# EVOLUTION OF SUBGENOMES IN TETRAPLOID SOUR CHERRY (*PRUNUS CERASUS* L.)

By

Kathleen E.B. Rhoades

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Plant Breeding, Genetics, and Biotechnology – Horticulture – Doctor of Philosophy

2023

## ABSTRACT

Sour cherry (*Prunus cerasus* L., $2n = 2x = 32$) is an allotetraploid fruit tree species native to Eastern Europe that resulted from the interspecific hybridization between the tetraploid ground cherry *P. fruticosa* Pall. ($2n = 4x = 32$) and the diploid sweet cherry *P. avium* L. ($2n = 2x = 16$). Sour cherry has been cultivated by humans for thousands of years, and the ability to clonally propagate through budwood cuttings means that many landraces have persisted for over 400 years and therefore their parental origins are unknown. Previous research on both crop and model polyploid species has shown that one subgenome of an allopolyploid may dominate the other(s) through a number of processes including gene expression bias and homoeologous exchanges and replacements favoring one subgenome over generations. By comparison, the subgenome dynamics of sour cherry have yet to be characterized. In the first chapter of this dissertation, I review the history of sour cherry, from the wild origins in Eastern Europe and Western Asia to the sour cherry industry in the US today, the current literature on polyploidy and its consequences for plant genomes and crop improvement, the progress made so far on sour cherry breeding and genetics in the modern era, and I give a brief overview of fruit development physiology for this crop. In the second chapter I present the first-ever reference genome of sour cherry, a chromosome-scale assembly of the sour cherry cultivar Montmorency that reveals three subgenomes, two inherited from the polyploid progenitor *P. fruticosa* (newly discovered to be allopolyploid, not auto- as previous literature suggests), and one inherited from the diploid progenitor *P. avium*. I was able to differentiate the subgenomes based on kmer content and further demonstrate the ancestral origin of the subgenomes through coding sequence phylogeny of a set of dormancy-associated MADS box genes associated with bloom timing in *Prunus.* The third chapter investigates the subgenome dynamics of a panel of four sour cherry landraces and

two bred cultivars by examining homoeolog dosage and overall subgenome expression bias. I identified 24 unique homoeologous exchange events in three of the six accessions, five whole-homoeolog replacements in two of these three accessions, and subgenome expression bias in favor of the ground cherry-derived A and A' subgenomes over the sweet cherry-derived B subgenome in all of the accessions. I partially validated the subgenome dosage of each accession using the previously-characterized *S*-alleles and demonstrated the implications of subgenome dosage and expression bias for a set of four expansins previously associated with fruit softening in sour cherry. Altogether this work represents substantial progress in genome resources and application of genomics techniques for sour cherry and a unique addition to the body of work on polyploid genomics.

This dissertation is dedicated in memory of my father, Allen Bitter, who passed away unexpectedly in 2019. He was a dairy goat breeder for close to 50 years and, I think without either of us fully realizing it, taught me to appreciate the art and science of breeding and agriculture. I am a plant breeder today because of him.

## ACKNOWLEDGEMENTS

First and foremost I must thank my advisor, Dr. Amy Iezzoni. In a phone call the week before I turned in this dissertation Amy and I talked about the results and discussion of my third chapter, and it struck me how excited she *still* is about this work, and how excited I *still* am about this work, and I realized how incredibly lucky I am to have both of these things at the end of a PhD. Amy has been a patient, diligent, caring mentor, who held me to a high standard as a scientist even when I was less than inclined to believe I could meet it.

I must also thank my committee, especially Dr. Pat Edger, who guided me through the wild world of polyploid research analysis and always made sure to acknowledge when I had worked hard. What he described as "fun conversations" about research were incredibly helpful to my development as a scientist. Dr. Courtney Hollender and Dr. Dave Douches were also valuable members of my committee and their enthusiasm for my research has always been much appreciated.

My dear, dear husband Samuel, a Texan by way of Tennessee, followed me up to the cold state of Michigan and has supported me through every step of the arduous process of graduate school. I could not have imagined a better partner to get through this thing called life, and I am thankful for him every day. These last few months he has quite literally kept me upright and functioning, and I love him more than words can say.

Endless thanks to Meghan Hill, our beloved graduate secretary, who is holding each and every one of us Horticulture graduate students together and deserves a six week vacation in a warm place with no internet.

To my sister in science, my esteemed collaborator, my favorite person with whom to pick apart preprints, and the only person who also speaks our twin language, Charity Goeckeritz, I

have so much gratitude. Charity is one of the smartest, hardest-working people I've ever met, she set an impossible standard for the rest of us at MSU to meet, and I have almost forgiven her for moving so far away. None of this work would have happened without our collaborations, our couch-based coworking, and her stubborn insistence that *something* was up with that 1.0 version of the sour cherry genome.

To the myriad other pseudo-siblings I have picked up in the department of horticulture over the years, in no particular order: Songwen Zhang, Chris Gottschalk, Josh VanderWeide, Phil Engelgau, Shujun Ou, Beth Alger, Annika Kohler, Andrea Kohler (no relation), Ying-chen Lin, Stephanie Rett-Cadman, Kevin Bird, Alan Yocca, Jeremy Pardo, Anna Pardo (yes relation), Kellie Walters, Mckenna Wilson, Thilani Jayakody, Joseph Hill, Elise Tomaszewski, Mallory St. Clair, and anyone I may have forgotten (for which I will kick myself thoroughly and buy you fries to make up for it). I love them all dearly and they have made my time at MSU such a pleasure. I wouldn't be half the scientist I am today without them all.

To Emily and Kait, my bestest friends, I love you profoundly and I so appreciate our ability to hold each other up and celebrate each other's wins. It's rare to have friends as good as you and I treasure it every day.

To my family: Mom, Susan, Tom, and Betsy, I love you guys so much. I love how eccentric and smart and strange our family is. I love you for supporting me in my weird endeavors. I love you for thinking I'm interesting and smart, something I don't always believe about myself.

And thanks, finally to my wealth of non-academia friends in Lansing, who have made this city feel like home.

## TABLE OF CONTENTS

# CHAPTER 1: LITERATURE REVIEW

**Part 1: Origins of sour cherry and a short history of sour cherry breeding efforts**

Sour cherry (*Prunus cerasus* L.) is a woody perennial fruit crop endemic to Eastern Europe. In its native range it can be found cultivated as landraces and bred cultivars and the fruit used for fresh eating and processed products such as jam, juice, pies, pastries, and fermented beverages (Faust et al. 2011). Sour cherry is an allotetraploid (2n = 4*x* = 32), originating in the wild as the result of an interspecific cross between the sweet cherry *P. avium* L. (2n = 2*x* = 16) and the tetraploid wild ground cherry *P. fruticosa* Pall. (2n = 4*x* = 32) (Olden and Nybom 1968; Hillig and Iezzoni 1988). Generally speaking, *P. avium* is endemic to southern and central Europe and is adapted to warmer climates, while *P. fruticosa* is native to Eastern Europe and Western Asia, reaching into Siberia (Faust et al. 2011). Where their ranges overlap, in the Carpathian Basin, the Balkans, and Anatolia, sour cherry resulted (Surányi 2021). Indeed, sour cherry exhibits a range of phenotypes in growth habit, climate adaptation, and fruit quality that is greater than the simple range between its two parental species (Hillig and Iezzoni 1988; Iezzoni et al. 1991; Surányi 2021).

