitton, B., Keeling, J., Payne, W., Ling, T. J., Ross, J. J., Hallett, I. C., Gunaseelan, K., Dayatilake, G. A., Diak, R., Breen, K. C., Tustin, D. S., Costes, E., Chagné, D., Schaffer, R. J., & David, K. M. (2012). A genomics approach to understanding the role of auxin in apple (Malus x domestica) fruit size control. *BMC Plant Biology*, *12*. https://doi.org/10.1186/1471-2229-12-7

Dowson, A. J., Lloyd, A. J., Cuming, A. C., Roper, D. I., Frigerio, L., & Dowson, C. G. (2022). Plant peptidoglycan precursor biosynthesis: Conservation between moss chloroplasts and Gram-negative bacteria. *Plant Physiology*, *190*(1), 165–179. https://doi.org/10.1093/plphys/kiac176

Duan, N., Wang, N., Xiang, L., Mao, L., Legall, N., Mao, Z., Fazio, G., Wang, K., Sun, H., Wang, X., Wang, Y., Liu, J., Xu, Y., Jiao, C., Wang, D., Yin, C., Jiang, S., Fei, Z., Wan, S., … Chao, T. (2017). Genome re-sequencing reveals the history of apple and supports a two-stage model for fruit enlargement. *Nature Communications*, *8*(1). https://doi.org/10.1038/s41467-017-00336-7

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, *6*(5), 1–10. https://doi.org/10.1371/journal.pone.0019379

Falginella, L., Cipriani, G., Monte, C., Gregori, R., Testolin, R., Velasco, R., Troggio, M., & Tartarini, S. (2015). A major QTL controlling apple skin russeting maps on the linkage group 12 of 'Renetta Grigia di Torriana.' *BMC Plant Biology*, *15*(1), 150. https://doi.org/10.1186/s12870-015-0507-4

Farneti, B., Masuero, D., Costa, F., Magnago, P., Malnoy, M., Costa, G., Vrhovsek, U., & Mattivi, F. (2015). Is There Room for Improving the Nutraceutical Composition of Apple? *Agricultural and Food Chemistry*, *63*, 2750–2759. https://doi.org/10.1021/acs.jafc.5b00291

Fu, M., & Ma, F. (2012). Characterization of the genetic relationships among biotypes of Malus prunifolia using simple sequence repeat marker. *Scientia Horticulturae*, *146*, 169–174. https://doi.org/10.1016/j.scienta.2012.08.020

Gardner, K. M., Brown, P., Cooke, T. F., Cann, S., Costa, F., Bustamante, C., Velasco, R., Troggio, M., & Myles, S. (2014). Fast and cost-effective genetic mapping in apple using next-generation sequencing. *G3: Genes, Genomes, Genetics*, *4*(9), 1681–1687. https://doi.org/10.1534/g3.114.011023

Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., & Buckler, E. S. (2014). TASSEL-GBS: A High Capacity Genotyping by Sequencing Analysis Pipeline. *PLOS ONE*, *9*(2), e90346. https://doi.org/10.1371/JOURNAL.PONE.0090346

Göthel, S. F., & Marahiel, M. A. (1999). Review Peptidyl-prolyl cis-trans isomerases, a superfamily of ubiquitous folding catalysts. In *CMLS, Cell. Mol. Life Sci* (Vol. 55).

Guan, Y., Peace, C., Rudell, D., Verma, S., & Evans, K. (2015). QTLs detected for individual sugars and soluble solids content in apple. *Molecular Breeding*, *35*(6), 1–13. https://doi.org/10.1007/s11032-015-0334-1

Gutierrez, B. L., Arro, J., Zhong, G. Y., & Brown, S. K. (2018). Linkage and association analysis of dihydrochalcones phloridzin, sieboldin, and trilobatin in *Malus*. *Tree Genetics and Genomes*, *14*(6). https://doi.org/10.1007/s11295-018-1304-7

Gutierrez, B. L., Zhong, G. Y., & Brown, S. K. (2018a). Genetic diversity of dihydrochalcone content in *Malus* germplasm. *Genetic Resources and Crop Evolution*, *65*(5), 1485–1502. https://doi.org/10.1007/s10722-018-0632-7

Gutierrez, B. L., Zhong, G. Y., & Brown, S. K. (2018b). Increased phloridzin content associated with russeting in apple (*Malus domestica* (Suckow) Borkh.) fruit. *Genetic Resources and Crop Evolution*, *65*(8), 2135–2149. https://doi.org/10.1007/s10722-018-0679-5

Hampson, C. R., Sanford, K., & Cline, J. (2001). Preferences of Canadian consumers for apple fruit size. *Canadian Journal of Plant Science*, 165–167.

He, J., Zhao, X., Laroche, A., Lu, Z. X., Liu, H. K., & Li, Z. (2014). Genotyping-by-sequencing (GBS), An ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. In *Frontiers in Plant Science* (Vol. 5, Issue SEP). Frontiers Media S.A. https://doi.org/10.3389/fpls.2014.00484

Howard, N. P., Tillman, J., Vanderzande, S., & Luby, J. J. (2019). Correction: Genetics of zonal leaf chlorosis and genetic linkage to a major gene regulating skin anthocyanin production (MdMYB1) in the apple (*Malus* × *domestica*) cultivar Honeycrisp (PLoS ONE (2019) 14: 1 (e0210611) DOI: 10.1371/journal.pone.021061. *PLoS ONE*, *14*(3), 1–17. https://doi.org/10.1371/journal.pone.0213799

Jaywant, S. A., Singh, H., & Arif, K. M. (2022). Sensors and Instruments for Brix Measurement: A Review. In *Sensors* (Vol. 22, Issue 6). MDPI. https://doi.org/10.3390/s22062290

Juniper, B. E., & Mabberley, D. J. (2006). *The Story of the Apple*. Timber Press.

Kenis, K., Keulemans, J., & Davey, M. W. (2008). Identification and stability of QTLs for fruit quality traits in apple. *Tree Genetics & Genomes*, *4*(4), 647–661. https://doi.org/10.1007/s11295-008-0140-6

Khan, M. A., Olsen, K. M., Sovero, V., Kushad, M. M., & Korban, S. S. (2015). Fruit Quality Traits Have Played Critical Roles in Domestication of the Apple. *The Plant Genome*, *7*(3), 0. https://doi.org/10.3835/plantgenome2014.04.0018

Khan, S. A., Chibon, P.-Y., de Vos, R. C. H., Schipper, B. A., Walraven, E., Beekwilder, J., van Dijk, T., Finkers, R., Visser, R. G. F., van de Weg, E. W., Bovy, A., Cestaro, A., Velasco, R., Jacobsen, E., & Schouten, H. J. (2012). Genetic analysis of metabolites in apple fruits indicates an mQTL hotspot for phenolic compounds on linkage group 16. *Journal of Experimental Botany*, *63*(8), 2895–2908. https://doi.org/10.1093/jxb/err464

Kikuchi, T., Kasajima, I., Morita, M., & Yoshikawa, N. (2017). Practical DNA markers to estimate apple (*Malus* × *domestica* Borkh.) skin color, ethylene production and pathogen resistance. *J Hortic*, *4*(4), 211. https://doi.org/10.4172/2376-0354.1000211

Kouassi, A. B., Durel, C. E., Costa, F., Tartarini, S., van de Weg, E., Evans, K., Fernandez-Fernandez, F., Govan, C., Boudichevskaja, A., Dunemann, F., Antofie, A., Lateur, M., Stankiewicz-Kosyl, M., Soska, A., Tomala, K., Lewandowski, M., Rutkovski, K., Zurawicz, E., Guerra, W., & Laurens, F. (2009). Estimation of genetic parameters and prediction of breeding values for apple fruit-quality traits using pedigreed plant material in Europe. *Tree Genetics and Genomes*, *5*(4), 659–672. https://doi.org/10.1007/s11295-009-0217-x

Kumar, S., Chagné, D., Bink, M. C. A. M., Volz, R. K., & Whitworth, C. (2012). Genomic selection for fruit quality traits in apple (*Malus x domestica* Borkh.). *PLoS ONE*, *7*(5), 36674. https://doi.org/10.1371/journal.pone.0036674

Kumar, S., Richard K. Volz, David Chagné, & Susan Gardiner. (2014). Breeding for apple (Malus× domestica Borkh.) fruit quality traits in the genomics era. In *Genomics of Plant Genetic Resources: Volume 2. Crop productivity, food security and nutritional quality* (Vol. 2, pp. 387–416).

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods 2012 9:4*, *9*(4), 357–359. https://doi.org/10.1038/NMETH.1923

Larsen, B., Migicovsky, Z., Jeppesen, A. A., Gardner, K. M., Toldam-Andersen, T. B., Myles, S., Ørgaard, M., Petersen, M. A., & Pedersen, C. (2019). Genome-wide association studies in apple reveal loci for aroma volatiles, sugar composition, and harvest date. *The Plant Genome*, *12*(2), 180104. https://doi.org/10.3835/plantgenome2018.12.0104

Le Roux, P. M., Flachowsky, H., Hanke, M. V., Gessler, C., & Patocchi, A. (2012). Use of a transgenic early flowering approach in apple (Malus × domestica Borkh.) to introgress fire blight resistance from cultivar Evereste. *Molecular Breeding*, *30*(2), 857–874. https://doi.org/10.1007/s11032-011-9669-4

Legay, S., Guerriero, G., André, C., Guignard, C., Cocco, E., Charton, S., Boutry, M., Rowland, O., & Hausman, J. F. (2016). MdMyb93 is a regulator of suberin deposition in russeted apple fruit skins. *New Phytologist*, *212*(4), 977–991. https://doi.org/10.1111/nph.14170

Liebhard, R., Kellerhals, M., Pfammatter, W., Jertmini, M., & Gessler, C. (2003). Mapping quantitative physiological traits in apple (*Malus x domestica* Borkh.). *Plant Molecular Biology*, *52*(3), 511–526. https://doi.org/10.1023/A:1024886500979

Lo Piccolo, E., Landi, M., Massai, R., Remorini, D., Conte, G., & Guidi, L. (2019). Ancient apple cultivars from Garfagnana (Tuscany, Italy): A potential source for 'nutrafruit' production. *Food Chemistry*, *294*, 518–525. https://doi.org/10.1016/j.foodchem.2019.05.027

Ma, B., Chen, J., Zheng, H., Fang, T., Ogutu, C., Li, S., Han, Y., & Wu, B. (2015). Comparative assessment of sugar and malic acid composition in cultivated and wild apples. *Food Chemistry*, *172*, 86–91. https://doi.org/10.1016/J.FOODCHEM.2014.09.032

Mansoor, S., Sharma, V., Mir, M. A., Mir, J. I., un Nabi, S., Ahmed, N., Alkahtani, J., Alwahibi, M. S., & Masoodi, K. Z. (2020). Quantification of polyphenolic compounds and relative gene expression studies of phenylpropanoid pathway in apple (Malus domestica Borkh) in response to Venturia inaequalis infection. *Saudi Journal of Biological Sciences*, *27*(12), 3397–3404. https://doi.org/10.1016/j.sjbs.2020.09.007

McClure, K. A., Gardner, K. M., Douglas, G. M., Song, J., Forney, C. F., DeLong, J., Fan, L., Du, L., Toivonen, P. M. A., Somers, D. J., Rajcan, I., & Myles, S. (2018). A Genome-Wide Association Study of Apple Quality and Scab Resistance. *The Plant Genome*, *11*(1), 1–14. https://doi.org/10.3835/plantgenome2017.08.0075

McClure, K. A., Gong, Y., Song, J., Vinqvist-Tymchuk, M., Campbell Palmer, L., Fan, L., Burgher-MacLellan, K., Zhang, Z., Celton, J.-M., Forney, C. F., Migicovsky, Z., & Myles, S. (2019). Genome-wide association studies in apple reveal loci of large effect controlling apple polyphenols. *Horticulture Research*, *6*(1), 107. https://doi.org/10.1038/s41438-019-0190-y

McCord, P. (2018). *Advances in marker-assisted breeding of apples* (pp. 109–132). https://doi.org/10.19103/as.2017.0035.06

Migicovsky, Z., Gardner, K. M., Money, D., Sawler, J., Bloom, J. S., Moffett, P., Chao, C. T., Schwaninger, H., Fazio, G., Zhong, G.-Y., & Myles, S. (2016). Genome to Phenome Mapping in Apple Using Historical Data. *The Plant Genome*, *9*(2), plantgenome2015.11.0113. https://doi.org/10.3835/PLANTGENOME2015.11.0113

Migicovsky, Z., Gardner, K. M., Richards, C., Chao, C. T., Schwaninger, H. R., Fazio, G., Zhong, G.-Y., & Myles, S. (2021). Genomic consequences of apple improvement. *Horticulture Research*. https://doi.org/10.1038/s41438-020-00441-7

Mikulič Petkovšek, M., Štampar, F., & Veberič, R. (2009). Accumulation of phenolic compounds in apple in response to infection by the scab pathogen, *Venturia inaequalis*. *Physiological and Molecular Plant Pathology*, *74*(1), 60–67. https://doi.org/10.1016/j.pmpp.2009.09.003

Orcheski, B., Parker, R., & Brown, S. (2015). Pale green lethal disorder in apple (*Malus*) is caused by a mutation in the PHYLLO gene which is essential for phylloquinone (vitamin K 1) biosynthesis. *Tree Genetics & Genomes*, *11*(131). https://doi.org/10.1007/s11295-015-0956-9

Paniagua, C., Blanco-Portales, R., Barceló-Muñoz, M., García-Gago, J. A., Waldron, K. W., Quesada, M. A., Muñoz-Blanco, J., & Mercado, J. A. (2016). Antisense down-regulation of the strawberry β-galactosidase gene FaβGal4 increases cell wall galactose levels and reduces fruit softening. *Journal of Experimental Botany*, *67*(3), 619–631. https://doi.org/10.1093/jxb/erv462

Pascoal, A., Quirantes-Piné, R., Fernando, A. L., Alexopoulou, E., & Segura-Carretero, A. (2015). Phenolic composition and antioxidant activity of *Malus domestica* leaves. *Industrial Crops and Products*, *78*, 116–123. https://doi.org/10.1016/j.indcrop.2015.10.028

Pereira-Lorenzo, S., Fischer, M., Ramos-Cabrer, A. M., & Castro, I. (2018). Apple (*Malus* spp.) Breeding: Present and Future. In *Advances in Plant Breeding Strategies: Fruits* (pp. 3–29). Springer International Publishing. https://doi.org/10.1007/978-3-319-91944-7_1

Platten, J. D., Cobb, J. N., & Zantua, R. E. (2019). Criteria for evaluating molecular markers: Comprehensive quality metrics to improve marker-assisted selection. *PLoS ONE*, *14*(1), 1–20. https://doi.org/10.1371/journal.pone.0210529

Potts, S. M., Khan, M. A., Han, Y., Kushad, M. M., & Korban, S. S. (2014). Identification of Quantitative Trait Loci (QTLs) for Fruit Quality Traits in Apple. *Plant Molecular Biology Reporter*, *32*(1), 109–116. https://doi.org/10.1007/s11105-013-0634-8

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

Rymenants, M., van de Weg, E., Auwerkerken, A., De Wit, I., Czech, A., Nijland, B., Heuven, H., De Storme, N., & Keulemans, W. (2020). Detection of QTL for apple fruit acidity and sweetness using sensorial evaluation in multiple pedigreed full-sib families. *Tree Genetics and Genomes*, *16*(5). https://doi.org/10.1007/s11295-020-01466-8

Steyn, N. (2020). *Red colour improvement in apple fruit (Malus domestica Borkh .) By*. *December*.