Cherry pits have been recovered from cave dwellings in Europe dating back to 4000-5000 BCE, and cherry was probably first cultivated in Greece (Iezzoni et al. 1991). As Hedrick puts it in *The Cherries of New York* (1915), "…from the time tillage of plants was first practiced in the Old World, this fruit has been under cultivation, feeble, obscure, and interrupted by war and chase though its cultivation may have been. Certainly the history of cherry is as old as that of agriculture in the southern European countries and is interwritten with it." Very little is known about human selection of cherries before the 16$^{th}$ century, but landraces of sweet and sour cherries arose all over Europe, identified by their town or region of popularity. Most modern cherry cultivars today have at least one landrace parent (Iezzoni et al. 1991).

Sour cherries were introduced to North America through European colonization, and moved west with white settlers. French colonists in Canada and Michigan and English colonists in New England planted fruit trees (Hedrick 1915). Sour cherry cultivation in Michigan began in the late 19th century with a few growers who established orchards in the Traverse City area (Faust et al. 2011). The industry in Michigan grew with the establishment of the first cherry packing facility in 1912, and the number of sour cherry trees in Michigan increased from less than 500,000 in 1911 to over four million by the early 1960s (Faust et al. 2011).

In 2022, Michigan produced 181 million pounds of sour cherries on 23,000 acres, making up 74 percent of total sour cherry production nationwide (Michigan Fruit Production 2022 2023). The industry's biggest biological struggles in 2023 are shared with most other fruit crops in Michigan and in the US: pest pressure, especially from the spotted wing drosophila, disease pressure, and late spring freezes that kill blossoms (Wilson 2019; Longstroth and Irish-Brown 2021 Mar 30; Diseases). The vast majority (>90 percent) of sour cherries grown in the U.S. are 'Montmorency,' a 400+ year-old cultivar which was first introduced to North America in the colonial era (Hedrick 1915).

Economically, one of the larger challenges to the U.S. sour cherry industry has been extremely cheap imports from Turkey, which U.S. grower groups allege are the result of the Turkish government subsidizing production for growers and allowing them to set prices below the cost of production (Noble 2018 Oct 22). In 2019, four Michigan cherry processors and a Utah-based fruit growers co-op filed a petition with the US Department of Commerce and the US International Trade Commission (ITC) against Turkey to ask for a 650 percent tariff on all Turkish sour cherry imports, citing the damage cheap Turkish imports have done to the domestic sour cherry industry by driving prices below the cost of production for American growers (Noble

2019 Jun 5). Despite the Department of Commerce finding that Turkey was in fact subsidizing sour cherry production so that cherries could be exported to the U.S. and sold below market value in December 2019, the ITC ruled that Turkish imports were not a threat to the U.S. sour cherry industry and declined to levy a tariff (Bartkowiak, Jr. 2020 Jan 15). According to a 2022 report from Michigan State University Extension, production costs for Michigan sour cherries were between 40 and 44 cents per pound, while the USDA reports that growers received between 15 and 22 cents per pound of sour cherries between 2017 and 2019, before smaller crops in 2020 and 2021 (due to late spring freezes) drove prices up to 48 and 61 cents per pound, respectively (Bardenhagen et al. 2022; Michigan Agricultural Statistics 2021-2022, Fruit). As labor and fuel costs continue to rise and exceptional weather events become the norm due to climate change, the sour cherry industry in Michigan finds itself in a precarious position.

The sour cherry breeding program at Michigan State University was established in 1982 with the hiring of Dr. Amy Iezzoni as breeder. It was immediately obvious that the sour cherry germplasm in the U.S. at that time was not diverse enough to be able to produce meaningful improvements through breeding. In order to establish a germplasm collection diverse enough to be useful for improvement efforts, Dr. Iezzoni needed to go back to the area of origin of sour cherry to collect material, in Eastern Europe and the countries that made up the former USSR. Over the 1980s and 1990s Dr. Iezzoni made numerous trips to Eastern Europe in order to evaluate the diversity present in sour cherry's native range and collect pollen, seeds, and budwood to bring back to the U.S. (Iezzoni 2004). The germplasm collection at Michigan State is today the most diverse in North America.

The most well-known release from the MSU sour cherry breeding program is 'Balaton,' originally named 'Újfehértói fürtös' in its native Hungary. 'Balaton' was released in the U.S. in

1998 and royalty payments for the cultivar are shared with cherry researchers in Hungary (Charles 2013 May 27). ‘Balaton’ cherries are firmer, more flavorful, and have darker flesh than ‘Montmorency’ fruit, and in Michigan ‘Balaton’ is popular as a wine and juice cherry. Two other Hungarian cultivars ‘Erdi Jubileum’ and ‘Erdi Bötermö’ were also released in the U.S. at this time under the cultivar names ‘Jubileum’ and ‘Danube,’ respectively. Dr. Iezzoni also released five cherry rootstocks in 2019 with increased precocity and dwarfing behavior over previous cherry rootstocks, compatible with both sweet and sour cherry. In recognition of this accomplishment she was awarded MSU Innovator of the Year in 2019 (MSU Professor Recognized for the Fruits of Her Labor – MSU Innovation Center Annual Report).

**Part 2: Sour cherry genetics and genomics**

Early on, cytological studies of sour cherry foretold the challenges to come for breeders. As early as 1950, researchers observed chromosomes “pairing” in ones, threes, fours, and fives during meiosis (Hruby, K. 1950). Cytological studies in sour cherry of chromosome pairing at meiosis in the 21$^{st}$ century classify 8.5 to 14.5 percent of microspores as irregular in a variety of genotypes (Schuster and Wolfram 2004; Popovska et al. 2005; Akšić et al. 2016). Dechun Wang’s doctoral work showed abundant examples of aberrant chromosome pairing in ‘Montmorency,’ ‘Erdi Bötermö,’ and ‘Reinische Schattenmorelle’ pollen mother cells (1998).