Sun, H. H., Zhao, Y. B., Li, C. M., Chen, D. M., Wang, Y., Zhang, X. Z., & Han, Z. H. (2012). Identification of markers linked to major gene loci involved in determination of fruit shape index of apples (*Malus domestica*). *Euphytica*, *185*(2), 185–193. https://doi.org/10.1007/s10681-011-0515-x

Sun, J., Janisiewicz, W. J., Nichols, B., Jurick II, W. M., & Chen, P. (2017). Composition of phenolic compounds in wild apple with multiple resistance mechanisms against postharvest blue mold decay. *Postharvest Biology and Technology*, *127*, 68–75. https://doi.org/10.1016/j.postharvbio.2017.01.006

Tan, Y., & Wang, L. (2020). MpDGK2, a Novel Diacylglycerol Kinase from Malus prunifolia, Confers Drought Stress Tolerance in Transgenic Arabidopsis. *Plant Molecular Biology Reporter*, 1–9. https://doi.org/10.1007/s11105-020-01209-y

Tsao, R., Yang, R., Young, J. C., & Zhu, H. (2003). Polyphenolic profiles in eight apple cultivars using high-performance liquid chromatography (HPLC). *Journal of Agricultural and Food Chemistry*, *51*(21), 6347–6353. https://doi.org/10.1021/jf0346298

Verdu, C. F., Guyot, S., Childebrand, N., Bahut, M., Celton, J. M., Gaillard, S., Lasserre-Zuber, P., Troggio, M., Guilet, D., & Laurens, F. (2014). QTL analysis and candidate gene mapping for the polyphenol content in cider apple. *PLoS ONE*, *9*(10). https://doi.org/10.1371/journal.pone.0107103

Wang, S., Liang, D., Shi, S., Ma, F., Shu, H., & Wang, R. (2011). Isolation and Characterization of a Novel Drought Responsive Gene Encoding a Glycine-rich RNA-binding Protein in Malus prunifolia (Willd.) Borkh. *Plant Molecular Biology Reporter*, *29*(1), 125–134. https://doi.org/10.1007/s11105-010-0221-1

Waterhouse, A. (2012). Folin-Ciocalteau Micro Method for Total Phenol in Wine Department of Viticulture & Enology University of California, Davis. Available in: http://waterhouse. ucdavis. edu/phenol/folinmicro. htm, accessed April.

Watts, S., Migicovsky, Z., Mcclure, K. A., Yu, C. H. J., Amyotte, B., Baker, T., Bowlby, D., Butler, L., Donald, R., Fan, L., Fillmore, S., Flewelling, J., Gardner, K., Hodges, M., Hughes, T., Jagadeesan, V., Lewis, N., Macdonell, E., Macvicar, L., … Myles, S.

(2021). *Quantifying apple diversity : A phenomic characterization of Canada ' s Apple Biodiversity Collection*. *April*, 1–14. https://doi.org/10.1002/ppp3.10211

Wojdyło, A., Oszmiański, J., & Laskowski, P. (2008). Polyphenolic compounds and antioxidant activity of new and old apple varieties. *Journal of Agricultural and Food Chemistry*, *56*(15), 6520–6530. https://doi.org/10.1021/jf800510j

Xu, Dr. H., Blatt, Dr. S., & Ediger, Ms. D. (2022). Tools for climate resilience in tree fruit I: Large-dwarfing rootstocks can alleviate sunburn damage in 'Buckeye Gala' apple. *Https://Doi.Org/10.1139/CJPS-2022-0080*. https://doi.org/10.1139/CJPS-2022-0080

Yao, J. L., Xu, J., Cornille, A., Tomes, S., Karunairetnam, S., Luo, Z., Bassett, H., Whitworth, C., Rees-George, J., Ranatunga, C., Snirc, A., Crowhurst, R., De Silva, N., Warren, B., Deng, C., Kumar, S., Chagné, D., Bus, V. G. M., Volz, R. K., … Gleave, A. P. (2015). A microRNA allele that emerged prior to apple domestication may underlie fruit size evolution. *Plant Journal*, *84*(2), 417–427. https://doi.org/10.1111/tpj.13021

Zhang, C., & Hao, Y. J. (2020). Advances in Genomic, Transcriptomic, and Metabolomic Analyses of Fruit Quality in Fruit Crops. *Horticultural Plant Journal*, *6*(6), 361–371. https://doi.org/10.1016/j.hpj.2020.11.001

Zhang, L., Hu, J., Han, X., Li, J., Gao, Y., Richards, C. M., Zhang, C., Tian, Y., Liu, G., Gul, H., Wang, D., Tian, Y., Yang, C., Meng, M., Yuan, G., Kang, G., Wu, Y., Wang, K., Zhang, H., … Cong, P. (2019). A high-quality apple genome assembly reveals the association of a retrotransposon and red fruit colour. *Nature Communications*, *10*(1494). https://doi.org/10.1038/s41467-019-09518-x

CHAPTER 3

APPLE FRUIT STORAGE DYNAMICS

INTRODUCTION

Ensuring that fruit quality is maintained during fruit storage is an integral part of the food production system. Fruit quality has been shown to be a major determinate of consumers' willingness to eat fruit, with minor defects resulting in low willingness to consume which can result in an estimate of 45% to 55% of all fruit and vegetables being discarded along the production system (L. C. Argenta et al., 2021; Jaeger et al., 2018). Apple is a unique case amongst fruit crops because harvested fruit of certain cultivars can be stored up to 12 months before being sold to consumers without significant loss of quality (Gapper et al., 2023). Consequently, apple has also been reported to have more documented storage disorders than any other commodity, reflecting the wide planting regions and various storage schemes from open air to controlled atmospheric storage (Watkins & Mattheis, 2019). Apple fruit are also marketed by cultivar, meaning that there is slow acceptance of new cultivars that might have improved storability as the other important consumer-based traits (Chapter Acidity and Fruit Quality) override the desire for longer storage capacity.

As a result of this market pressure, significant research has been done to optimize storage climate conditions, fruit quality, and conditioning, in order to lengthen storage time and improve storage quality. A small proportion of studies have approached the problem through cultivar specific research examining the underlying genetics of storage capacity (Cheng et al., 2022; Costa et al., 2005; Harada et al., 2000; Oraguzie et al., 2004b; Zhu & Barritt, 2008).

Storage time in commercial production is recognized as a complex trait that is composed of many factors including human impacts and genetic and physiological characteristics. Traits include mineral composition of the fruit at harvest, handling during processing, presence of pathogens in storage, storage equipment, storage conditions, storage methods, fruit texture and genetics present in the stored fruit (Vicente et al., 2007). This complexity has resulted in much of the research on storage of fruit pertaining to the human controlled inputs such as storage temperature and pretreatments to prevent shock and disorders (Watkins & Mattheis, 2019).

Variation in resistance to storage disorders vary by cultivar and species. The most common post-harvest pathogens are Blue Mold *Penicillium expansum,* and Bitter Rot *Colletotrichum acutatum*, *resistance* to these post-harvest pathogens has been identified descending from wild species which have not been used in commercial cultivar breeding (Jurick et al., 2011). Blue mold is regarded as the most prevalent post-harvest pathogen present in apple storage and processing (Cappellini, 1987). In a cross of *M. domestica* 'Royal Gala' x *M.sieversii* PI613981 (GMAL4593) two QTLs associated with blue mold resistance have been identified: one on Chromosome 10 descending from 'Royal Gala', and one on chromosome 3, primarily descending from *M. sieversii*. Genetic resistances to other post-harvest pathogens have yet to be identified (Norelli et al., 2017a).

Watercore is a physiological disorder that is associated with a softening of fruit flesh and a "water-soaked" appearance in the flesh. This disorder initially appears around the vascular bundles but can extend from the core to the skin's surface (Marlow & Loescher, 2011). Watercore has been correlated with High K and Mg and

low Ca concentration in fruit (Sharples, 1967). Fruit that display watercore symptoms have flooded intercellular spaces, low levels of reducing sugars (e.g. glucose and fructose) and high concentrations of sorbitol (Bowen & Watkins, 1997; Yamada et al., 2012). Intercellular airspace volume is decreased in watercore tissues, ethanol and acetaldehyde can accumulate, and the Co2 concentration of internal tissues can increase (L. Argenta et al., 2002). In some markets it is prized because of the flavor and sweetness that it confers, for example the cultivar 'Fuji', which is prone to watercore without texture degradation. In a resequencing and haplotype study of Fuji using 115 'Fuji' relations, a QTL that associated with degree of watercore was identified on chromosome 14 in the haplotype derived from Fuji's parent 'Delicious' (Kunihisa et al., 2016). The genetic architecture of watercore remains an area of active research.

Genes associated with fruit texture, crispness, storability, and harvest date have been identified in mapping populations and Genome Wide Association Studies (GWAS). Using GWAS a SNP associated with fruit firmness after storage, was identified on chromosome 10 using 172 cultivars commonly grown in in Nova Scotia (McClure et al., 2018). Also located on chromosome 10 is another SNP identified using GWAS that associated with fruit 'crisp' texture in apple by using 233 accessions (Di Guardo et al., 2017). A SNP that correlates with harvest date was also able to be identified using historical data and a GWAS was used to identify a haplotype that associated with harvest data for apple on chromosome 3 (Larsen et al., 2019).

Ethylene has been identified as the major plant hormone associated with ripening and changes in fruit quality during storage. The pathway of synthesis and

effects of ethylene on fruit has been well documented (Lau et al., 1986). In the initial steps of the pathway, methionine is converted into S-adenosyl-L-methionine (SAM) by *SAM synthetase* using ATP (D. 0 Adams et al., 1977). SAM is converted into 1-aminocycloproplane-1-carboxylic acid (ACC) and 5'-methylthioadenosie (MTA) by *ACC-synthase* (*ACS*) (D. 0 Adams et al., 1979; Boller et al., 1979; Murr & Yang, 1975). MTA is recycled back to methionine by the Yang Cycle to avoid depletion during high ethylene production times (Murr & Yang, 1975). In the second and final step, ethylene is released from ACC by *ACC-oxidase* (*ACO*), requiring molecular oxygen (Burg & Burg, n.d.; Hamilton et al., n.d.; Ververidis & John, 1991). **Figure 3.1**

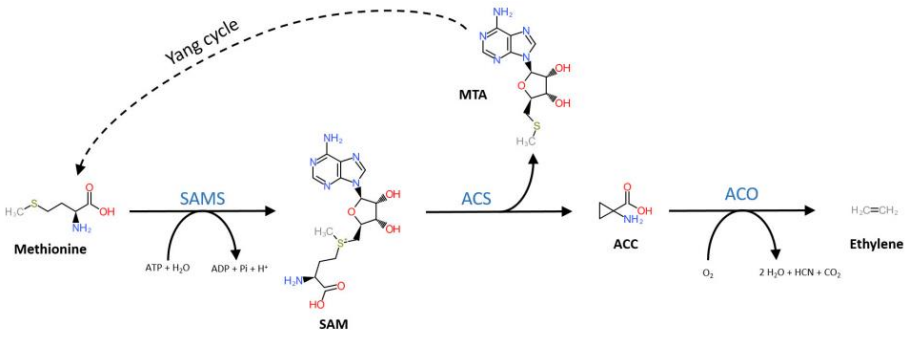


Figure 3.1 Adapted from Houben & Van de Poel, (2019) The ethylene biosynthesis pathway.

MdACS1 encodes a *-aminocyclopropane-1-carboxylic acid synthase* (ACS) gene and has been localized to chromosome 15 in *Malus*. This gene is understood to be the rate-limiting-step in ethylene synthesis and generates 1-aminocyclopropane-1-carboxylic acid (ACC) from S-adenosyl-L-methionine (SAM) early in the ethylene biosynthesis pathway (Houben & Van de Poel, 2019). Two alleles of ACS1 have been identified and verified (ACS1-1 and ACS1-2). These alleles have been observed in

three arrangements; ACS1-1/1, 1-1/2 and 1-2/2 correlating with high, medium, and low ethylene production (Costa et al., 2005; Harada et al., 2000; Oraguzie et al., 2004; Sunako et al., 1999).

MdACO1 encodes a 1-aminocyclopropane-1-carboxylate oxidase which is one of the final steps in the conversion of 1-aminocyclopropane-1-carboxylic acid (ACC) into ethylene. This gene has been located on Chromosome 10 in linkage studies using cultivars 'Prima' × 'Fiesta' and 'Fuji' × 'Mondial Gala' within the confidence region of a previously identified firmness QTL (King et al., 2000). Two markers for *ACO1* have been developed and verified (*ACO1-1* and *ACO1-2*) with *ACO1-1* correlating with higher ethylene production and *ACO1-2* with lower ethylene production(Costa et al., 2005; Zhu & Barritt, 2008) Using reciprocal crosses of the cultivars 'Golden Delicious' x 'James Grieve', 'Golden Delicious' x 'Jonathan', 'Cox's Orange Pippin' x 'Golden Delicious', and 'Cox's Orange Pippin' x 'Jonathan', three allelic forms of *ACO1* were identified (Lyzhin & Savelyeva, 2020). Two of these forms (*a* and *c*) correlate with the previously identified *ACO1-1*, with a single base pair difference while allele *b* correlates to *ACO-2* which confers lower ethylene production. The *a* and *c* allelic forms have not been tested to determine if they have any difference in physiological effect.

A key gene of at the interface between ethylene and storability is *MdPG1*. *MdPG1* encodes an *ethylene-dependent polyglacturonase* gene. *PG* function is known to be dependent on the plant hormone ethylene, with a base level of ethylene needed to induce *PG* transcription, and the accumulation of *PG* is directly regulated by the level of ethylene present (Brummell & Harpster, 2001; Costa et al., 2010).

Polygalacuronase (*PG*) is well established as one of the genes that is active in the break-down of pectin by catalyzing the hydrolytic cleavage of a α(1-4) galacturonan bond in the pectin molecule. The breakdown of pectin results in a fruit that is softer and less firm in storage, as well as cell degradation resulting in shorter storage times (Longhi et al., 2013). *PG1* has been located on Chromosome 10 in *Malus* near other significant gene associated with storage ability, *MdACO1* (discussed above), and has been associated with a loss in firmness during storage as well as differences in firmness at harvest (Bassil et al., 2019; Costa et al., 2010; Longhi et al., 2013). Markers for *PG1* have been made that differentiate between the two alleles known, and an allelic dosage model has been put forward to explain the genotypes of *PG1* with Md-PG1$_{SNP}$-G allele conferring lower firmness during storage (Longhi et al., 2013).

Of these three genes *ACO1* and *ACS1* have been primarily associated with storage time of apples while *PG1* has been more closely related to texture and firmness changes during storage, which influences storability but isn't the only characteristic determining it (Vicente et al., 2007). Both *ACO1* and *ACS1* show low transcription at time of harvest, with their transcription levels increasing over storage time, and increasing due to adverse or injurious storage conditions (Tian et al., 2002). *ACO1* and *ACS1* both appear to be a part of a feedback loop of producing, and being induced by, ethylene. In closely related pear (*Pyrus ussuriensis*) similar orthologs of *ACS1* and *ACO1* in apple were induced by the treatment of synthetic ethylene (Ethephon) and repressed by the application of the ethylene uptake inhibitor 1-MCP (Yuan et al., 2020).

While plants that produce fleshy fruit vary widely in their appearance and ripening types, the controls of ripening appear to be well conserved across the plant kingdom. *MdACO1* and *MdACS* shows high similarity to those found in tomato (Holdsworth et al., 1988), melon (Lasserre et al., 1996), banana (L6pez-G6mez 'aa et al., 1997), and peach (Ruperti et al., 2001) all of which have also been reported to influence storage and ripening. In Guava (*Psidium guajava* L.) the combination of *ACS* and *ACO* genotypes for low ethylene has been shown to be responsible for the non-climacteric cultivar 'Jen-Ju Bar' which remains firm in long storage (K.-E. Chen et al., 2023). Again, in Guava *ACS* appears to be primarily responsible for the phenotype of long vs short storing, with *ACO* showing effects within genotypes but not accounting for large amounts of ethylene variation. These three genes' (*ACO*, *ACS*, and *PG*) effects on storability and texture are variable in effect depending on measurement technique and population, so there is still a lot of work needed to fully understand the genetic controls of storage in apple.

Objectives

The objective of this research is to identify underlying genetic causes of segregating storage times using a biparental interspecific cross of an advanced selection of *Malus domestica* (hereafter NY Co1) in the Cornell Apple breeding program and *Malus prunifolia* PI 589816 (hereafter 589816).

NY Co1 and 589816 have contrasting storage time with 589816 presenting a short-store phenotype, losing fruit firmness after less than three months in common storage of 4 °C, while NYCo1 remaining sound up to 12 months after harvest in the same storage conditions. This project offers the opportunity to examine the role of known

genes in the ethylene biosynthesis pathway, as well as the effects that ethylene has on fruit during storage. This also adds context to our understanding of interspecific apple breeding by examining genes that have primarily been fixed in breeding populations.

MATERIALS AND METHODS

Plant Material

This research was completed using a bi-parental, full sibling population of *M. domestica* (NYCo1) x *M. prunifolia* (PI 589816) grown on their own roots at the Lucy Research Farm located at Cornell Agritech in Geneva, NY (42.868536, -77.044668). The cross was made in 2013 with 600 seeds collected, stratified, and planted once the radical emerged. 150 germinating pale green lethal, which is expected in material heterozygous for the disorder, resulting in a population of 445 individuals (Orcheski et al., 2015). These seedlings were grown in the greenhouse for 2 years before being planted outside. NYCo1 is heterozygous for *Co,* the columnar gene, and produces large (>80 mm) yellow fruit with a pink blush where exposed to the sun with minimal russet outside of the stem well. *M. prunifolia* PI 589516 has standard architecture and produced small (<15 mm) yellow fruit, primarily with a waxy cuticle, but occasionally displays partial to full russet.

Fruit was harvested starting at the beginning of September and extending to the end of October each year. Ripeness was determined by evaluating the ease of abscission from the tree as well as ground color change. At least 30 fruit were randomly harvested from each tree, harvesting from different regions of the tree canopy. The fruit was placed into paper bags to allow gas exchange and labeled with the row number and tree number; this number was then used to track the fruit through

data collection. The bags of fruit were then stored at 4°C in a commercial cooler with limited light exposure, under ambient atmospheric gases.

Storage Observations

30 fruit were harvested from trees between September 1 to October 31, and placed in paper bags. For storage evaluation, at least 30 fruit from each accession were placed in a commercial cold storage room (4°C) in paper bags to allow gas exchange. No other storage treatments were applied. Fruit soundness, determined by applying thumb pressure to fruit to evaluate firmness, was evaluated periodically during the storage period, noting when over 50% of stored fruit exhibited the breakdown phenotype. The percentage of fruit in each sample that had senesced, defined by the appearance of discoloration, cracking, peeling, or otherwise rotting, was estimated visually or from counting proportion of fruit. Post-harvest pathogens present in each sample were noted, as well as any physiological defects such as cracking, water core, rots, and senescence of mesocarp tissue.

Storage Genotyping

Two alleles previously associated with storability *ACO1* and *ACS1* were evaluated in the parents using Polymerase chain reaction (PCR). PCR was performed using the protocol from Zhu & Barritt, (2008); in a final mix of 25µl containing 50 ng of template DNA, 0.25 mM of each deoxyribonucleotide triphosphate(dNTP), 1 mM $MgCl_2$, 0.2 µM of each primer (forward and reverse), 2.5µl 10 × PCR buffer and 1 U of GoTaq HotStart polymerase (New England Biolabs, Ipswich, MA). The thermal cycler (MastercyclerVREP; Eppendorf, Hamburg, Germany) performed the following protocol: 94 °C for 2 min, 35 cycles of 94 °C for 45 s, 58 °C for *ACS1* primers and 65

°C for *ACO1* primers for 45 s, 72 °C for 2 min, followed by a final extension at 72 °C for 7 min. The PCR products were separated on a 2% agarose gel and band visualizing using SYBR florescence tagging. The sequence of the *ACS1* primers followed Harada et al., (2000) and the *ACO1* primers by Costa et al., (2005). *ACO1* has expected fragments sizes around 525 (*ACO1-1*) and 587 bp (*ACO1-2*), with the 525 bp fragment associating with lower ethylene production, longer storage life, and maintained fruit firmness. *ACS1* has expected fragment sizes of 489 bp (*ACS1-1*) and 655 bp (*ACS1-2*) with the smaller band associated with higher ethylene production (Sunako et al., 1999).

Genetic Association

Marker-trait association analyses between 51442 Genotype-by-Sequencing markers and measured traits were performed using TASSEL v5 software (Bradbury et al. 2007). Using the Q method, a general lineal model (GLM) was done using numeric data joined to genotype data and principal component analysis (PCA) as population membership estimates, along with *Ma1* and *Ma3* genotypes, and tree architecture as covariates (*Ma1Ma1* = 1, *Ma1ma1* =0; *Ma3ma3* = 1, *ma3ma3* = 0; *Coco* = 1, *coco* =0). Data was exported from TASSEL and results were visualized using the 'qqman' R package (D. Turner, 2018). Significant SNPs were determined using the Bonferroni-corrected threshold for GWAS (Log of Odds > 7).

Ethylene Measurements

In 2021 a subset of the population consisting of 60 individuals (**Table 3.1)**, including the two parents, were measured for internal ethylene concentration (IEC). 10 fruit from each accession were measured for IEC on day of harvest and after storing at

20 C for 7 days after harvest to check for changes in ethylene production. IEC was determined by withdrawing a 1 mL gas sample from the core of each apple using a 1 mL plastic syringe. Gas samples were analyzed using a Hewlett Packard 5890 series II (Hewlett Packard Co., Wilmington, Del.) equipped with a stainless-steel column packed with 60/80 mesh alumina F-1 (2-m × 2-mm, i.d.) and an FID was used.

Blue Mold Resistance Evaluation

In 2021 the same subset of the population used for ethylene measurement were also used to evaluate Blue Mold resistance. Each sample in the subset consisted of groups of ten apples per accession. The fruit were taken at harvest and standardized for ripeness using Starch Iodine values of 8 (complete starch breakdown). The fruit were then punctured once using a nail to a depth and diameter of 8 mm using a self-made wounding device (a nail driven through a piece of wood). The fruit were then inoculated via this wound with a *P. exansum* suspension of strain R19 with 1x10^5 spores mL^{-1}. Fruit were then placed into a plastic Rubbermaid container with damp paper towels to ensure humidity. The fruit were checked daily starting 4 days after inoculation and presence of infection and time to complete infection were noted, as well as presence of sporulation. Storage temperature was kept at room temperature approximately 20°C for the duration of the experiment.

RNAseq

To identify and confirm genes that are active during storage and could be causal genes underlying the segregation of storage times, a transcriptional profiling experiment was designed, contrasting long and short storing progeny, as well as the parental genotypes of the mapping population. Fruit peel samples were taken at

harvest and following a time series design. Using an 8 mm circular leaf punch 1-2 peel disks were removed from each fruit on the day of harvest. The peel tissue was then separated from the fruit flesh and immediately placed in a 2 mL centrifuge tube and dropped into liquid Nitrogen, until being placed in – 80 for storage. This collection was repeated with three collections occurring in the short storing individuals, and five collections in the long storing individuals (**Figure 3.2**) These time series sample points were chosen to match the physiological time points of the senescing fruit in the two classes as well as provide more detail of the physiological processes in the long storing individuals.

Frozen tissue was pulverized in a tissue homogenizer (miniG Genogrinder, SPEX) and total mRNA was isolated using standard protocols (Sigma Spectrum plant RNA, Sigma Aldrich). Extracted RNA was evaluated for quantity and quality using a nanodrop (Nano-UV Vis Spectrophotometer)

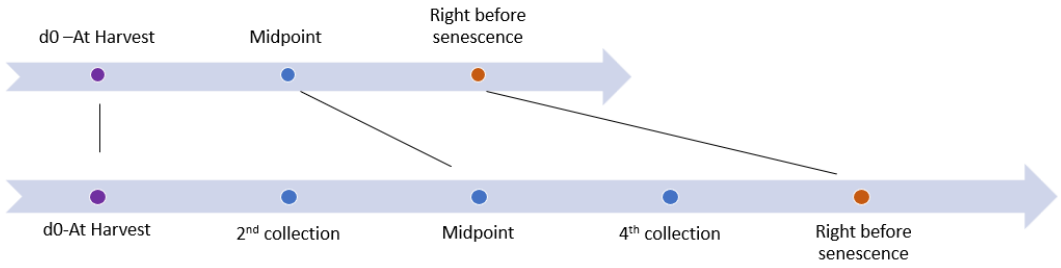


Figure 3.2 Depiction of RNAseq sample collection timepoints

Extracts of total RNAs (n=92) were submitted to the Cornell BioHPC genomics facility where 3' RNA libraries were constructed (Tandonnet & Torres, 2017). RNA was sequenced on a NextSeq500 as 85bp single end reads. Barcodes and poor-quality bases were removed from reads using Trimmomatic (Bolger et al., 2014) removing the first and final 10 bp, reducing the read length from 85 bp to 65 bp.

Following trimming, sample reads were evaluated for quality using fastQC (Andrews, 2010). Cleaned reads were then aligned to the golden delicious reference genome and reference transcriptome using STAR (Dobin et al., 2013) with the following presets: Maximum number of multiple alignments allowed = 2, Maximum number of mismatches allowed in a single alignment = 10. The RNA count matrix was then evaluated using DESeq2 (Love et al., 2014) Metadata consisted of "group" (long/short), "name" (accession number), "time" (in days after harvest), "rep" (biological rep), and "TimeG" (standardized time point 0-1 as a ratio of total storage time). DeSeq2 model was outlined as ~group + group:TimeG.

The false discovery rate (FDR) was set to 1% to determine the threshold of *P*-value by manipulating the FDR. *P*<001 and the absolute value of log2Ratio > 1.5 were used as the threshold to determine significance of the gene expression difference (Audic & Claverie, 1997).

Genes associated with the ethylene biosynthesis and reception pathway were identified by manually searching the GDDH13 genome (https://www.rosaceae.org/). Gene lists were made, and genes identified to be differentially expressed were plotted as eigengene expression patterns over storage time.

WGCNA

A weighted gene co-expression network analysis (WGCNA) done in R (Langfelder & Horvath, 2008) was used to identify gene modules for pathway enrichment analysis using a signed type of topological overlap matrix (TOM), soft-thresholding power = 12, a minimal module size of 50, and a branch merge cut height of 0.25. Gene modules were manually curated to identify modules that correlated with

expected changes in ethylene and synesis focusing on modules that were upregulated in short verse long storing groups or the inverse.

RESULTS

Storage Observations

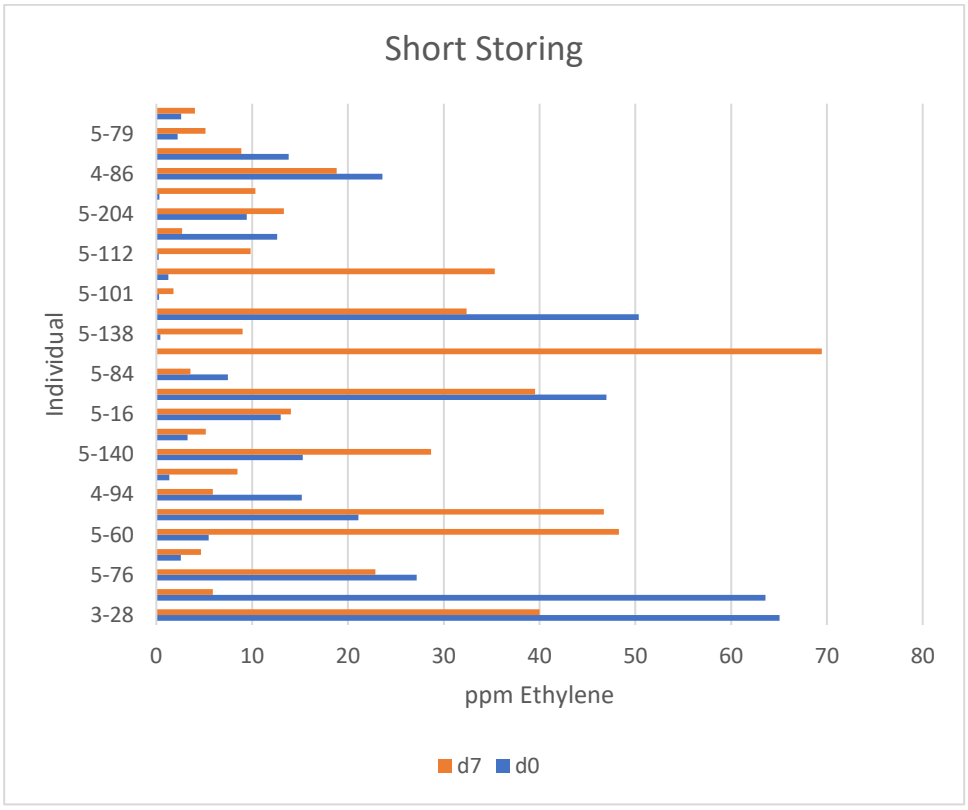
Storage phenotypes were recorded in 2020, 2021, and 2022. The number of fruiting individuals in each year varied; in 2020; 292 individuals, 2021; 345 individuals, 2022; 309. Samples were sound when put into storage and were checked periodically to remove and note fully rotten samples to not affect neighboring samples. After 9 months a final observation was taken noting all samples that remained sound and all samples that had senesced. In 2020, 129 samples remained sound, with 156 samples senescing. 2021, 179 samples remained sound, with 103 samples senescing. 2022, 196 samples remained sound, with 139 samples senescing. Samples that contained post-harvest pathogens or disorders were also noted including blue and white mold, fruit splitting, skin peeling, black rot, and disorders that were unable to be easily identified. The segregation ratios for each year of sound: senescent work out to 2020; 129: 156, 1:1.21, 2021; 179:103 0.58:1, and 2022; 196:139, 1:1.41.

Storage Genotyping

The parents of the population were evaluated for ACO and ACS. The *M. domestica* parent, NYCo1, is homozygous for the low ethylene allelic forms of both ACO and ACS. The *prunifolia* parent is homozygous at ACS but is heterzygeous at ACO, with one low ethylene and one high ethylene gene. Given the parental genotypes, it is expected that the progeny should segregate at a 1:1 ratio for long vs short storing individuals.

Ethylene Measurements

There was not a clear pattern noted by measuring internal ethylene in the subset of the population. 60 individuals were measured, 33 were in the “long storing” group and 26 were in the “short storing” group. The results of the ethylene measurements in 2021 are depicted in **Figure 3.3**. In the short storing group 16 of the 26 accessions increased ethylene concentration from day 0 to day 7, while 10 accessions decreased. In the long storing group 17 accessions increased ethylene concentration, while 16 decreased concentrations from day 0 to day 7.


Short Storing
Individual
5-79
4-86
5-204
5-112
5-101
5-138
5-84
5-16
5-140
4-94
5-60
5-76
3-28
0
10
20
30
40
50
60
70
80
ppm Ethylene
d7
d0

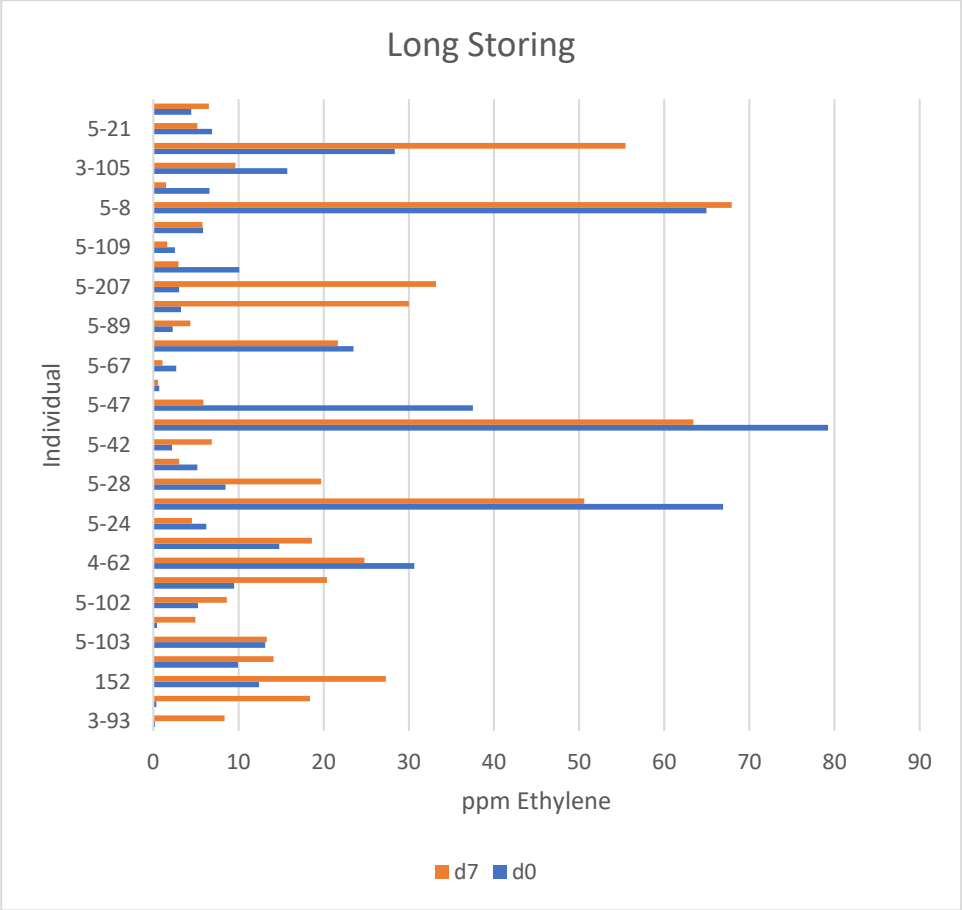


Figure 3.3 Line graph of change in internal ethylene concentration (ppm) from day of harvest (d0) to seven days after harvest (d7). 5-113 was removed from chart due to its abberantly high measurment.