Allopolyploids tend to exhibit disomic inheritance patterns as a result of chromosomes pairing exclusively with their respective homologs and not pairing with homoeologs from different subgenomes, while autopolyploids, which do not have separate subgenomes, exhibit tetrasomic inheritance because all homolog copies having equal likelihood of pairing with each other. The difference in expected progeny genotype ratios between disomic and tetrasomic inheritance schemes is illustrated quite well in Beaver and Iezzoni (1993). Consistent with

frequent irregular pairing at meiosis, marker data shows sour cherry to exhibit both disomic and tetrasomic inheritance (Beaver and Iezzoni 1993; Tsukamoto et al. 2010), indicating that the chromosomes do not always pair with their respective homologs. Meiotic irregularities caused by unequal subgenome balance may also be heightened by the interfertility and lack of reproductive isolation of sour cherry from its progenitors. Progenitor introgression in sour cherry in the wild has been directly observed (Macková et al. 2018).

This lack of bivalent pairing results in low gamete (pollen and egg) viability and fruit set. The consequence of the low maternal gamete fertility in cherry is particularly severe, as cherry is not a multi-seeded fruit (i.e. with a fleshy carpel) such as tomato or blueberry. Instead, cherry and all *Prunus* (i.e., stone fruits) have just two ovules, and one typically degenerates, leaving just one seed inside the pit. Fruit set requires successful fertilization and the initiation of embryo development. Therefore, any disruption that results in an inviable ovule or zygote will result in the flower not producing a fruit.

The first sour cherry linkage maps were developed with RFLPs in the late 1990s, using markers common to other *Prunus* species and numbering linkage groups according to homology with peach and almond (Wang, Dechun 1998). Subsequent QTL work in sour cherry was done with SSR and RFLP markers (Wang et al. 2000; Olmstead 2006). This work was limited due to the relative difficulty of working with an allotetraploid at this point in time, and sour cherry researchers mostly relied on the more abundant QTL work being done in sweet cherry, the diploid progenitor (Sooriyapathirana et al., 2010; Zhang et al., 2010). In 2012 the first SNP array for sweet and sour cherry was released by the RosBREED project (Peace et al. 2012). This allowed for more precise QTL mapping in both species, and in combination with the injection of research funds provided by RosBREED, many more QTL studies in sweet and sour cherry were

made possible in the following years (Campoy et al. 2016; Cai et al. 2017; Cai et al. 2018; Cai et al. 2019; Calle et al. 2020).

The first peach (*P. persica*) genome, constructed using Sanger sequencing and a *Prunus* reference map for chromosome scaffolding, was published in 2013 (Verde et al.). It was updated in 2017 using new linkage maps and short-read sequencing of the original Lovell double-haploid individual to correct misassembled sections, fix SNPs and indels, and fill in gaps (Verde et al.). Given the high synteny among *Prunus* species, the peach reference was a useful tool for sweet and sour cherry breeders to use in conjunction with their own linkage maps and markers.

The first sweet cherry reference sequence was released in 2017, and it and the subsequent few reference assemblies were constructed with Illumina short reads scaffolded to the peach reference (Shirasawa et al. 2017; Pinosio et al. 2020; Calle et al. 2021). The first sweet cherry reference constructed with long reads and not scaffolded to any other reference was released in 2020, and this assembly of cv. 'Tieton' indicates that all chromosomes of *P. avium* are longer than their orthologs in *P. persica* (Wang et al. 2020). This difference in chromosome length between species underlines the importance of assembling reference genomes according to that species' specific genome size and structure. While the 'Tieton' v2.0 reference was extremely helpful to sour cherry researchers, there was still a need for a sour cherry reference genome that would allow breeders and researchers to separate homoeologous genes and compare the subgenomes of sour cherry.

The first reference genome for sour cherry, see Chapter 2 of this dissertation, is an assembly of cv. 'Montmorency' produced using PacBio long reads, Illumina 2x150bp short reads for error correction, and HiC sequencing to guide chromosome scaffolding. In the process of assembling this genome we discovered that the *P. fruticosa* progenitor of sour cherry is likely

itself an allopolyploid, and it passed two subgenomes down to sour cherry. The final assembly has 771.8 Mb of sequence scaffolded into 24 chromosomes, and the overall homoeolog dosage of 'Montmorency' is 1A (*P. fruticosa*-derived):1A' (*P. fruticosa*-derived):2B (*P. avium*-derived). Kmer abundance analysis separated the A and A' chromosome sets from each other and from the B chromosome set, reflecting differences in repeat content between the subgenomes. We validated the ancestry of these subgenomes using protein phylogenies of 'Montmorency' genes alongside protein sequences from *P. avium* cv. 'Tieton' (Wang et al. 2020) and a draft assembly of *P. fruticosa* we assembled and released with the Montmorency genome (Goeckeritz et al. 2023).

A second reference genome for cv. 'Schattenmorelle' was released in a preprint in March of this year (Wöhner et al. 2023). It was assembled using Oxford Nanopore long reads and Illumina short reads, and the authors assembled two subgenomes, one originating from *P. fruticosa* and one originating from *P. avium.* Using a combination of sequence similarity approaches, the authors identified 28 regions of homoeologous exchanges (HE) between the *P. fruticosa* and *P. avium*-derived subgenomes, an interesting contrast to the lack of HE identified by our group in cv. 'Montmorency' (Goeckeritz et al. 2023; Wöhner et al. 2023).

Sour cherry genetic research has made substantial gains over the past decade, but very little is currently known about sour cherry as a polyploid system. Polyploidy research has also made substantial progress in the last decade, and so we next review the current allopolyploid literature as a basis for beginning to understand the implications of allopolyploidy for sour cherry.

**Part 3: Allopolyploidy and its consequences**

Allopolyploidy, the presence of three or more haploid sets of chromosomes resulting from a cross between two different species, is a common phenomenon in plants (Barker et al. 2016). The evolutionary advantage of polyploidy is thought to lie in the increased capacity to adapt to novel environments and circumstances. Polyploids are often more resilient to biotic and abiotic stresses than their diploid progenitors (Van de Peer et al. 2021). A 2017 review of polyploid literature found a large number of polyploidy events originating with the K-Pg extinction event, commonly known for the massive asteroid collision with earth that resulted in rapid climate changes and extinction of 60 to 70 percent of all animal and plant life on earth (Van de Peer et al. 2017). Studies in multiple kingdoms, including plants, have shown that polyploids are more prevalent at latitudes further from the equator, correlating with colder temperatures (Rice et al. 2019; David 2022).

The sudden presence of two genomes in a single nucleus presents a threat to stable meiosis and therefore a threat to the new polyploid's reproductive fitness, despite any adaptive advantages to the individual organism. This type of occurrence, referred to as a "genomic shock," triggers a sequence of events to re-stabilize meiotic activities within the newly-formed polyploid (neopolyploid) and ensure long-term proliferation. Indeed, newly-formed polyploids have been observed to be extremely genetically unstable, owing to the challenge of re-establishing stable meiosis after an allopolyploidy event (Mason and Wendel 2020). Stable meiosis in allopolyploids is achieved through preferential pairing of homologs (originating from the same progenitor) over homoeologs (syntenic but originating from different progenitors), and active suppression of homoeolog pairing. Despite this selective pressure, over 50 percent of allopolyploids exhibit some level of multivalent pairing at meiosis (Li et al. 2021).