Assessment of Blue Mold Resistance

There was no resistance to blue mold observed in the subset of the population screened, however small differences in time to sporulation were observed. In all fruit regardless of storage group blue mold infection took place before the first observation point at d4 after inoculation, by d7 sporulation was seen in all samples.

Table 3.1 Subset of population used in Internal Ethylene measurements, and Blue Mold resistance testing "Tree" designates the individual location of the tree in the field, "d0" is the internal ethylene content at day of harvest in ppm ethylene, "d7" is the IEC on day 7, and "Group" designates if they are considered short or long storing types.

Tree	d0	d7	Group
3-28	65.05736	40.018	Short
4-37	63.58758	5.9066	Short
5-76	27.1724	22.8776	Short
5-200	2.563818	4.6648	Short
5-60	5.4602	48.2948	Short
5-2	21.0918	46.7234	Short
4-94	15.184	5.9034	Short
5-170	1.372545	8.4652	Short
5-140	15.2688	28.6674	Short
5-64	3.261636	5.1746	Short
5-16	12.97133	14.0372	Short
4-40	46.98436	39.54782	Short
5-84	7.4552	3.570727	Short
5-99		69.47691	Short
5-138	0.41	9	Short
5-19	50.3728	32.3938	Short
5-101	0.293	1.793	Short
5-53	1.2586	35.31267	Short
5-112	0.262	9.8276	Short
5-191	12.6114	2.692	Short
5-204	9.4258	13.32825	Short
5-15	0.329	10.3314	Short
4-86	23.5968	18.8286	Short
3-32	13.8162	8.877	Short
5-79	2.2284	5.1448	Short
Prunifolia	2.579	4.043111	Short
3-93	0.192425	8.3752	Long
5-93	0.350771	18.4112	Long
152	12.4176	27.3178	Long
5-32	9.9396	14.1096	Long
5-103	13.1322	13.3124	Long

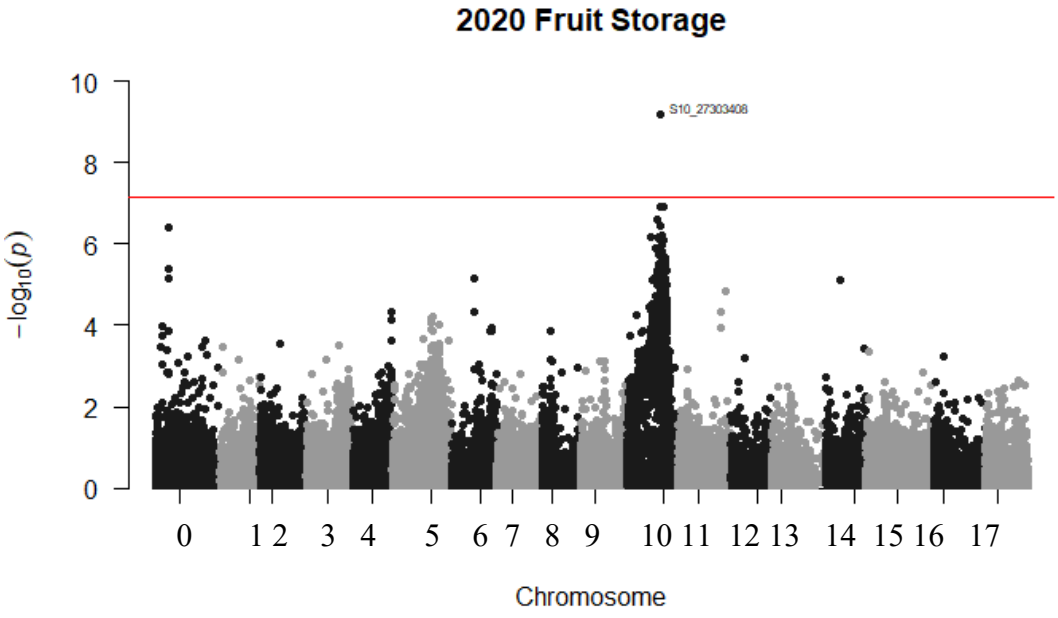
5-35	0.4524	4.9278	Long
5-102	5.2326	8.6238	Long
5-7	9.5072	20.3966	Long
4-62	30.6598	24.7716	Long
4-58	14.7808	18.6178	Long
5-24	6.2096	4.542	Long
3-47	66.90345	50.59123	Long
5-28	8.4682	19.73231	Long
5-45	5.176545	3.042182	Long
5-113	220.555	169.6988	Long
5-42	2.205	6.8668	Long
5-38	79.2498	63.4294	Long
5-47	37.5438	5.8916	Long
5-18	0.7154	0.556	Long
5-67	2.6958	1.0678	Long
5-86	23.5058	21.6638	Long
5-89	2.27875	4.3332	Long
5-29	3.2496	29.993	Long
5-207	3.043	33.19333	Long
4-1	10.1094	2.958667	Long
5-109	2.5406	1.639778	Long
4-74	5.863	5.7566	Long
5-8	64.94145	67.91782	Long
4-78	6.588	1.4932	Long
3-105	15.7178	9.6612	Long
5-1	28.3718	55.459	Long
5-21	6.9158	5.1622	Long
5-110	4.4488	6.5166	Long

Genetic Association

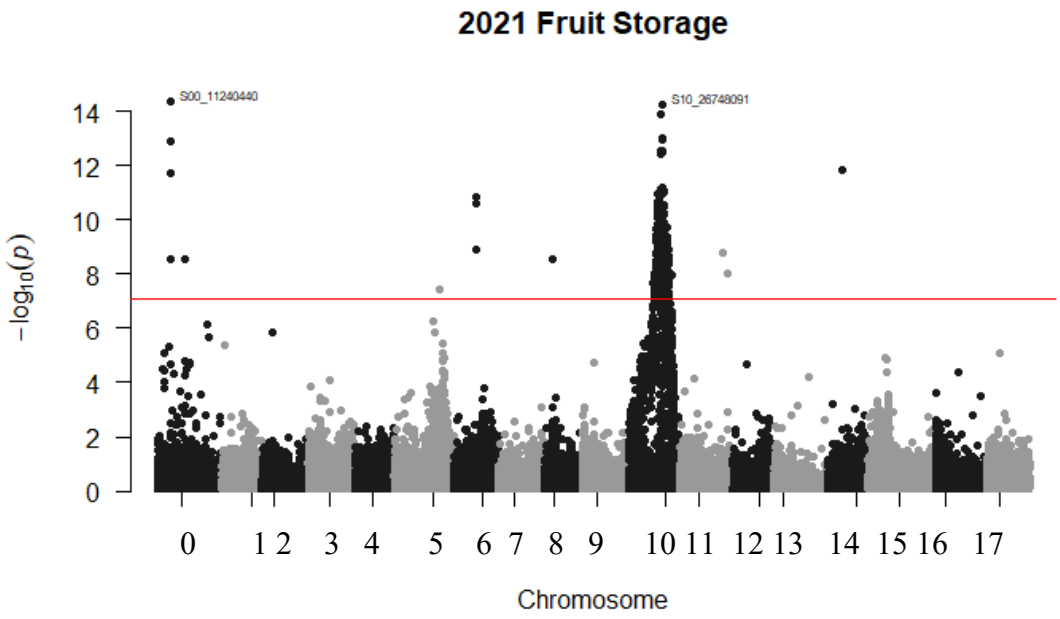
Significant SNP associations for increased storage time were found on Chromosome 10 all three years collected. In 2020 the SNP was located at 27303408, in 2021 26748091, in 2022 28780687. Located near this SNP are three genes with start and stop positions of 28068311 – 28070024, 28095661 – 28097111, and 28108962 – 28110702 (**Figure 3.4**). All three of these genes are homologs of 1-aminocyclopropane-1-carboxylate oxidase in *Arabidopsis* and two are predicted 1-

aminocyclopropane-1-carboxylate oxidase genes in *M. domestica* with the latter one being a predicted 1-aminocyclopropane-1-carboxylate oxidase gene in *Pyrus x bretschneideri*. No other SNP regions associated with storage in any of the years. While water core was observed each year of study, it's prevalence in the population was too low (50 with watercore in 2019, 51 in 2020, and 46 in 2021) to identify any SNPs that associated with its development.

A



B


2021 Fruit Storage
S00_11240440
S10_26748091
−log10(p)
0
2
4
6
8
10
12
14
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Chromosome

C

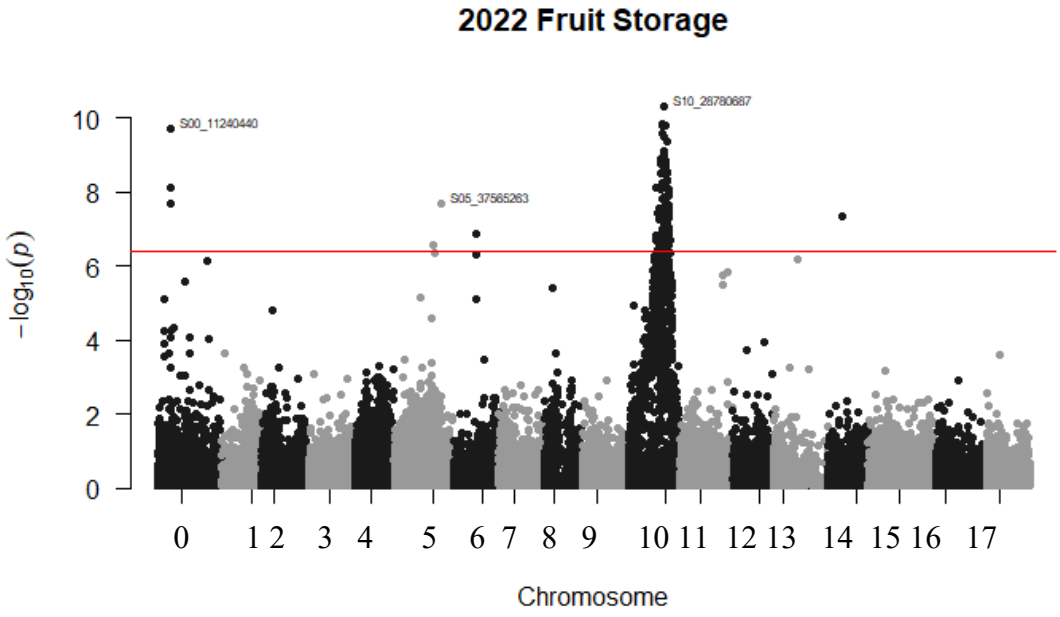

2022 Fruit Storage
S00_11240440
S05_37565263
S10_26780687
−log10(p)
0
2
4
6
8
10
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Chromosome

Figure 3.4 Manhattan Plots showing genetic associations determined by General Linear Model of Fruit Storage in **A**. 2020, **B**. 2021, **C**. 2022. Significant LOD is marked at 7 with a red line.

RNAseq

After sequencing the extracted RNAs (n=92), reads were filtered to remove low quality and poor alignment reads using Trimmomatic and STAR. The resulting filtering step reduced library number to 80 samples. A further two samples were removed after DeSeq2 as outliers to the data distribution. Resulting in 78 samples used for further profiling. There were 39 samples in the long storing group and 39 in the short storing group.

The average number of reads in the quality filtered samples was over 3 million, with an average percentage of uniquely mapped reads of 81.43%. After alignment to the GDDH13 annotation, 30485 genes were observed in this experiment out of a total of 52741 total annotations, representing 57.8% of total annotated genes being captured from the peel tissue. After filtering for lowly-expressed genes (sum of gene counts across all samples $\leq$ total samples = 78) there were 24195 highly expressed genes.

A PCA was used to examine the underlying gene expression structure of the dataset, 17.4% of variation was explained by PC1, 11.4% by PC2 and 7.1% by PC3, resulting in 35.9% variance explained by these three components. There was clear physical separation between groups (long, short storing) along PC2, as well as by collection time along PC1, with earlier times clustering on the right side of the PCA and later collections in both groups clustering on the left (**Figure3.5**).

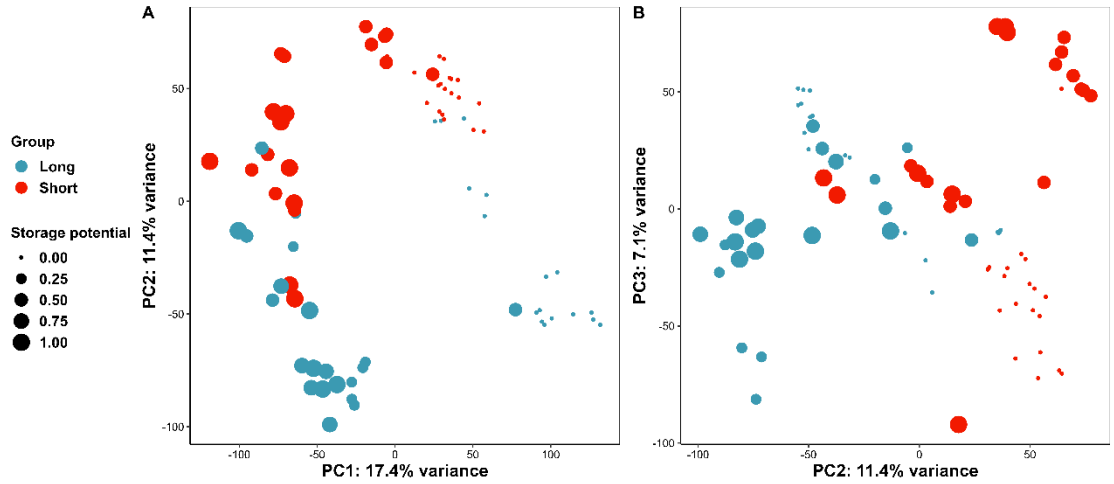


Figure 3.5 Principal Component Analysis of RNA samples. PC1 vs PC2 on the left and PC2 vs PC3 on the right. Long storing samples are coded as blue, short storing samples are coded as red, percentage of total storage time is given by the dot size.

WGCNA Cluster identification and Gene Ontology (GO)

Based on average hierarchical clustering and dynamic tree clipping a total of 32 gene co-expression modules were detected (**Figure 3.6**) In 'turquoise' there were 4116 genes, 'blue' with 3636, 'grey' with 3158, 'yellow' with 1941, 'green' with 1503, 'red' with 1488, 'black' with 962, 'pink' with 787, 'magenta' with 710, 'purple' with 447, 'greenyellow' with 286, 'tan' with 267, 'salmon' with 263, 'cyan' with 218, 'midnightblue' with 181, 'lightcyan' with 173, 'grey60' with 158, 'lightgreen' with 156, 'lightyellow' with 150, 'royalblue' with 147, 'darkred' with 141, 'darkgreen' with 129, 'darkturquoise' with 124, 'darkgrey' with 118, 'orange' with 113, 'darkorange' with 105, 'white' with 103, 'skyblue' with 96, 'saddlebrown' with 83, 'steelblue' with 66, 'paleturquoise' with 65. Eigengene expression plots were also made to model clusters over time **Figure 3.7**.

The 'turquoise' cluster showed a strong positive correlation with TimeG, (0.83) and a GO term pathway enrichment was conducted to identify any

overrepresented gene function in this cluster. After filtering the cluster contained 149 genes which were associated with 3 significant GO terms, ‘iron-sulfur cluster binding” with 5 genes from the cluster “metal cluster binding” also with 5 genes from the cluster and “nucleus” with 75 genes from the cluster.

The ‘green’ cluster showed a strong negative correlation with “group” (0.78) (group is either “short” or “long”). After filtering there were 282 genes, which associated with one GO term “response to oxygen-containing compound” with 56 genes from the cluster.

The ‘yellow’ cluster showed a strong positive correlation with “group” (0.78). After filtering there were 342 genes, which associated with 89 significant GO terms (**Table 3.2**). The GO terms that had the highest number of genes from the cluster were “intercellular” with 260 genes, “intracellular part” with 259 genes, “intracellular organelle” with 233 genes, “organelle” with 233 genes. No cluster indicated a clear involvement in senescence or ethylene reception or ethylene biosynthesis upon examination of the gene ontology.

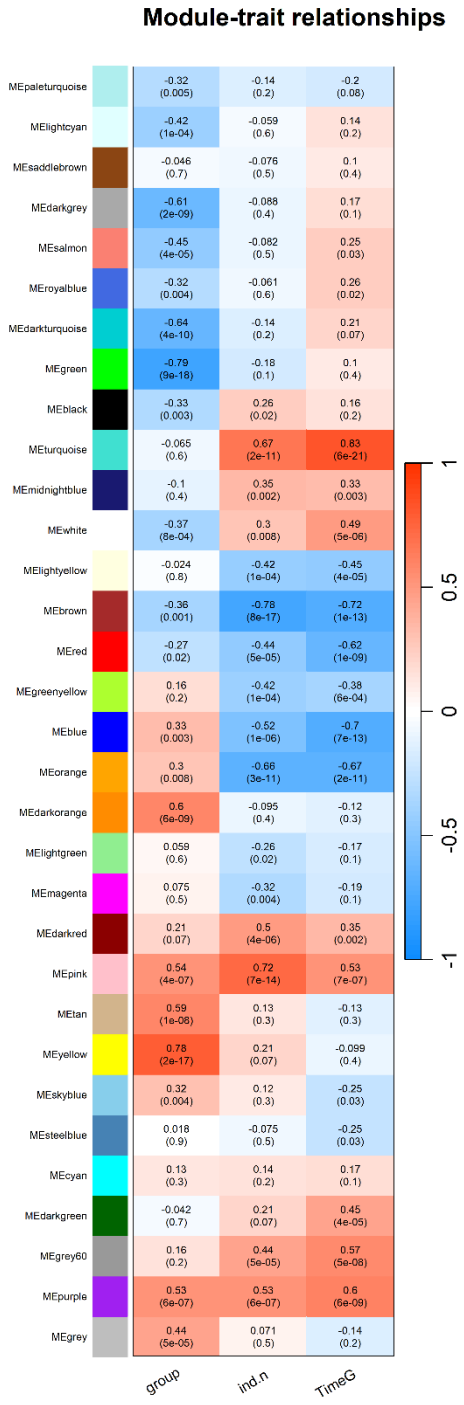


Figure 3.6 Weighted Gene Co-expression Network Analysis (WGCNA) for RNA storage samples. Clusters are identified by color names on the left, correlations with variables indicated red or blue. Red indicates a positive correlation; blue indicates a negative correlation. r^2 is shown in the top of each cell and *p*-value is below.

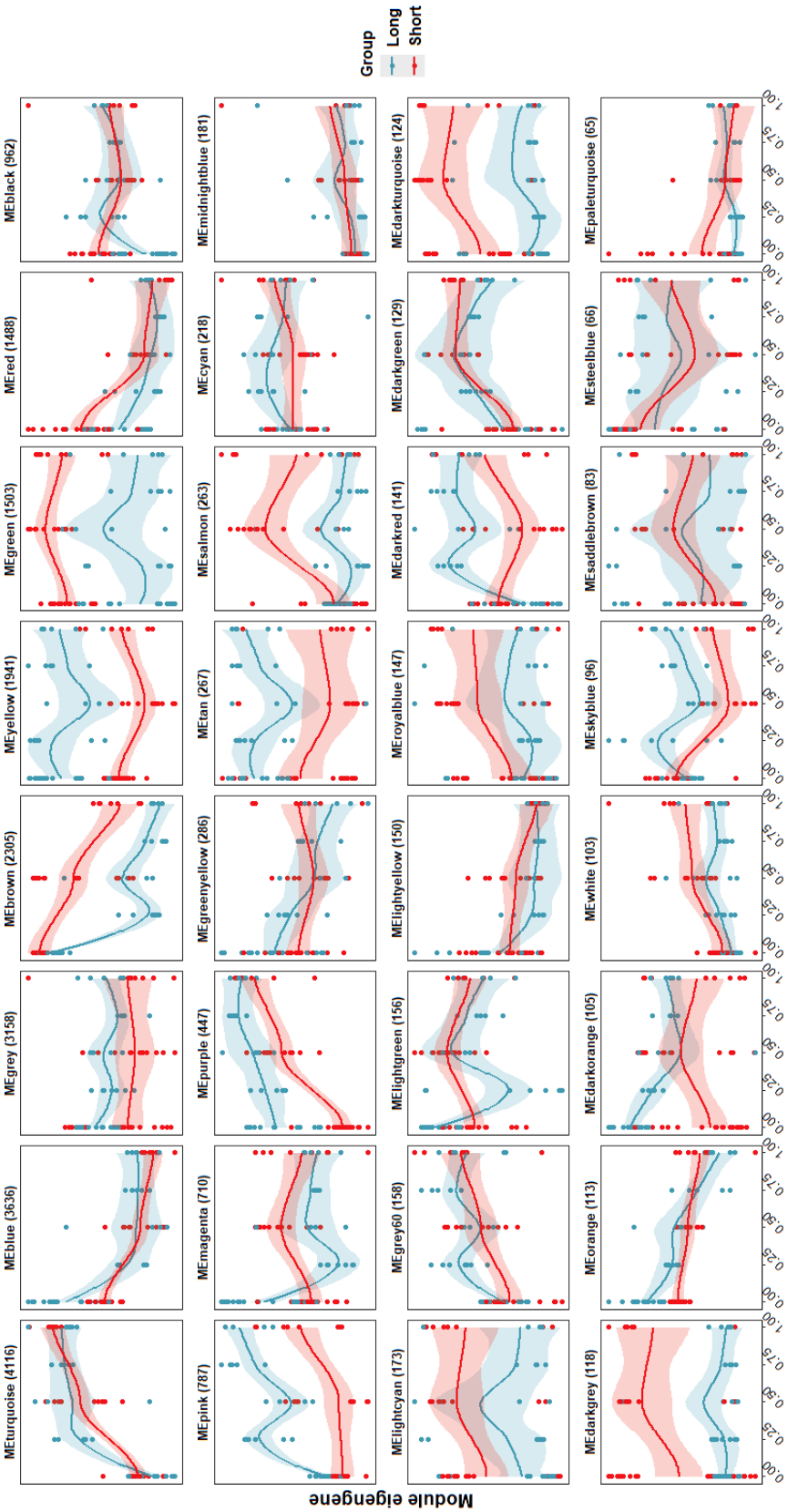


Figure 3.7 Modeled eigengene trends for each cluster developed by WGCNA. Long storing samples are represented by blue and short storing samples by red. The normalized percentage of storage time is on the x axis.

Table 3.2 Gene Ontology output for 'yellow' cluster identified by WGCNA. "GO term" is the Gene Ontology term number, "Description" is the explanation of what that GO term is associated with, "Number in input list" is the number of genes from the 'yellow' cluster were found in that GO term.

GO term	Description	Number in input list
GO:0046483	heterocycle metabolic process	122
GO:0034641	cellular nitrogen compound metabolic process	126
GO:0006139	nucleobase-containing compound metabolic process	115
GO:0006996	organelle organization	71
GO:0090304	nucleic acid metabolic process	100
GO:1901360	organic cyclic compound metabolic process	125
GO:0006807	nitrogen compound metabolic process	129
GO:0006725	cellular aromatic compound metabolic process	119
GO:0010467	gene expression	91
GO:0016070	RNA metabolic process	88
GO:0032259	Methylation	29
GO:0043414	macromolecule methylation	28
GO:0044271	cellular nitrogen compound biosynthetic process	84
GO:0006259	DNA metabolic process	32
GO:0051169	nuclear transport	15
GO:0018130	heterocycle biosynthetic process	74
GO:0051276	chromosome organization	32
GO:0034654	nucleobase-containing compound biosynthetic process	68
GO:0006913	nucleocytoplasmic transport	15
GO:0006396	RNA processing	34
GO:1901362	organic cyclic compound biosynthetic process	77
GO:0071840	cellular component organization or biogenesis	90
GO:0016043	cellular component organization	83
GO:1901293	nucleoside phosphate biosynthetic process	18
GO:0009220	pyrimidine ribonucleotide biosynthetic process	11
GO:0009218	pyrimidine ribonucleotide metabolic process	11
GO:0019438	aromatic compound biosynthetic process	71
GO:0043170	macromolecule metabolic process	148
GO:0006221	pyrimidine nucleotide biosynthetic process	11
GO:0006220	pyrimidine nucleotide metabolic process	11
GO:0044260	cellular macromolecule metabolic process	139
GO:0009165	nucleotide biosynthetic process	17
GO:0072528	pyrimidine-containing compound biosynthetic process	11
GO:0016458	gene silencing	18

GO:0072527	pyrimidine-containing compound metabolic process	11
GO:0006325	chromatin organization	23
GO:1902593	single-organism nuclear import	10
GO:0006606	protein import into nucleus	10
GO:0044744	protein targeting to nucleus	10
GO:0051170	nuclear import	10
GO:1901566	organonitrogen compound biosynthetic process	45
GO:0009560	embryo sac egg cell differentiation	9
GO:0034504	protein localization to nucleus	10
GO:0007292	female gamete generation	9
GO:0009561	Megagametogenesis	10
GO:0006306	DNA methylation	11
GO:0006304	DNA modification	11
GO:0006305	DNA alkylation	11
GO:0044728	DNA methylation or demethylation	11
GO:0031047	gene silencing by RNA	15
GO:0009260	ribonucleotide biosynthetic process	12
GO:0046390	ribose phosphate biosynthetic process	12
GO:1901137	carbohydrate derivative biosynthetic process	22
GO:0035194	posttranscriptional gene silencing by RNA	11
GO:0034645	cellular macromolecule biosynthetic process	83
GO:0022412	cellular process involved in reproduction in multicellular organism	9
GO:0040029	regulation of gene expression, epigenetic	16
GO:0010605	negative regulation of macromolecule metabolic process	22
GO:0010629	negative regulation of gene expression	20
GO:0007005	mitochondrion organization	8
GO:0016441	posttranscriptional gene silencing	11
GO:0009059	macromolecule biosynthetic process	84
GO:0016246	RNA interference	10
GO:0003676	nucleic acid binding	59
GO:0003677	DNA binding	29
GO:0005488	Binding	164
GO:1901363	heterocyclic compound binding	109
GO:0097159	organic cyclic compound binding	109
GO:0003723	RNA binding	22
GO:0044424	intracellular part	259
GO:0005622	Intracellular	260
GO:0043229	intracellular organelle	233
GO:0043226	Organelle	233
GO:0043227	membrane-bounded organelle	227

GO:0043231	intracellular membrane-bounded organelle	227
GO:0043232	intracellular non-membrane-bounded organelle	32
GO:0043228	non-membrane-bounded organelle	32
GO:0005634	Nucleus	135
GO:1990234	transferase complex	15
GO:0005856	Cytoskeleton	10
GO:0044422	organelle part	70
GO:0044446	intracellular organelle part	70
GO:0032991	macromolecular complex	42
GO:0031975	Envelope	25
GO:0044428	nuclear part	20
GO:0031967	organelle envelope	25
GO:0009528	plastid inner membrane	5
GO:0043234	protein complex	32
GO:0044430	cytoskeletal part	8

Gene identification and profiling

By searching the GDDH13 genome and annotation for predicted gene function related to ethylene reception and biosynthesis, a list of over 300 genes were identified. Of those, 125 were found expressed in the peel-storage RNAseq data set with 22 passing the significance threshold of absolute value Log2Fold >1.5 and adjusted *p*-value < 0.001 Of these 22 genes, 9 were observed in the ‘pink’ cluster but do not appear in the pathway enrichment results for the ‘pink’ module.

The 22 genes that were determined to be differentially expressed included 9 ethylene-responsive transcription factors, 7 SAM related, and 2 ethylene synthesis and 4 with known overlap with ethylene processes but have different biological processes.

Ethylene associated genes in this study were separated into roughly three groups based on gene annotations from the GDDG13 reference annotation and published papers on gene function (Alonso et al., 2003; Chao & Rothenberg, 1997; Cheng et al., 2022; Di Matteo et al., 2010; Feng et al., 2020; Hall et al., 2007; Jiang et

al., 2022; Li et al., 2019; Licausi et al., 2013; Moffat et al., 2012; Van de Poel et al., 2013; Xu et al., 2022; Zhang et al., 2019; Zwack et al., 2013) genes upregulated by ethylene, genes downregulated by ethylene, and ethylene biosynthesis and feedback genes.

In the category of upregulated genes there were 8 genes identified that were differentially expressed (MD03G1020800, MD04G1058200, MD08G1176000, MD12G1245100, MD13G1130700, MD15G1078900, MD15G1275900, MD15G1286400). (**Figure 3.8)** This group included a predicted ethylene insensitive-3-like protein, predicted ERF5-like, ERF ABR1-related, ethylene sensor protein EIN4, ERF113-like, ERF011-like, ethylene receptor, and AP2-like ERF TOE3.

In the category of down regulated genes, there were 4 genes (MD02G1265300, MD06G1215500, MD15G1326800, MD16G1043500). (**Figure 3.9)** This group included ERF WRI1, ERF CRF5-like isoform, ERF060, predicted ERF118-like.
In the category of ethylene biosynthesis and feedback genes there were 9 genes identified that were differentially expressed in this dataset (MD06G1223800, MD09G1079800, MD09G1079900, MD09G1254000, MD10G1187000, MD13G1141700, MD15G1143500, MD16G1057400, MD17G1270700). (**Figure 3.10)** This group included genes annotated as; SAM dependent Methyltransferase (3), SAMDC(2), ACO, SAM synthase, SAMDC proenzyme, and ethylene biosynthesis V (engineered).

Other genes known to be a part of the ethylene biosynthesis pathway or known to affect storage time were also identified. An additional *ACS* gene was identified from the genome (MD15G1203500) but was not differentially expressed, *PG1*

(MD10G1179100) was differentially expressed between the two treatment groups.

(**Figure 3.11**)

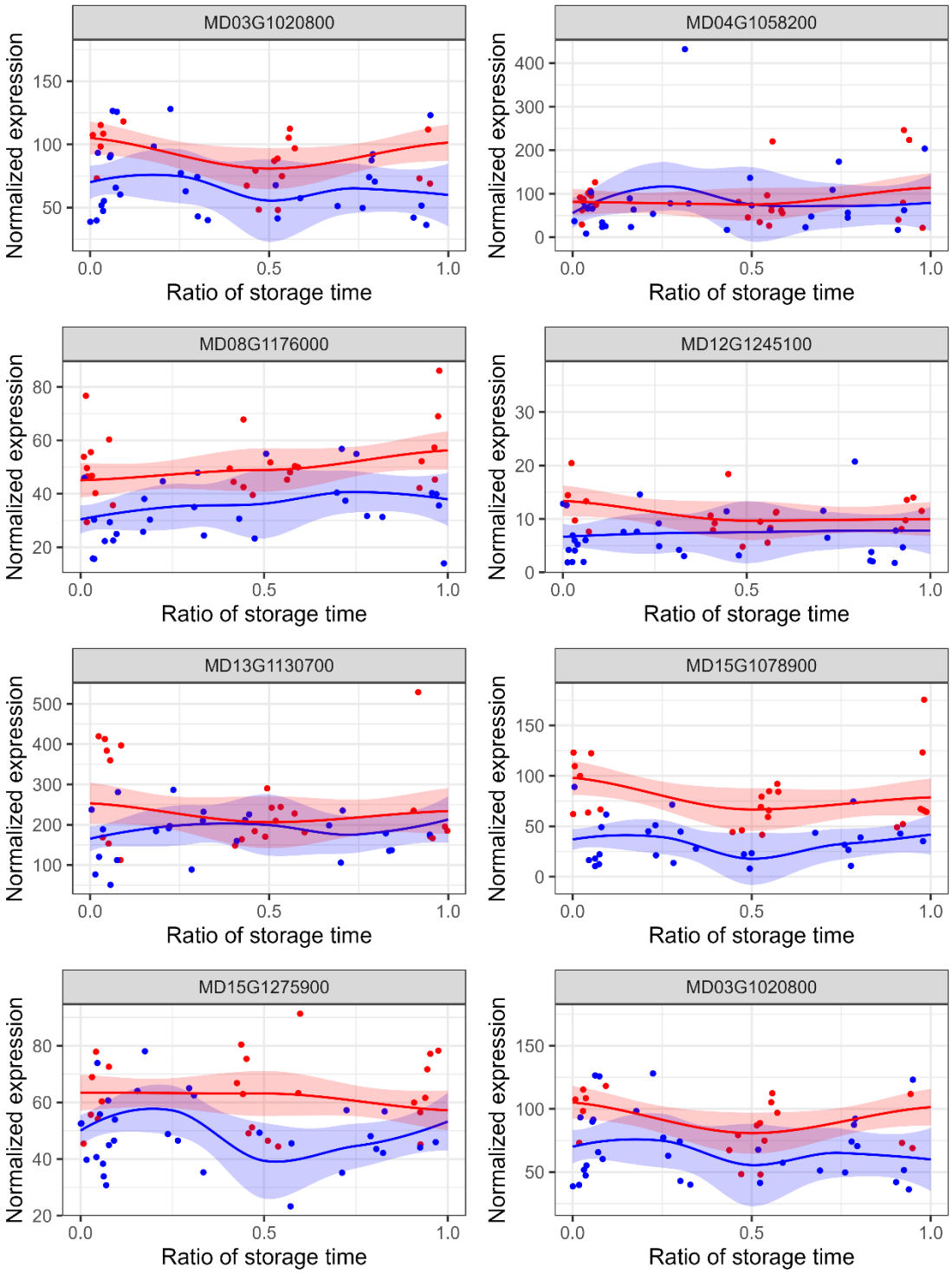


Figure 3.8 Eigengene expression plots of genes upregulated by ethylene. Short storing samples are indicated in red; the long storing samples are blue. Gene names are displayed at the top of each eigengene expression plot. Normalized eigengene expression is on the y axis and the ratio of total storage time on the x axis.