As reviewed in Soltis et al. (2016), plants have a common set of mechanisms for genome behavior post-polyploidy event, but the details will vary from clade to clade. Generally speaking, allopolyploid genomes return to a diploid-like state over time through the combined processes of genome downsizing, genome rearrangement, changes in gene expression levels, neo- and subfunctionalization of duplicated genes, methylation repatterning, and the expansion or contraction of transposable elements (TEs). All allopolyploids do not take the same path towards diploidization, and the characteristics of the whole genome duplication event itself appear to heavily influence subsequent evolution and diploidization processes. Within these various diploidization processes there is the possibility that one subgenome can dominate the other(s) and will eventually come to make up the majority of the diploidized genome (Bird et al. 2018). For our purposes we will discuss subgenome dominance through two lenses: structural changes in the genome, and subgenome expression bias.

Homoeologous exchanges (HE), structural changes resulting from homoeologs pairing during meiosis and the subsequent crossover events, are an important evolutionary path for adaptation and divergence in polyploids as well as a method of diploidization. HE and other forms of "illegitimate recombination" (i.e. recombination between non-homologous chromosomes) are the main mechanism for gene deletion and genome size reduction in allopolyploids (Li et al. 2021). Patterns of HE and gene loss within a clade tend to be broadly consistent. In Brassicaceae, 13 separate mesopolyploid events were identified among ten Brassicaceae tribes, and the patterns of gene loss and retention were similar across tribes in terms of gene functions (Mandáková et al. 2017). But wild individuals of *Arabidopsis suecica*, an allotetraploid dating back about 16,000 years, show almost no signs of genome rearrangement. This is in contrast to resynthesized *A. suecica* individuals, suggesting that natural selection may

have acted against individuals exhibiting extensive genome rearrangement (Burns et al. 2021). Burns et al. (2021) speculate that these results imply that domesticated polyploids may not always be representative of the way natural selection acts upon polyploids. In a comparison of post-hexaploidy genome evolution of four Solanaceae genomes, biased fractionation resulted in the most recently-added subgenome having the fewest genes, a surprising contrast to Brassicaceae (McRae et al. 2022). The authors hypothesize that higher repetitive element content may be driving gene loss, similar to what was found in broomcorn millet (Sun et al. 2023), but they also consider that subgenome expression bias, which they did not examine in their study, may play a role in biased gene loss. Octoploid strawberry exhibits subgenome dominance in gene loss and homoeologous exchanges biased towards the *F. vesca*-derived subgenome (Edger et al. 2019).

The subclades of Poaceae exhibit remarkable variation in their structural responses to polyploidy. Within these groups, however, gene and chromosome loss tend to follow the same routes. Ozkan et al. (2001) found that synthetic and natural allohexaploid wheat had consistent and reproducible loss of genome-specific and chromosome-specific sequences, even when using different genotypes, ploidies, and reciprocal crosses. Thirteen years later the first genome assemblies of hexaploid bread wheat would reinforce this finding (The International Wheat Genome Sequencing Consortium (IWGSC) 2014). In maize, the 26 diverse genomes making up the parents of the maize nested association mapping population were found to have population-specific homoeolog retention patterns when comparing non-stiff-stalk temperate, tropical, and flint-derived lines, although overall fractionation is consistently biased against the M2 subgenome (Hufford et al. 2021). Annual bluegrass shows no evidence of biased fractionation, but one subgenome was found to have a preferential increase in gene number through HE events

(Benson et al. 2023). Broomcorn millet, a relatively recent polyploid (~0.48MYA), exhibits more gene loss in the subgenome with larger repeat content, and a significant portion of those genes were lost to TE insertions (Sun et al. 2023). Teff shows remarkably stable subgenome content and no evidence of large structural rearrangements or biased fractionation (VanBuren et al. 2020).

Generally speaking, the dosage balance hypothesis (Birchler and Veitia 2007; Birchler and Veitia 2012) does appear to be a factor in loss or retention of genes in polyploids. Transferases, protein kinases, and transcription factors are among those categories that tend to be preferentially retained over successive rounds of gene loss (Soltis et al. 2016). Dosage balance may also explain why a higher number of HE events and chromosome replacements is associated with lower pollen germination rates and lower overall fertility (Gaeta and Chris Pires 2010; Xiong et al. 2011; Gaebelein et al. 2019).

Subgenome expression bias, another phenomenon of allopolyploids, is the preferential expression of one subgenome over another independent of homoeolog dosage. It is hypothesized to be a step in the resolution of genetic and epigenetic conflicts between parental genomes in an allopolyploid, as genes that are lowly or not expressed tend to be lost over time (Colle et al. 2019). The earliest work observing possible subgenome expression bias comes from Wendel's group in 2003, where they examined a panel of 40 homoeologous gene pairs in allopolyploid *Gossypium* spp. and identified a minority showing biased expression, which in six cases switched subgenomes in different organs of the plant (Adams et al. 2003). They observed these expression differences in synthetic neopolyploids, indicating that expression changes occur immediately following the polyploidy event, a hypothesis that has gained support in the decades since (Wang et al. 2004; Edger et al. 2017; Bird et al. 2021).

Similar to the case with structural changes due to polyploidy, the presence, absence, and/or consistency of subgenome expression bias tends to vary by clade. Hexaploid bread wheat does not show evidence of any prevailing subgenome expression dominance, although synthetic hexaploid lines do show a strong bias against the D genome in the first generations post-polyploidy (The International Wheat Genome Sequencing Consortium (IWGSC) 2014; Vasudevan et al. 2023). Maize exhibits subgenome expression bias that varies by tissue (Walsh et al. 2020). Teff exhibits subgenome expression bias that remains fairly constant across tissue types (VanBuren et al. 2020). Tetraploid blueberry has been shown to exhibit subgenome expression bias, and the dominant subgenome changes in developing fruit tissue (Colle et al. 2019). In six resynthesized lines of *Brassica napus*, expression dominance favored the same subgenome (BnC) in all lines, although in *B. napus* seeds transcription factors in the BnA subgenome are observed to dominate expression in the seed coat, while BnC transcription factors dominate in the developing embryo (Bird et al. 2021; Khan et al. 2022). *B. napus* also exhibits a higher number of biased homoeolog pairs overall in response to fungal pathogen infection, although the BnC subgenome remains dominant overall (de Jong and Adams 2023). As one of the more commonly-studied allopolyploid systems, *B. napus* literature demonstrates how much variation in subgenome expression bias within one species.