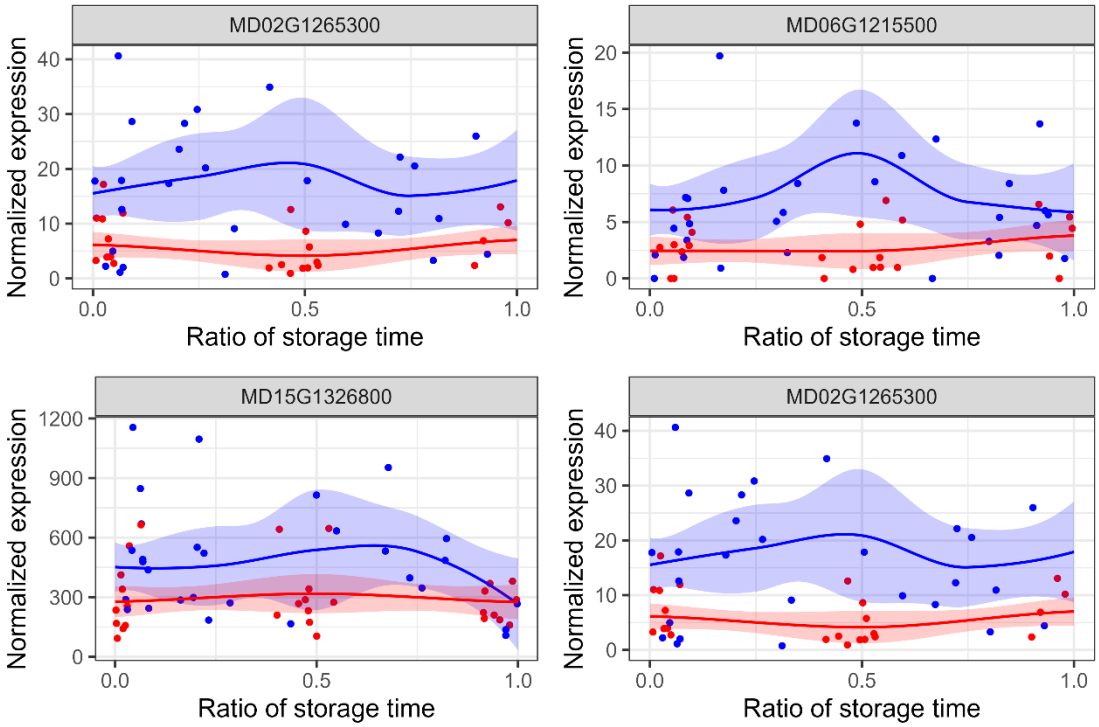


Figure 3.9 Eigengene expression plots of genes downregulated by ethylene. Short storing samples are indicated in red; the long storing samples are blue. Gene names are displayed at the top of each eigengene expression plot. Normalized eigengene expression is on the y axis and the ratio of total storage time on the x axis.

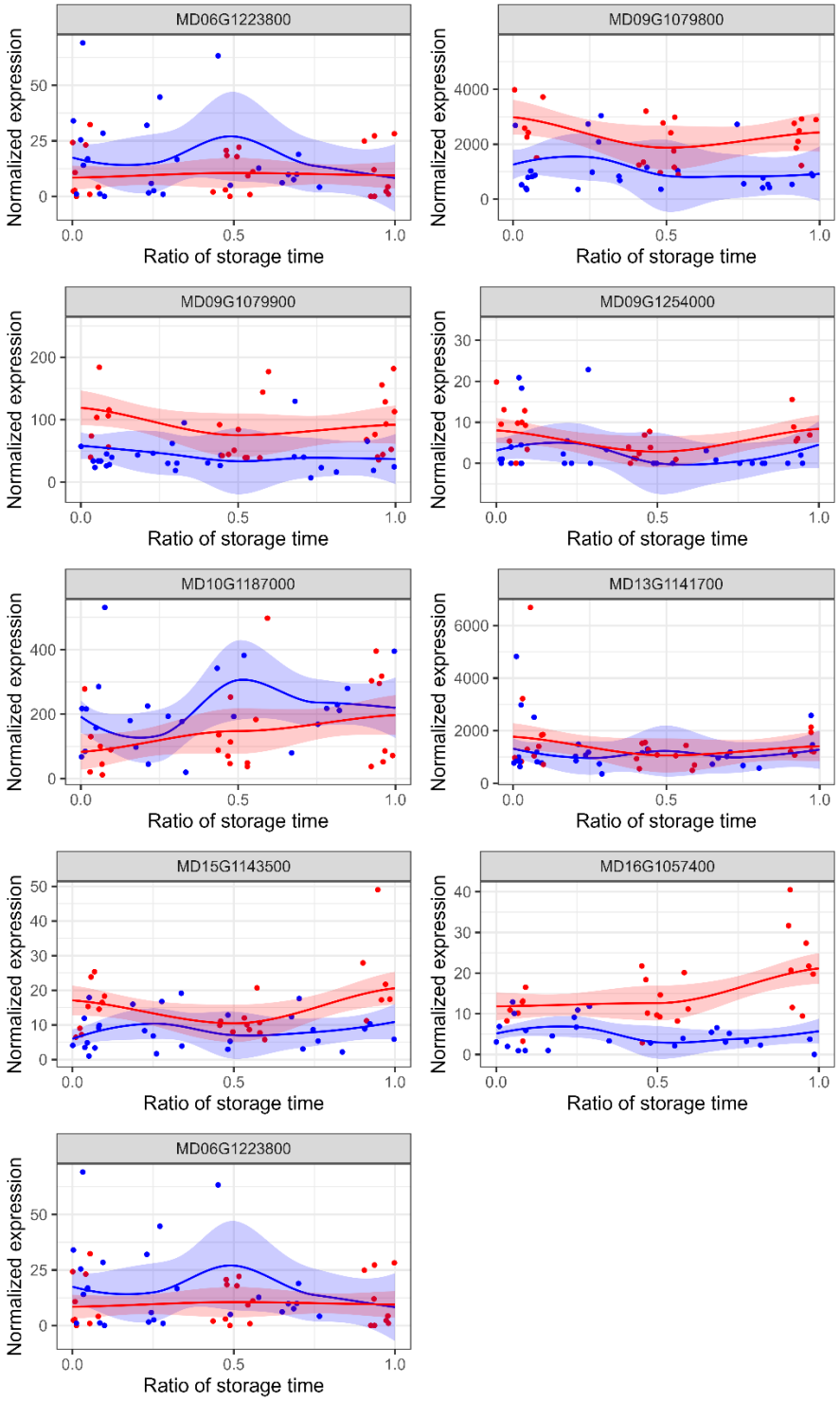


Figure 3.10 Eigengene expression plots of genes involved in the biosynthesis and feedback of ethylene. Short storing samples are indicated in red; the long storing samples are blue. Gene names are displayed at the top of each eigengene expression plot. Normalized eigengene expression is on the y axis and the ratio of total storage time on the x axis.

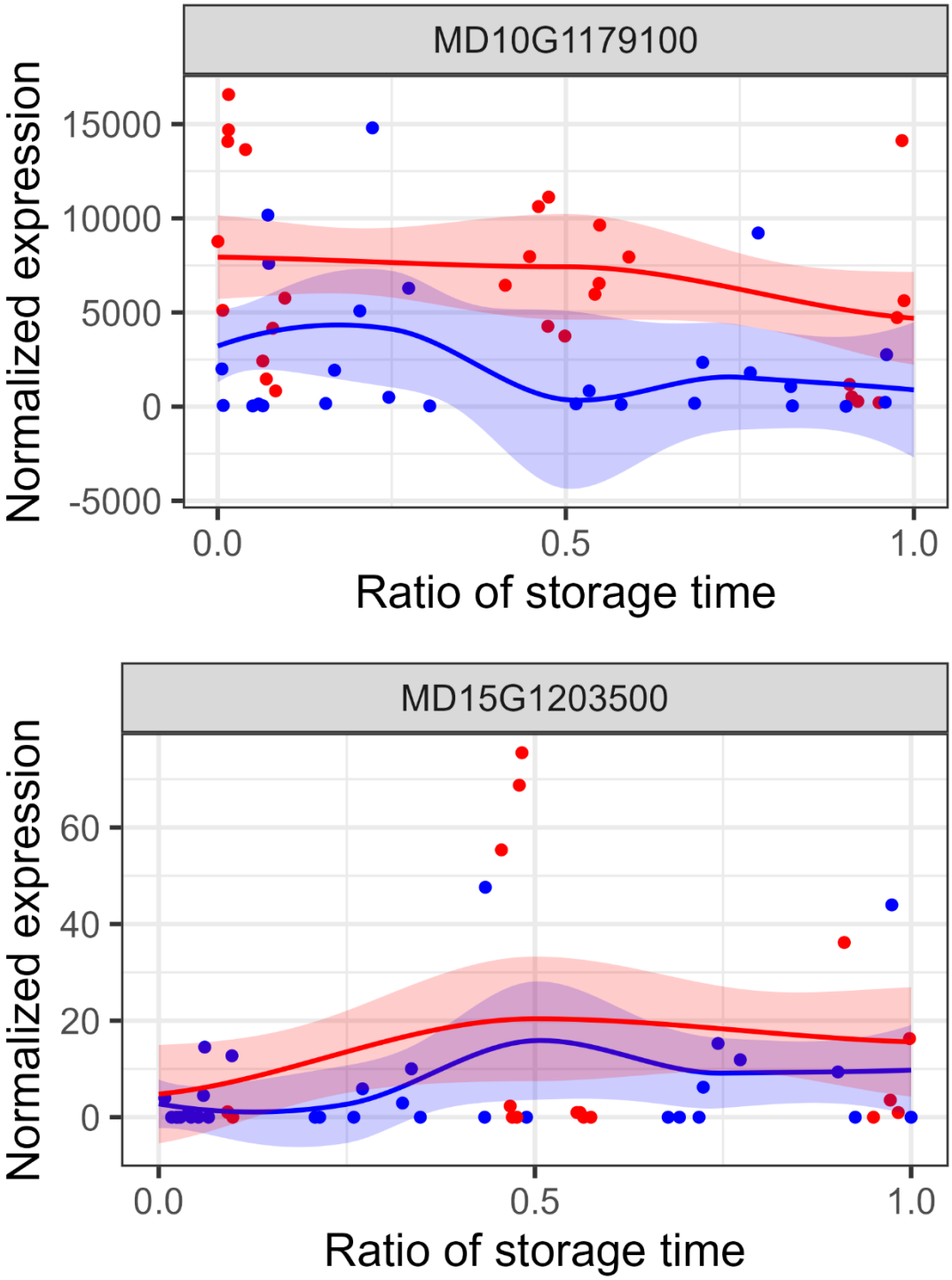


Figure 3.11 Eigengene expression plots of *ACS1*(MD151203500) and *PG1*(10G1179100). *PG1* was differentially expressed *ACS1* was not. Short storing samples are indicated in red; the long storing samples are blue. Gene names are displayed at the top of each eigengene expression plot. Normalized eigengene expression is on the y axis and the ratio of total storage time on the x axis.

DISCUSSION

Storage is an important trait with impacts on economics and sustainability. Most studies in apple do not investigate the genetic control of storage, instead research focuses on optimizing energy expensive storage control methods due to the cultivar specific marketing practices in apple due to the ability to clonally propagate (see discussion in introduction). This study worked to understand the genetics of storage capability in this introgressed mapping population to evaluate if any new loci coming from wild species might be useful. This study also offered a unique chance to examine the genetic control of storage and ethylene associated with an uncommon *ACO* genotype because of selection in common *M. domestica* cultivars.

In the initial harvest of a small number of fruiting individuals in 2018 that were not analyzed, it was observed that several individuals fruit stayed firm to thumb pressure much longer than others in the sample. This led to the observations of storage time in later years. NYCo1, while not an incredibly long storing individual (9 months), does store well enough when compared to other *M. domestica* cultivars and stores significantly longer than PI 589816 (3 months). Surprisingly, fruit of some of the trees were observed to senesce while still attached to the tree suggesting there was a strong segregation for this storage trait in this mapping population. This predisposition for progeny in this population to senesce on the tree resulted in some progeny being excluded due to their fruit appearing diseased, rotten, or otherwise not processable, and were not harvested. The discrepancy in segregation ratios has resulted from these excluded progeny and the higher numbers of long storing fruit being measured.

Based on previous research on climacteric fruit and their ripening and senescence patterns, it was hypothesized that the storage differences may be driven by differences in ethylene content or perception (D. O. Adams & Yang, 1981; Y. Chen et al., 2018; Pech et al., 2008). Ethylene is well understood to be necessary for senescence and climacteric fruit ripening (Adams-Phillips et al., 2004; Yang et al., 2013). Internal ethylene concentration has previously been used to quantify ripening and senescence in apple (Costa et al., 2005) Following on this expectation, the first characterization focused on measuring internal ethylene concentration. Despite testing 26 short and 33 long storing genotypes, no consistent pattern of internal ethylene could be attributed to the storage phenotype. There are several possible explanations for this observation. First, given the range of harvest dates in this mapping population, standardization of ripeness was elusive. Using methods such as a penetrometer (Symoneaux et al., 2005) was not possible given the large range in fruit size in this population, penetrometer readings are not reliable in small fruit because of the reduced thickness of the fruit flesh. Instead, ripeness was determined through ease of abscission and ground color change. Unfortunately, using this method could allow for variation in ripeness; with some fruit being overripe while others could have been still on the ripening curve. Additionally, some of the accessions were so small that drawing enough gas from the seed chamber proved to be a challenge. It is possible that these confounding factors obscured any correlation between storage phenotype and internal ethylene concentration.

The assessments of blue mold resistance offered no additional explanation for enhanced storability. Blue mold genetic resistance is a relatively new trait to be

identified and to date has only been identified in a population that used *M. siversii* as one of the parents, two QTLs for resistance to Blue Mold were identified on chromosome 3 and 10, with the one on chromosome 3 descending from *M. siversii* and the one on chromosome 10 descending from 'Royal Gala' (Norelli et al., 2017b). It is possible this trait is rare or only present in certain gene pools, or easily overcome by aggressive strains of the pathogen.