Subgenome expression bias appears to be influenced in part by DNA methylation, which is strongly correlated with transposable element (TE) content. TEs themselves do not appear to affect gene expression bias, but TEs are frequently methylated to prevent their proliferation and this methylation is hypothesized to have "spill-over" effects on the expression of nearby genes (Alger and Edger 2020). In *Mimulus peregrinus* the dominant subgenome has fewer TEs and lower levels of DNA methylation near both genes and TEs, and this difference in methylation

between subgenomes increases over generations, although it is unclear which is the causative factor between gene methylation and expression dominance (Edger et al. 2017). Gene methylation is positively correlated with expression dominance in resynthesized lines of *B. napus,* and the authors also note it is unclear whether the DNA methylation patterns are a cause or a result of expression bias (Bird et al. 2021). In cases where progenitor genomes have similar levels of TE abundance, such as Cucurbitaceae and *Capsella bursa-pastoris,* a distinct lack of subgenome expression bias is observed (Douglas et al. 2015; Sun et al. 2017).

As everything above suggests, biased structural changes and subgenome expression bias are not independent processes. The dosage balance hypothesis appears to have a role in both, as the expression of dosage-sensitive homoeolog pairs will exhibit less variance in response to HE events than dosage-insensitive homoeolog pairs, except in cases where the pair is biased towards the non-dominant subgenome (Bird et al. 2023). In rice, changes in expression and gene methylation associated with allopolyploidization have been found to be exacerbated by HE events (Li et al. 2019). Homoeologs that are lowly-expressed are also more likely to be lost over time to HE events or other structural changes, as their loss will have a smaller impact on overall phenotype than loss of a highly-expressed copy (Schnable et al. 2011).

**Part 4: A brief overview of sour cherry fruit development**

As detailed above, for sour cherry fruit to develop, it must have a viable embryo. As sour cherry will not set or develop fruit without a viable embryo, reduced fertility directly translates into production issues for growers. Sour cherry germplasm typically exhibits just ~2 percent fruit set while at least 25 percent fruit set is needed for a viable commercial crop. This low level of fertility in sour cherry is associated with  meiotic irregularities which are believed to reduce gamete viability (Iezzoni et al. 2005; Popovska et al. 2005; Schuster and Wolfram 2005; Akšić et

al. 2016). This is in contrast to the vast majority of allopolyploid crops where fruit or seeds are the food product, as these crops have been indirectly selected for mechanisms that ensure bivalent pairing at meiosis and therefore high fertility.

Following successful fertilization, sour cherry fruit development follows a double sigmoid growth curve that is divided into three stages (Willing 1960) (Figure 1.1). The first stage, beginning immediately after fertilization, is characterized by exponential growth driven by rapid cell division. Stage one is the only stage in which cell division takes place, so at the end of stage one the fruit has its final number of cells (Tukey and Young 1939). Final cell number is the major factor in determining final fruit size (Olmstead 2006). In stage two, growth plateaus and endocarp tissue hardens, forming the pit (Tukey and Young 1939). Stage three consists of a second exponential growth phase, which occurs entirely through cell expansion. It is in this stage that anthocyanins, sugars, and other secondary metabolites will accumulate in the fruit and red or blush-colored fruit will change color. Stage three is also when cell wall-modifying enzymes begin to break down cell walls to accommodate expansion, and the fruit softens as a result (Tukey and Young 1939).

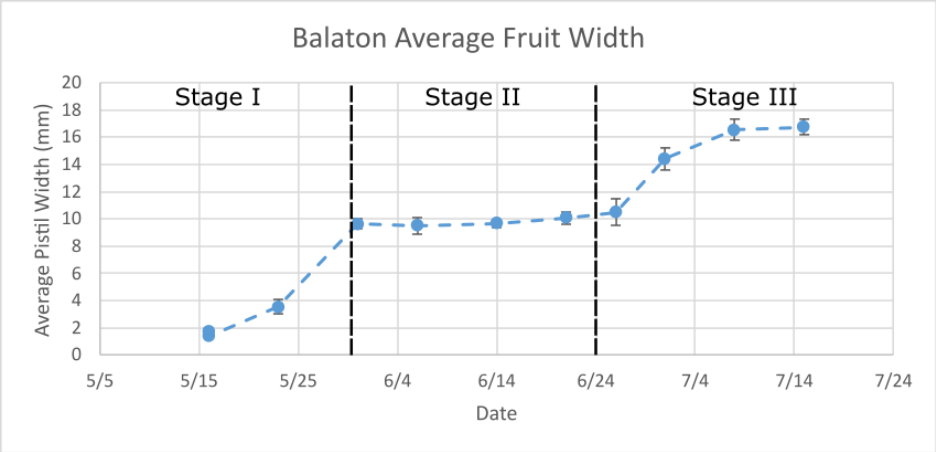
Over 20 years ago, a study of gene expression in ripening ‘Montmorency’ fruit resulted in the characterization of four genes encoding cell wall-modifying expansin enzymes (Yoo et al. 2003). Named exp1, exp2, exp4, and exp5, all four of these genes show peak expression in the final stages of fruit ripening, when fruit is softening through the breakdown of cell walls (Yoo et al. 2003). These are only a subset of cell wall-modifying enzymes implicated in fruit ripening processes, but Yoo et al. (2003) is one of the few studies to date examining fruit ripening in sour cherry specifically, and no follow-up work has been done on these expansins. Sour cherries, like sweet cherries, are non-climacteric fruit, so their ripening process is regulated by abscisic acid

(ABA), and they will not ripen further post-harvest with the application of ethylene (Ponce et al. 2021).

**Part 5: Objectives for this work**

My objectives for this dissertation work were as follows: (*i*) to assemble and annotate a high quality reference genome for sour cherry, (*ii*) to determine whether sour cherry exhibits differential subgenome dosage in a panel of six diverse accessions, (*iii*) to determine whether sour cherry exhibits subgenome expression bias in a panel of six diverse accessions, and (*iv*) to demonstrate the implications of the results of objectives ii and iii for a set of previously-characterized genes associated with fruit development.

**Figure 1.1:** A graph illustrating the *Prunus* sigmoidal fruit growth curve. The graph plots the width of the widest part of the developing fruit starting immediately after anthesis through maturity for cultivar Balaton in summer of 2019 with fruit development stages labeled.



## REFERENCES


Adams KL, Cronn R, Percifield R, Wendel JF. 2003. Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. Proc Natl Acad Sci. 100(8):4649–4654. doi:10.1073/pnas.0630618100.

Akšić MF, Cerović R, Ercişli S, Jensen M. 2016. Microsporogenesis and meiotic abnormalities in different 'Oblačinska' sour cherry (*Prunus cerasus* L.) clones. Flora. 219:25–34. doi:10.1016/J.FLORA.2015.12.009.

Alger EI, Edger PP. 2020. One subgenome to rule them all: underlying mechanisms of subgenome dominance. Curr Opin Plant Biol. 54:108–113. doi:10.1016/j.pbi.2020.03.004.