As a result of the ambiguous internal ethylene measures and lack of correlation with blue mold resistance, the mechanism behind the storage phenotype was examined using Genetic association analysis. The analysis was successful in identifying a set of SNPs on chromosome 10, near the presumed location for *ACO1* in the GDDH13 reference genome. The SNPs also fall near the annotated *ACO* genes in the *M. prunifolia* genome. In this same genomic region is the *PG1* gene. These two candidate genes are physically very close together and difficult to differentiate from each other in association studies. The interspecific nature of this population also makes distinguishing between these two candidates more difficult as it is not known if the same genomic regions of the *M. prunifolia* genome are syntenic or has structural variation. In the *M. prunifolia* reference genome a gene annotated as a polyglacturonase gene is located with start and stop points of 34679429 – 34682946, while in the GDDH13 genome *PG1* is located at 27285865 – 27288648, which overlaps the SNPS identified in this study.

Given the parental genotypes for *ACO1* and the significant association of SNPs in the *ACO1* region of LG 10, genes associated with the ethylene biosynthesis pathway and feedback system were examined using RNAseq and differentially

expressed genes. Examining the apple annotation for ethylene related genes identified 375 genes related to ethylene in the GDDH13 genome. Of which 22 were seen to be differentially expressed between short and long storing apples.

One of these genes *S-adenosyl-L-methionine-decarboxylase* (*SAMDC*) has been shown to be upregulated by the perception of ethylene in tomato (Van de Poel et al., 2013). In this study two different *SAMDC* genes (MD09G1079800, MD09G1079900) were upregulated in the short storing group, suggesting that ethylene levels may be increased in the short storing phenotype. A *SAMDC* protoenzyme gene (MD16G1057400) was also differentially expressed, suggesting that that the *SAMDC* synthesis pathway is also stimulated in short storing apples. The first step in the ethylene pathway (**Figure 3.1**, discussed above) involves SAM being converted to ACC, SAM is also the major methyl donor in plant biological processes, because of competition for SAM as a methyl donor as well as to produce ethylene, genes using SAM as a methyl donor were examined (Di Matteo et al., 2010). Three SAM methyltransferase genes were differentially expressed, two were upregulated in short storing and one was up regulated in the long storing group. SAM dependent methyltransferase have wide variation in structure and binding affinity (Boissier et al., 2006), so it is difficult to determine the expected expression for these three genes, as a result their differentially expressed profiled do not indicate that there are differences in SAM utilization.

The next step in ethylene biosynthesis is the conversion of SAM to ACC by *ACS*. In this study, the *ACS* gene was transcribed, however it was not differentially

expressed between the two groups, as expected given the presumed lack of segregation of this locus in this mapping population.

The next step is conversion of ACC by *ACO* into ethylene. In this study, two different *ACO* genes were transcribed, though only one of the genes were differentially expressed. The *ACO* homolog MD10G1187000 was differentially expressed, though the expression patterns overlap between short and long storing individuals. However, it is important to note that the causal mutations behind the allelic forms of *ACO* are not known to cause transcription level differences, but instead the allelic forms produce varying levels of functional enzymes (Houben & Van de Poel, 2019). Consequently, it is possible that even with similar levels of *ACO* transcription, individuals that are heterozygous at *ACO* may produce less functional enzyme and potentially less ethylene.

Ethylene Response Factors (ERF) are a broad class of transcription factors that cover a wide number of biological processes, often relating to stress response (Licausi et al., 2013). I have divided the ERF genes in this study into two groups; ERF genes that are upregulated by ethylene, and ERF genes that are repressed by the presence of ethylene.

In the repressed group four genes were identified (MD02G1265300, MD06G1215500, MD15G1326800, MD16G1043500) all fall into the broad category of ERF but are in different groups within that class. MD02G1265300 is annotated as *ERF WRI1* in Arabidopsis. This *WRI1* gene has been characterized using *Arabadopsis* and Oil Palm and has shown to be induced by darkness but suppressed significantly by endogenous ethylene (Zhang et al., 2019). This observation is similar to the gene

expression seen in this study where the short storing group shows lower transcription than the long storing group, presumably due to increased ethylene produced by the *ACO* genotype in the short storing group. MD06G1215500 is annotated as an *ERF CRF6-like isoform X1* in the GDDH13 genome. *Cytokinin response factor 6* (*CRF6*) has been associated with negatively controlling leaf and tissue senescence, as well as other abiotic stresses (Zwack et al., 2013). The gene expression pattern observed, with the long storing individuals having higher transcription of *CRF6*, suggests that the long storing group were able to delay senescence by inducing this *CRF6* gene. MD15G1326800 is annotated as ERF060, which has been associated with increased wax biosynthesis and increased drought tolerance (Jiang et al., 2022). Increased cuticle wax has been identified as a response to cold storage and ethylene, increasing the storage time and delaying senescence (Li et al., 2019). The last gene in the downregulated group is an *ERF118-like* gene that's transcription is induced by temperature stress (Xu et al., 2022)

The third group of differentially expressed genes identified, containing 8 genes) are genes we would expect to be upregulated by ethylene (MD03G1020800, MD04G1058200, MD08G1176000, MD12G1245100, MD13G1130700, MD15G1078900, MD15G1275900, MD15G1286400). One such gene is an ethylene receptor (MD15G1275900) that responds to changes in ethylene level and is involved in ethylene signaling (Hall et al., 2007). This group included two genes that are annotated as ethylene insensitive genes (MD03G1020800, MD12G1245100). Contrary to their name, ethylene insensitive genes have been long associated with the induction of integral ethylene signaling pathways (Alonso et al., 2003; Chao & Rothenberg,

1997). These two genes show lower transcript levels in the long storing group than the short storing group, suggesting that they are responding to a potential increase of more ethylene in the short storing group. The genes identified as ERFs that are upregulated by the presence of ethylene include an *ERF5-like, ERF ABR1-related, ERF113-like, AP2-like ERF,* and *ERF011-like* genes. These ERFs are involved in wounding response, senescence pathways, and disease resistance once induced by ethylene (Cheng et al., 2022; Feng et al., 2020; Jiang et al., 2022; Licausi et al., 2013; Moffat et al., 2012). Notably *ERF/ABR1* has been identified as an activator of *polygalacturonase* (*PG*) in peach, *PG* promotes softening and senescence of peach fruit during storage in the same mode as it does in *Malus* (Cheng et al., 2022). *PG1* was observed to be transcribed significantly higher in the short storing group, indicating that across storage time *PG1* is being induced presumably acting to breakdown pectin in fruit cell walls, causing more rapid senescence and softening in the short storing accessions. In this group of 8 genes, dominated by ERFs, all display higher transcription in the short storing group than the long, indicating that their transcription is being induced by higher levels of ethylene.

Together these findings suggest that the cause of the variation in storing time was due to segregation at *ACO1* located on Chromosome 10, and presumably it's increased levels of ethylene that is perceived by various ERFs to promote senescence, further outlined by the genes in the biosynthesis and feedback pathways responding accordingly. Additionally, enhanced cell breakdown is also likely due to the large differences in expression of *PG1*, with the short storing group displaying much higher expression.

Genetic causes of variation in storage are of interest because of the unique storing practices done with apple, and the global market that requires storage and transport around the world. Since storage of fruit has been so valuable to humans over time selection and fixation has occurred around many of these gene loci with *ACS* and *ACO* appearing in domestication sweeps and the region that contains *ACO* being associated with storage time and texture that has been selected for by breeders over time (Duan et al., 2017; Migicovsky et al., 2021). In most studies *ACS* is identified as the rate limiting step of ethylene biosynthesis and *ACO* has a minor effect (Costa et al, 2005; Oraguzie et al., 2004; Yuan et al., 2020; Zhu & Barritt, 2008), however in this study segregation at *ACO* and the change in ethylene perception is enough to cause fruit to senesce on the tree compared to fruit that stay firm and sound in storage up to a year after harvest. As interspecific breeding becomes more important and consequential it is important to identify areas of potential improvement and areas that more intense selection must occur to maintain the quality that is required for fresh market production.

REFERENCES

Adams, D. 0, Shang, A., & Yang, F. A. (1977). Methionine Metabolism in Apple Tissue IMPLICATION OF S-ADENOSYLMETHIONINE AS AN INTERMEDIATE IN THE CONVERSION OF METHIONINE TO ETHYLENE'. In *Plant Physiol* (Vol. 60). https://academic.oup.com/plphys/article/60/6/892/6075576

Adams, D. 0, Yang, S. F., & Stumpf, P. K. (1979). *Ethylene biosynthesis: Identification of 1-aminocyclopropane-1-carboxylic acid as an intermediate in the conversion o fmethionine to ethylene*. *76*(1), 170–174. https://www.pnas.org

Adams, D. O., & Yang, S. F. (1981). Ethylene the gaseous plant hormone: mechanism and regulation of biosynthesis. *TIBS*, 161–164.

Adams-Phillips, L., Barry, C., Kannan, P., Leclercq, J., Bouzayen, M., & Giovannoni, J. (2004). Evidence that CTR1-mediated ethylene signal transduction in tomato is encoded by a multigene family whose members display distinct regulatory features. *Plant Molecular Biology*, *5*(3), 387–404.

Alonso, J. M., Stepanova, A. N., Solano, R., Wisman, E., Ferrari, S., Ausubel, F. M., & Ecker, J. R. (2003). Five components of the ethylene-response pathway identified in a screen for weak ethylene-insensitive mutants in Arabidopsis. *PNAS*, *100*(5), 2992–2997.

Argenta, L. C., de Freitas, S. T., Mattheis, J. P., Vieira, M. J., & Ogoshi, C. (2021). Characterization and quantification of postharvest losses of apple fruit stored under commercial conditions. *HortScience*, *56*(5), 608–616. https://doi.org/10.21273/HORTSCI15771-21

Argenta, L., Fan, X., & Mattheis, J. (2002). Impact of watercore on gas permeance and incidence of internal disorders in "Fuji" apples. In *Postharvest Biology and Technology* (Vol. 24). www.elsevier.com/locate/postharvbio

Audic, S., & Claverie, J.-M. (1997). *The Significance of Digital Gene Expression Profiles*.

Bassil, N. V., Chagné, D., Vanderzande, S., Volz, R. K., Weskett, R., Kirk, C., Profitt, N., Gardiner, S. E., & Peace, C. P. (2019). Validation of SNP markers for fruit quality and disease resistance loci in apple (*Malus* × *domestica* Borkh.) using the OpenArray® platform. *Horticulture Research*, *6*(1). https://doi.org/10.1038/s41438-018-0114-2

Boissier, F., Bardou, F., Guillet, V., Uttenweiler-Joseph, S., Daffé, M., Quémard, A., & Mourey, L. (2006). Further insight into S-adenosylmethionine-dependent methyltransferases: Structural characterization of Hma, an enzyme essential for the

biosynthesis of oxygenated mycolic acids in mycobacterium tuberculosis. *Journal of Biological Chemistry*, *281*(7), 4434–4445. https://doi.org/10.1074/jbc.M510250200

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Boller, T., Herner, R. C., & Kende, H. (1979). Assay for and enzymatic formation of an ethylene precursor, 1-aminocyclopropane-1-carboxylic acid. *Planta*, *145*, 293–303.

Bowen, J. H., & Watkins ', C. B. (1997). Fruit maturity, carbohydrate and mineral content with watercore in "Fuji" apples. In *Postharvest Biology and Technology* (Vol. 11).

Brummell, D. A., & Harpster, M. H. (2001). Cell wall metabolism in fruit softening and quality and its manipulation in transgenic plants. In *Plant Molecular Biology* (Vol. 47).

Burg, S. P., & Burg, E. A. (n.d.). *Ethylene Action and the Ripening of Fruits Ethylene influences the growth and development of plants and is the hormone which initiates fruit ripening*. https://www.science.org

Cappellini, R. A. (1987). Disorders in Apple and Pear Shipments to the New York Market. *The American Phytopathological Society*, *71*(9), 852–856.

Chao, Q., & Rothenberg, M. (1997). Activation of the Ethylene Gas Response Pathway in Arabidopsis by the Nuclear Protein ETHYLENE-INSENSITIVE3 and Related Proteins. In *Cell* (Vol. 89).

Chen, K.-E., Liu, Y.-F., Liu, Y.-C., Sulistio, M., Chen, C.-C., & Wu, C.-T. (2023). Nonclimacteric ripening characteristics of 'Jen-Ju Bar' guava conferred by a defect in the expression of the system-2 ACC synthase gene PgACS1. *Postharvest Biology and Technology*, *195*, 112147. https://doi.org/10.1016/j.postharvbio.2022.112147

Chen, Y., Grimplet, J., David, K., Castellarin, S. D., Terol, J., Wong, D. C. J., Luo, Z., Schaffer, R., Celton, J. M., Talon, M., Gambetta, G. A., & Chervin, C. (2018). Ethylene receptors and related proteins in climacteric and non-climacteric fruits. In *Plant Science* (Vol. 276, pp. 63–72). Elsevier Ireland Ltd. https://doi.org/10.1016/j.plantsci.2018.07.012

Cheng, C., Liu, J., Wang, X., Wang, Y., Yuan, Y., & Yang, S. (2022). PpERF/ABR1 functions as an activator to regulate PpPG expression resulting in fruit softening during storage in peach (Prunus persica). *Postharvest Biology and Technology*, *189*. https://doi.org/10.1016/j.postharvbio.2022.111919

Costa, F., Peace, C. P., Stella, S., Serra, S., Musacchi, S., Bazzani, M., Sansavini, S., & Van de Weg, W. E. (2010). QTL dynamics for fruit firmness and softening around an ethylene-dependent polygalacturonase gene in apple (MalusDomestica Borkh.). *Journal of Experimental Botany*, *61*(11), 3029–3039. https://doi.org/10.1093/jxb/erq130

Costa, F., Stella, S., Van De Weg, W. E., Guerra, W., Cecchinel, M., Dallavia, J., Koller, B., & Sansavini, S. (2005). Role of the genes Md-ACO1 and Md-ACS1 in ethylene production and shelf life of apple (Malus domestica Borkh). *Euphytica 2005 141:1*, *141*(1), 181–190. https://doi.org/10.1007/S10681-005-6805-4

Di Guardo, M., Bink, M. C. A. M., Guerra, W., Letschka, T., Lozano, L., Busatto, N., Poles, L., Tadiello, A., Bianco, L., Visser, R. G. F., Van De Weg, E., & Costa, F. (2017). Deciphering the genetic control of fruit texture in apple by multiple family-based analysis and genome-wide association. *Journal of Experimental Botany*, *68*(7), 1451–1466. https://doi.org/10.1093/jxb/erx017

Di Matteo, A., Sacco, A., Anacleria, M., Pezzotti, M., Delledonne, M., Ferrarini, A., Frusciante, L., & Barone, A. (2010). *The ascorbic acid content of tomato fruits is associated with the expression of genes involved in pectin degradation*. http://www.biomedcentral.com/1471-2229/10/163

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15–21. https://doi.org/10.1093/bioinformatics/bts635

Duan, N., Wang, N., Xiang, L., Mao, L., Legall, N., Mao, Z., Fazio, G., Wang, K., Sun, H., Wang, X., Wang, Y., Liu, J., Xu, Y., Jiao, C., Wang, D., Yin, C., Jiang, S., Fei, Z., Wan, S., … Chao, T. (2017). Genome re-sequencing reveals the history of apple and supports a two-stage model for fruit enlargement. *Nature Communications*, *8*(1). https://doi.org/10.1038/s41467-017-00336-7

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, *6*(5), 1–10. https://doi.org/10.1371/journal.pone.0019379

Feng, K., Hou, X. L., Xing, G. M., Liu, J. X., Duan, A. Q., Xu, Z. S., Li, M. Y., Zhuang, J., & Xiong, A. S. (2020). Advances in AP2/ERF super-family transcription factors in plant. In *Critical Reviews in Biotechnology*. Taylor and Francis Ltd. https://doi.org/10.1080/07388551.2020.1768509

Gapper, N. E., Bowen, J. K., & Brummell, D. A. (2023). Biotechnological approaches for predicting and controlling apple storage disorders. *Current Opinion in Biotechnology*, *79*. https://doi.org/10.1016/j.copbio.2022.102851

Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., & Buckler, E. S. (2014). TASSEL-GBS: A High Capacity Genotyping by Sequencing Analysis Pipeline. *PLOS ONE*, *9*(2), e90346. https://doi.org/10.1371/JOURNAL.PONE.0090346

Hall, B. P., Shakeel, S. N., & Schaller, G. E. (2007). Ethylene receptors: Ethylene perception and signal transduction. In *Journal of Plant Growth Regulation* (Vol. 26, Issue 2, pp. 118–130). https://doi.org/10.1007/s00344-007-9000-0

Hamilton, A. J., Lycett, G. W., & Griereon, D. (n.d.). Antisense gene that inhibits synthesis of the hormone ethylene in transgenic plants. *Nature*, *346*, 284–287.