Bardenhagen C, Rothwell N, Colella F, Nugent J. 2022. Michigan Tart Cherry Cost of Production Study 2022. Michigan State University Extension Report No.: E3474. [accessed 2023 Oct 21]. https://www.canr.msu.edu/cherries/uploads/files/Final-2022-Michigan-Tart-Cherries.pdf.

Barker MS, Arrigo N, Baniaga AE, Li Z, Levin DA. 2016. On the relative abundance of autopolyploids and allopolyploids. New Phytol. 210(2):391–398.

Bartkowiak, Jr. D. 2020 Jan 15. Michigan cherry growers lose in fight to place tariffs on Turkish exporters. Click Detroit. [accessed 2023 Oct 21]. https://www.clickondetroit.com/news/local/2020/01/15/michigan-cherry-growers-lose-in-fight-to-place-tariffs-on-turkish-exporters/.

Beaver JA, Iezzoni AF. 1993. Allyzyme Inheritance in Tetraploid Sour Cherry (*Prunus cerasus* L.). J Am Soc Hortic Sci. 118(6):873–877.

Benson CW, Sheltra MR, Maughan PJ, Jellen EN, Robbins MD, Bushman BS, Patterson EL, Hall ND, Huff DR. 2023. Homoeologous evolution of the allotetraploid genome of *Poa annua* L. BMC Genomics. 24(1):350. doi:10.1186/s12864-023-09456-5.

Birchler JA, Veitia RA. 2007. The Gene Balance Hypothesis: From Classical Genetics to Modern Genomics. Plant Cell. 19(2):395–402. doi:10.1105/tpc.106.049338.

Birchler JA, Veitia RA. 2012. Gene balance hypothesis: Connecting issues of dosage sensitivity across biological disciplines. Proc Natl Acad Sci. 109(37):14746–14753. doi:10.1073/pnas.1207726109.

Bird KA, Niederhuth CE, Ou S, Gehan M, Pires JC, Xiong Z, VanBuren R, Edger PP. 2021. Replaying the evolutionary tape to investigate subgenome dominance in allopolyploid *Brassica napus*. New Phytol. 230(1):354–371. doi:10.1111/nph.17137.

Bird KA, Pires JC, VanBuren R, Xiong Z, Edger PP. 2023. Dosage-sensitivity shapes how genes transcriptionally respond to allopolyploidy and homoeologous exchange in resynthesized *Brassica napus.* Genetics. 225(1):iyad114. doi:10.1093/genetics/iyad114.

Bird KA, VanBuren R, Puzey JR, Edger PP. 2018. The causes and consequences of subgenome dominance in hybrids and recent polyploids. New Phytol. 220(1):87–93. doi:10.1111/nph.15256.

Burns R, Mandáková T, Gunis J, Soto-Jiménez LM, Liu C, Lysak MA, Novikova PY, Nordborg M. 2021. Gradual evolution of allopolyploidy in *Arabidopsis suecica.* Nat Ecol Evol. 5(10):1367–1381. doi:10.1038/s41559-021-01525-w.

Cai L, Quero-García J, Barreneche T, Dirlewanger E, Saski C, Iezzoni A. 2019. A fruit firmness QTL identified on linkage group 4 in sweet cherry (*Prunus avium* L.) is associated with domesticated and bred germplasm. Sci Rep. 9(1):5008. doi:10.1038/s41598-019-41484-8.

Cai L, Stegmeir T, Sebolt A, Zheng C, Bink MCAM, Iezzoni A. 2018. Identification of bloom date QTLs and haplotype analysis in tetraploid sour cherry (*Prunus cerasus*). Tree Genet Genomes. 14(22). https://link.springer.com/content/pdf/10.1007%2Fs11295-018-1236-2.pdf.

Cai L, Voorrips RE, van de Weg E, Peace C, Iezzoni A. 2017. Genetic structure of a QTL hotspot on chromosome 2 in sweet cherry indicates positive selection for favorable haplotypes. Mol Breed. 37(7):85. doi:10.1007/s11032-017-0689-6.

Calle A, Balas F, Cai L, Iezzoni A, López-Corrales M, Serradilla MJ, Wünsch A. 2020. Fruit size and firmness QTL alleles of breeding interest identified in a sweet cherry 'Ambrunés' × 'Sweetheart' population. Mol Breed. 40(9):86. doi:10.1007/s11032-020-01165-1.

Calle A, Grimplet J, Le Dantec L, Wünsch A. 2021. Identification and Characterization of DAMs Mutations Associated With Early Blooming in Sweet Cherry, and Validation of DNA-Based Markers for Selection. Front Plant Sci. 12:1018. doi:10.3389/fpls.2021.621491.

Campoy JA, Lerigoleur-Balsemin E, Christmann H, Beauvieux R, Girollet N, Quero-García J, Dirlewanger E, Barreneche T. 2016. Genetic diversity, linkage disequilibrium, population structure and construction of a core collection of *Prunus avium* L. landraces and bred cultivars. BMC Plant Biol. 16(1):49. doi:10.1186/s12870-016-0712-9.

Charles D. 2013 May 27. A Hungarian Cherry Tree's Long Trek To Michigan. NPR. [accessed 2023 Oct 6]. https://www.npr.org/sections/thesalt/2013/05/27/186270141/a-hungarian-cherry-trees-long-trek-to-michigan.

Colle M, Leisner CP, Wai CM, Ou S, Bird KA, Wang J, Wisecaver JH, Yocca AE, Alger EI, Tang H, et al. 2019. Haplotype-phased genome and evolution of phytonutrient pathways of tetraploid blueberry. GigaScience. 8(3). doi:10.1093/gigascience/giz012. [accessed 2020 Oct 23]. https://academic.oup.com/gigascience/article/doi/10.1093/gigascience/giz012/5304886.

David KT. 2022. Global gradients in the distribution of animal polyploids. Proc Natl Acad Sci. 119(48):e2214070119. doi:10.1073/pnas.2214070119.

Diseases. Cherries. [accessed 2023 Oct 21]. https://www.canr.msu.edu/cherries/pest_management/diseases.

Douglas GM, Gos G, Steige KA, Salcedo A, Holm K, Josephs EB, Arunkumar R, Ågren JA, Hazzouri KM, Wang W, et al. 2015. Hybrid origins and the earliest stages of diploidization in the highly successful recent polyploid *Capsella bursa-pastoris.* Proc Natl Acad Sci. 112(9):2806–2811. doi:10.1073/pnas.1412277112.

Edger PP, Poorten TJ, VanBuren R, Hardigan MA, Colle M, McKain MR, Smith RD, Teresi SJ, Nelson ADL, Wai CM, et al. 2019. Origin and evolution of the octoploid strawberry genome. Nat Genet. 51(3):541–547. doi:10.1038/s41588-019-0356-4.

Edger PP, Smith R, McKain MR, Cooley AM, Vallejo-Marin M, Yuan Y, Bewick AJ, Ji L, Platts AE, Bowman MJ, et al. 2017. Subgenome Dominance in an Interspecific Hybrid, Synthetic Allopolyploid, and a 140-Year-Old Naturally Established Neo-Allopolyploid Monkeyflower. Plant Cell. 29(9):2150–2167. doi:10.1105/tpc.17.00010.

Faust M, Timon B, Surányi D, Nyujtó F, Gradziel TM. 2011. Origin and dissemination of *Prunus* crops: peach, cherry, apricot, plum and almond. Orig Dissem Prunus Crops Peach Cherry Apric Plum Almond. https://www.cabdirect.org/cabdirect/abstract/20123346667.

Gaebelein R, Schiessl SV, Samans B, Batley J, Mason AS. 2019. Inherited allelic variants and novel karyotype changes influence fertility and genome stability in *Brassica* allohexaploids. New Phytol. 223(2):965–978. doi:10.1111/nph.15804.

Gaeta RT, Chris Pires J. 2010. Homoeologous recombination in allopolyploids: the polyploid ratchet. New Phytol. 186(1):18–28. doi:10.1111/j.1469-8137.2009.03089.x.

Goeckeritz CZ, Rhoades KE, Childs KL, Iezzoni AF, VanBuren R, Hollender CA. 2023. Genome of tetraploid sour cherry (*Prunus cerasus* L.) 'Montmorency' identifies three distinct ancestral Prunus genomes. Hortic Res. 10(7):uhad097. doi:10.1093/hr/uhad097.

Hedrick UP. 1915. The Cherries of New York. Albany, NY: J.B. Lyon (Report of the New York Agricultural Experiment Station for the Year 1914).

Hillig KW, Iezzoni AF. 1988. Multivariate analysis of a sour cherry germplasm collection [PhD Thesis]. [American Society for Horticultural Science]. http://agris.fao.org/agris-search/search.do?recordID=US8901862.

Hruby, K. 1950. The cytology of tetraploid cherries. Stud Bot Cechoslov-. 11(3):87–97.

Hufford MB, Seetharam AS, Woodhouse MR, Chougule KM, Ou S, Liu J, Ricci WA, Guo T, Olson A, Qiu Y, et al. 2021. De novo assembly, annotation, and comparative analysis of 26 diverse maize genomes. Science. 373(6555):655–662. doi:10.1126/science.abg5289.

Iezzoni A, Schmidt H, Albertini A. 1991. CHERRIES (PRUNUS). Acta Hortic.(290):111–176. doi:10.17660/ActaHortic.1991.290.4.

Iezzoni AF. 2004. Acquiring cherry germplasm from Central and Eastern Europe. HortScience. 40(2):304–308.

Iezzoni AF, Sebolt AM, Wang D. 2005. Sour cherry breeding program at Michigan state university. Acta Hortic. 667:131–134. doi:10.17660/ActaHortic.2005.667.18.

de Jong GW, Adams KL. 2023. Subgenome-dominant expression and alternative splicing in response to Sclerotinia infection in polyploid *Brassica napus* and progenitors. Plant J. 114(1):142–158. doi:10.1111/tpj.16127.

Khan D, Ziegler DJ, Kalichuk JL, Hoi V, Huynh N, Hajihassani A, Parkin IAP, Robinson SJ, Belmonte MF. 2022. Gene expression profiling reveals transcription factor networks and subgenome bias during *Brassica napus* seed development. Plant J. 109(3):477–489. doi:10.1111/tpj.15587.

Li N, Xu C, Zhang A, Lv R, Meng X, Lin X, Gong L, Wendel JF, Liu B. 2019. DNA methylation repatterning accompanying hybridization, whole genome doubling and homoeolog exchange in nascent segmental rice allotetraploids. New Phytol. 223(2):979–992. doi:10.1111/nph.15820.

Li Z, McKibben MTW, Finch GS, Blischak PD, Sutherland BL, Barker MS. 2021. Patterns and Processes of Diploidization in Land Plants. Annu Rev Plant Biol. 72(1):387–410. doi:10.1146/annurev-arplant-050718-100344.

Longstroth M, Irish-Brown A. 2021 Mar 30. Freeze damage depends on tree fruit stage of development. Mich State Univ Ext. [accessed 2023 Oct 21]. https://www.canr.msu.edu/news/freeze_damage_depends_on_tree_fruit_stage_of_development.

Macková L, Vít P, Urfus T. 2018. Crop-to-wild hybridization in cherries-Empirical evidence from *Prunus fruticosa*. Evol Appl. 11(9):1748–1759. doi:10.1111/eva.12677.

Mandáková T, Li Z, Barker MS, Lysak MA. 2017. Diverse genome organization following 13 independent mesopolyploid events in *Brassicaceae* contrasts with convergent patterns of gene retention. Plant J. 91(1):3–21. doi:10.1111/tpj.13553.

Mason AS, Wendel JF. 2020. Homoeologous Exchanges, Segmental Allopolyploidy, and Polyploid Genome Evolution. Front Genet. 11. [accessed 2023 Jan 24]. https://www.frontiersin.org/articles/10.3389/fgene.2020.01014.

McRae L, Beric A, Conant GC. 2022. Hybridization order is not the driving factor behind biases in duplicate gene losses among the hexaploid *Solanaceae*. Proc R Soc B Biol Sci. 289(1985):20221810. doi:10.1098/rspb.2022.1810.

Michigan Agricultural Statistics 2021-2022, Fruit. USDA NASS Annual Statistical Bulletin. [accessed 2023 Oct 21]. https://www.nass.usda.gov/Statistics_by_State/Michigan/Publications/Annual_Statistical_Bulletin/stats22/Fruit.pdf.

Michigan Fruit Production 2022. 2023. East Lansing, MI: USDA NASS Report No.: NR-23-27. [accessed 2023 Jul 5]. https://www.nass.usda.gov/Statistics_by_State/Michigan/Publications/Current_News_Release/2023/nr2327mi.pdf.

MSU Professor Recognized for the Fruits of Her Labor – MSU Innovation Center Annual Report. [accessed 2023 Oct 23]. https://annualreport.msuinnovationcenter.com/editions/msu-professor-recognized-for-the-fruits-of-her-labor/.

Noble B. 2018 Oct 22. Michigan cherry industry sour over Turkish imports. Detroit News. [accessed 2023 Oct 21]. https://www.detroitnews.com/story/business/2018/10/22/michigan-cherry-industry-sour-over-turkish-imports/1526726002/.

Noble B. 2019 Jun 5. Michigan cherry processors request 650% tariff on Turkey. Detroit News. [accessed 2023 Oct 21]. https://www.detroitnews.com/story/business/2019/06/05/michigan-cherry-processors-tariff-turkey-dried-tart-cherries/3684775002/.

Olden EJ, Nybom N. 1968. On the Origin of *Prunus cerasus* L. Hereditas. 59(2-3):327–345. doi:10.1111/J.1601-5223.1968.TB02181.X.