Harada, T., Sunako, T., Wakasa, Y., Soejima, J., Satoh, T., & Niizeki, M. (2000). An allele of the 1-aminocyclopropane-1-carboxylate synthase gene (Md-ACS1) accounts for the low level of ethylene production in climacteric fruits of some apple cultivars. *Theoretical and Applied Genetics*, *101*(5–6), 742–746. https://doi.org/10.1007/s001220051539

Holdsworth, M. J., Schuch, W., & Grierson, D. (1988). Organisation and expression of a wound/ripening-related small muitigene family from tomato. In *Plant Molecular Biology* (Vol. 11).

Houben, M., & Van de Poel, B. (2019). 1-aminocyclopropane-1-carboxylic acid oxidase (ACO): The enzyme that makes the plant hormone ethylene. In *Frontiers in Plant Science* (Vol. 10). Frontiers Media S.A. https://doi.org/10.3389/fpls.2019.00695

Jaeger, S. R., Machín, L., Aschemann-Witzel, J., Antúnez, L., Harker, F. R., & Ares, G. (2018). Buy, eat or discard? A case study with apples to explore fruit quality perception and food waste. *Food Quality and Preference*, *69*, 10–20. https://doi.org/10.1016/j.foodqual.2018.05.004

Jiang, Q., Wang, Z., Hu, G., & Yao, X. (2022). Genome-wide identification and characterization of AP2/ERF gene superfamily during flower development in Actinidia eriantha. *BMC Genomics*, *23*(1). https://doi.org/10.1186/s12864-022-08871-4

Jurick, W. M., Janisiewicz, W. J., Saftner, R. A., Vico, I., Gaskins, V. L., Park, E., Forsline, P. L., Fazio, G., & Conway, W. S. (2011). Identification of wild apple germplasm (Malus spp.) accessions with resistance to the postharvest decay pathogens Penicillium expansum and Colletotrichum acutatum. *Plant Breeding*, *130*(4), 481–486. https://doi.org/10.1111/j.1439-0523.2011.01849.x

King, G. J., Maliepaard, C., Lynn, J. R., Alston, F. H., Durel, C. E., Evans, K. M., Griffon, B., Laurens, F., Manganaris, A. G., Schrevens, E., Tartarini, S., & Verhaegh, J.

(2000). Quantitative Genetic Analysis and Comparison of Physical and Sensory Descriptors Relating to Fruit Flesh Firmness in Apple (Malus pumila Mill.). *Theor Appl Genet* , *100*, 1074–1084.

Kunihisa, M., Moriya, S., Abe, K., Okada, K., Haji, T., Hayashi, T., Kawahara, Y., Itoh, R., Itoh, T., Katayose, Y., Kanamori, H., Matsumoto, T., Mori, S., Sasaki, H., Matsumoto, T., Nishitani, C., Terakami, S., & Yamamoto, T. (2016). Genomic dissection of a 'Fuji' apple cultivar: Re-sequencing, SNP marker development, definition of haplotypes, and QTL detection. *Breeding Science*, *66*(4), 499–515. https://doi.org/10.1270/jsbbs.16018

Langfelder, P., & Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics*, *9*. https://doi.org/10.1186/1471-2105-9-559

Larsen, B., Migicovsky, Z., Jeppesen, A. A., Gardner, K. M., Toldam-Anderson, T. B., Myles, S., Orgaard, M., Petersen, M. A., & Pedersen, C. (2019). Genome-Wide Association Studies in Apple Reveal Loci for Aroma Volatiles, Sugar Composition, and Harvest Date. *The Plant Genome* , *12*(2), 1–15.

Lasserre, E., Bouquin, T., Hernandez J Bull, J. A., Pech, J., Balagu, C., Lasserre, E., Bouquin, T., & Hernandez J-C Pech -C Balagu, J. A. (1996). *Structure and expression of three genes encoding ACC oxidase homologs from melon (Cucumis melo L.)*. Springer-Verlag.

Lau, O. L., Liu, Y., & Yang, S. F. (1986). *Effects of Fruit Detachment on Ethylene Biosynthesis and Loss of Flesh Firmness Skin Color, and Starch in Ripening "Golden Delicious" Apples*.

Li, F., Min, D., Ren, C., Dong, L., Shu, P., Cui, X., & Zhang, X. (2019). Ethylene altered fruit cuticular wax, the expression of cuticular wax synthesis-related genes and fruit quality during cold storage of apple (Malus domestica Borkh. c.v. Starkrimson) fruit. *Postharvest Biology and Technology*, *149*, 58–65. https://doi.org/10.1016/j.postharvbio.2018.11.016

Licausi, F., Ohme-Takagi, M., & Perata, P. (2013). APETALA2/Ethylene Responsive Factor (AP2/ERF) transcription factors: mediators of stress responses and developmental programs. In *The New phytologist* (Vol. 199, Issue 3, pp. 639–649). https://doi.org/10.1111/nph.12291

Longhi, S., Hamblin, M. T., Trainotti, L., Peace, C. P., Velasco, R., & Costa, F. (2013). A candidate gene based approach validates Md-PG1 as the main responsible for a QTL impacting fruit texture in apple (Malus x domestica Borkh). *BMC Plant Biology*, *13*(1). https://doi.org/10.1186/1471-2229-13-37

Lopez-Gomez, R., Campbell, A., Dong, J.-G., Yang, S.-F., & Gomez-Lim, M. A. (1997). Ethylene biosynthesis in banana fruit: isolation of a genomic clone to ACC oxidase and expression studies. In *Plant Science* (Vol. 123).

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12). https://doi.org/10.1186/s13059-014-0550-8

Marlow, G. C., & Loescher, W. H. (2011). Watercore. In *Horticultural Reviews* (pp. 189–251). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118060797.ch6

McClure, K. A., Gardner, K. M., Douglas, G. M., Song, J., Forney, C. F., DeLong, J., Fan, L., Du, L., Toivonen, P. M. A., Somers, D. J., Rajcan, I., & Myles, S. (2018). A Genome-Wide Association Study of Apple Quality and Scab Resistance. *The Plant Genome*, *11*(1), 1–14. https://doi.org/10.3835/plantgenome2017.08.0075

Migicovsky, Z., Yeats, T. H., Watts, S., Song, J., Forney, C. F., Burgher-MacLellan, K., Somers, D. J., Gong, Y., Zhang, Z., Vrebalov, J., Giovannoni, J. G., C Rose, J. K., & Myles, S. (2021). *Apple ripening is controlled by a NAC transcription factor Running title: NAC18.1 controls apple ripening*. 1–43.

Moffat, C. S., Ingle, R. A., Wathugala, D. L., Saunders, N. J., Knight, H., & Knight, M. R. (2012). ERF5 and ERF6 play redundant roles as positive regulators of JA/Et-mediated defense against botrytis cinerea in arabidopsis. *PLoS ONE*, *7*(4). https://doi.org/10.1371/journal.pone.0035995

Murr, D. P., & Yang, S. F. (1975). CONVERSION OF S-METHYLTHIOADENOSINE TO METHIONINE BY APPLE TISSUE. In *Phytochemistry* (Vol. 14). Pergamon Press.

Norelli, J. L., Wisniewski, M., Fazio, G., Burchard, E., Gutierrez, B. L., Levin, E., & Droby, S. (2017a). Genotyping-by-sequencing markers facilitate the identification of quantitative trait loci controlling resistance to Penicillium expansum in Malus sieversii. *PLOS ONE* , 1–24. https://doi.org/10.1371/journal.pone.0172949

Norelli, J. L., Wisniewski, M., Fazio, G., Burchard, E., Gutierrez, B., Levin, E., & Droby, S. (2017b). Genotyping-by-sequencing markers facilitate the identification of quantitative trait loci controlling resistance to Penicillium expansum in Malus sieversii. *PLoS ONE*, *12*(3). https://doi.org/10.1371/journal.pone.0172949

Oraguzie, N. C., Iwanami, · H, Soejima, · J, Harada, · T, & Hall, · A. (2004a). O R I G I N A L P A P E R Inheritance of the Md-ACS1 gene and its relationship to fruit softening in apple (Malus ” domestica Borkh.). *Theor Appl Genet*, *108*, 1526–1533. https://doi.org/10.1007/s00122-003-1574-8

Oraguzie, N. C., Iwanami, H., Soejima, J., Harada, T., & Hall, A. (2004b). Inheritance of the Md-ACS1 gene and its relationship to fruit softening in apple (Malus x domestica Borkh.). *Theoretical and Applied Genetics*, *108*(8), 1526–1533. https://doi.org/10.1007/s00122-003-1574-8

Pech, J. C., Bouzayen, M., & Latché, A. (2008). Climacteric fruit ripening: Ethylene-dependent and independent regulation of ripening pathways in melon fruit. In *Plant Science* (Vol. 175, Issues 1–2, pp. 114–120). https://doi.org/10.1016/j.plantsci.2008.01.003

Ruperti, B., Bonghi, C., Rasori, A., Ramina, A., & Tonutti, P. (2001). Characterization and expression of two members of the peach 1-aminocyclopropane-1-carboxylate oxidase gene family. *Physiologia Plantarum*, *111*(3), 336–344. https://doi.org/10.1034/j.1399-3054.2001.1110311.x

Sharples, R. O. (1967). A Note on the Occurrence of Watercore Breakdown in Apples during 1966. *Plant Pathology*, *16*(3), 119–120. https://doi.org/10.1111/j.1365-3059.1967.tb00384.x

Sunako, T., Sakuraba, W., Senda, M., Akada, S., Ishikawa, R., Niizeki, M., & Harada, T. (1999). *An Allele of the Ripening-Specific 1-Aminocyclopropane-1-Carboxylic Acid Synthase Gene (ACS1) in Apple Fruit with a Long Storage Life 1*. https://academic.oup.com/plphys/article/119/4/1297/6096206

Symoneaux, R., Royer, G., Madieta, E., & Jourjon, F. (2005). *Acoustic and Sensory Measurements of Different Ripeness of Apples*.

Tandonnet, S., & Torres, T. T. (2017). Traditional versus 3′ RNA-seq in a non-model species. *Genomics Data*, *11*, 9–16. https://doi.org/10.1016/j.gdata.2016.11.002

Tian, M. S., Prakash, S., Zhang, N., & Ross, G. S. (2002). *Chilling-induced ethylene biosynthesis in Braeburn apples*.

Van de Poel, B., Bulens, I., Oppermann, Y., Hertog, M. L. A. T. M., Nicolai, B. M., Sauter, M., & Geeraerd, A. H. (2013). S-adenosyl-l-methionine usage during climacteric ripening of tomato in relation to ethylene and polyamine biosynthesis and transmethylation capacity. *Physiologia Plantarum*, *148*(2), 176–188. https://doi.org/10.1111/j.1399-3054.2012.01703.x

Ververidis, P., & John, P. (1991). COMPLETE RECOVERY IN vZTR0 OF ETHYLENE-FORMING ENZYME ACTIVITY. In *Phytochemistry* (Vol. 30, Issue 3).

Vicente, A. R., Saladié, M., Rose, J. K. C., & Labavitch, J. M. (2007). The linkage between cell wall metabolism and fruit softening: Looking to the future. In *Journal of the*

Science of Food and Agriculture (Vol. 87, Issue 8, pp. 1435–1448). https://doi.org/10.1002/jsfa.2837

Watkins, C. B., & Mattheis, J. P. (2019). *Postharvest Physiological Disorders in Fruits and Vegetables* (1st ed., Vol. 1).

Xu, S. da, Geng, X. min, Mao, L. feng, Yi, Y., Gong, J. yi, & Xu, X. rong. (2022). Transcriptome analysis and identification of the genes associated with the heat stress response in four rhododendron species. *Scientia Horticulturae*, *303*. https://doi.org/10.1016/j.scienta.2022.111176

Yamada, H., Mukai, T., & Fukasawa-Akada, T. (2012). Comparison of the sub-cellular compartmentation of sugars in mature apples of two cultivars susceptible to different types of watercore and grown in different climates. *Journal of Horticultural Science and Biotechnology*, *87*(1), 17–22. https://doi.org/10.1080/14620316.2012.11512824

Yang, X., Song, J., Campbell-Palmer, L., Fillmore, S., & Zhang, Z. (2013). Effect of ethylene and 1-MCP on expression of genes involved in ethylene biosynthesis and perception during ripening of apple fruit. *Postharvest Biology and Technology*, *78*, 55–66. https://doi.org/10.1016/j.postharvbio.2012.11.012

Yuan, H., Yue, P., Bu, H., Han, D., & Wang, A. (2020). Genome-wide analysis of ACO and ACS genes in pear (Pyrus ussuriensis). *In Vitro Cellular and Developmental Biology - Plant*, *56*(2), 193–199. https://doi.org/10.1007/s11627-019-10009-3

Zhang, Q., Sun, R., Zheng, Y., Yuan, Y., & Li, D. (2019). Isolation and characterization of the EgWRI1 promoter from oil palm (Elaeis guineensis Jacq.) and its response to environmental stress and ethylene. *PLoS ONE*, *14*(12). https://doi.org/10.1371/journal.pone.0225115

Zhu, Y., & Barritt, B. H. (2008). Md-ACS1 and Md-ACO1 genotyping of apple (Malus x domestica Borkh.) breeding parents and suitability for marker-assisted selection. *Tree Genetics & Genomes* , *4*, 555–562. https://doi.org/10.1007/s11295-007-0131-z

Zwack, P. J., Robinson, B. R., Risley, M. G., & Rashotte, A. M. (2013). Cytokinin response factor 6 negatively regulates leaf senescence and is induced in response to cytokinin and numerous abiotic stresses. *Plant and Cell Physiology*, *54*(6), 971–981. https://doi.org/10.1093/pcp/pct049

CONCLUSIONS

Apple is an important food crop, culturally, nutritionally, and environmentally. However, issues such as increased disease pressure and the changing climate challenge our ability to grow, store, and ship apples. Introgression from wild sources of disease resistance and novel traits offers solutions to many of these issues, though our understanding of introgression of fruit quality traits remains limited. This project was outlined with the main research questions surrounding understanding the genetic controls of commercially important traits when introgressing genetics from a wild species of *Malus* rarely used in breeding, to gain information and context for using this specific *Malus prunifolia* accession in future breeding Acid levels in the progeny population have higher acidity, but apples of acceptable acidity are present in the population in the first generation. This higher acidity is not due to secondary acids such as citric acid, but this study proposed a more complicated system for the controls of acidity, and offers the chance for future research to understand the controls of high acid in this population. Fruit quality traits like fruit size, sweetness, and phenolic concentration were all explored to see if there were unique genetics affecting them from *M. prunifolia.* All these traits appear to be polygenically controlled, with unique genetic associations possible from wild species. SNPs were identified in regions of the genome with no genes located nearby, showing areas of potential future research. Introgression studies offer a unique opportunity to examine the processes of domestication and selection by creating allelic combinations that are otherwise rare in the cultivar breeding pool, such as the downstream effects of increased ethylene due to *ACO*1 gene. By examining the inheritance and the downstream effects of segregation

at *ACO1* our understanding of the impact of the *ACO1* and understanding of traits to be aware of when introgressing from a non-selected species was increased.

Together these studies provide information and context for using wild species apples to address large challenges that plant breeders must face. New sources of resistance to diseases and climate change should be explored, and these studies combined with previous studies creates a launching pad for further breeding using this individual and population.

From measured information I have made a list of recommended individuals for future breeding work. These individuals were selected due to their large fruit size, and acceptable brix, phenolic, and acidity concentrations as well as their *ACO1* genotypes (**Table 4.1**)

Table 4.1 List of progeny that are recommended for future breeding efforts. Diameter (mm), Height (mm) Mass (g), Russet 1 = present in some level 0 = not observed, Brix, TA (g/L) Phenolics (g GAE/L). NA indicates that trait was not present to measure.

Individual	Diameter	Height	Mass	Russet	Brix	TA	Phenolics
26_7	83.57	70.501	261.373	0	12.3	NA	0.274
32_2	74.328	73.396	189.493	0	11.1	2.61	0.302
26_10	74.471	64.797	184.295	0	11	5.19	0.559
43_10	72.551	62.273	150.496	1	14	1.97	0.912
48_4	68.603	62.724	145.37	0	10.6	1.13	0.344