Olmstead JW. 2006. Linkage map construction and analysis of fruit size in sweet and sour cherry [Doctor of Philosophy]. Michigan State University.

Ozkan H, Levy AA, Feldman M. 2001. Allopolyploidy-Induced Rapid Genome Evolution in the Wheat (*Aegilops–Triticum*) Group. Plant Cell. 13(8):1735–1747. doi:10.1105/TPC.010082.

Peace C, Bassil N, Main D, Ficklin S, Rosyara UR, Stegmeir T, Sebolt A, Gilmore B, Lawley C, Mockler TC, et al. 2012. Development and Evaluation of a Genome-Wide 6K SNP Array for Diploid Sweet Cherry and Tetraploid Sour Cherry. PLoS ONE. 7(12). doi:10.1371/journal.pone.0048305.

Pinosio S, Marroni F, Zuccolo A, Vitulo N, Mariette S, Sonnante G, Aravanopoulos FA, Ganopoulos I, Palasciano M, Vidotto M, et al. 2020 Jun 21. A draft genome of sweet cherry (*Prunus avium* L.) reveals genome-wide and local effects of domestication. Plant J.:tpj.14809. doi:10.1111/tpj.14809.

Ponce C, Kuhn N, Arellano M, Time A, Multari S, Martens S, Carrera E, Sagredo B, Donoso JM, Meisel LA. 2021. Differential Phenolic Compounds and Hormone Accumulation Patterns between Early- and Mid-Maturing Sweet Cherry (*Prunus avium* L.) Cultivars during Fruit Development and Ripening. J Agric Food Chem. 69(31):8850–8860. doi:10.1021/acs.jafc.1c01140.

Popovska M, Angelova E, Popovski B. 2005. MICROSPOROGENESIS OF SOUR CHERRIES IN THE SKOPJE REGION. Acta Hortic.(667):111–116. doi:10.17660/ActaHortic.2005.667.14.

Rice A, Šmarda P, Novosolov M, Drori M, Glick L, Sabath N, Meiri S, Belmaker J, Mayrose I. 2019. The global biogeography of polyploid plants. Nat Ecol Evol. 3(2):265–273. doi:10.1038/s41559-018-0787-9.

Schnable JC, Springer NM, Freeling M. 2011. Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. Proc Natl Acad Sci. 108(10):4069–4074. doi:10.1073/pnas.1101368108.

Schuster M, Wolfram B. 2004. RESULTS OF SOUR CHERRY BREEDING IN DRESDEN-PILLNITZ. Acta Hortic.(663):911–914. doi:10.17660/ActaHortic.2004.663.167.

Schuster M, Wolfram B. 2005. New sour cherry cultivars from Dresden-Pillnitz. Acta Hortic. 795 PART 1:83–86. doi:10.17660/ActaHortic.2008.795.6.

Shirasawa K, Isuzugawa K, Ikenaga M, Saito Y, Yamamoto T, Hirakawa H, Isobe S. 2017. The genome sequence of sweet cherry (*Prunus avium*) for use in genomics-assisted breeding. DNA Res. 24(5):499–508. doi:10.1093/dnares/dsx020.

Soltis DE, Visger CJ, Marchant DB, Soltis PS. 2016. Polyploidy: Pitfalls and paths to a paradigm. Am J Bot. 103(7):1146–1166. doi:10.3732/ajb.1500501.

Sun H, Wu S, Zhang G, Jiao C, Guo S, Ren Y, Zhang J, Zhang H, Gong G, Jia Z, et al. 2017. Karyotype Stability and Unbiased Fractionation in the Paleo-Allotetraploid *Cucurbita* Genomes. Mol Plant. 10(10):1293–1306. doi:10.1016/j.molp.2017.09.003.

Sun Y, Liu Y, Shi J, Wang L, Liang C, Yang J, Chen J, Chen M. 2023. Biased mutations and gene losses underlying diploidization of the tetraploid broomcorn millet genome. Plant J. 113(4):787–801. doi:10.1111/tpj.16085.

Surányi D. 2021. Comparative analysis of sour cherry cultivars on their ecological and biological indicators. Int J Hortic Sci. 27:7-28. doi:10.31421/ijhs/27/2021/9810.

The International Wheat Genome Sequencing Consortium (IWGSC). 2014. A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. Science. 345(6194):1251788. doi:10.1126/science.1251788.

Tsukamoto T, Hauck NR, Tao R, Jiang N, Iezzoni AF. 2010. Molecular and genetic analyses of four nonfunctional S haplotype variants derived from a common ancestral S haplotype identified in sour cherry (*Prunus cerasus* L.). Genetics. 184(2):411–27. doi:10.1534/genetics.109.109728.

Tukey HB, Young JO. 1939. Histological Study of the Developing Fruit of the Sour Cherry. Bot Gaz. 100(4):723–749.

Van de Peer Y, Ashman T-L, Soltis PS, Soltis DE. 2021. Polyploidy: an evolutionary and ecological force in stressful times. Plant Cell. 33(1):11–26. doi:10.1093/plcell/koaa015.

Van de Peer Y, Mizrachi E, Marchal K. 2017. The evolutionary significance of polyploidy. Nat Rev Genet. 18(7):411–424. doi:10.1038/nrg.2017.26.

VanBuren R, Man Wai C, Wang X, Pardo J, Yocca AE, Wang H, Chaluvadi SR, Han G, Bryant D, Edger PP, et al. 2020. Exceptional subgenome stability and functional divergence in the allotetraploid Ethiopian cereal teff. Nat Commun. 11(1):884. doi:10.1038/s41467-020-14724-z.

Vasudevan A, Lévesque-Lemay M, Edwards T, Cloutier S. 2023. Global transcriptome analysis of allopolyploidization reveals large-scale repression of the D-subgenome in synthetic hexaploid wheat. Commun Biol. 6(1):1–14. doi:10.1038/s42003-023-04781-7.

Verde I, Abbott AG, Scalabrin S, Jung S, Shu S, Marroni F, Zhebentyayeva T, Dettori MT, Grimwood J, Cattonaro F, et al. 2013. The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. Nat Genet. 45(5):487–494. doi:10.1038/ng.2586.

Verde I, Jenkins J, Dondini L, Micali S, Pagliarani G, Vendramin E, Paris R, Aramini V, Gazza L, Rossini L, et al. 2017. The Peach v2.0 release: high-resolution linkage mapping and deep resequencing improve chromosome-scale assembly and contiguity. BMC Genomics. 18(1):225. doi:10.1186/s12864-017-3606-9.

Walsh JR, Woodhouse MR, Andorf CM, Sen TZ. 2020. Tissue-specific gene expression and protein abundance patterns are associated with fractionation bias in maize. BMC Plant Biol. 20(1):4. doi:10.1186/s12870-019-2218-8.

Wang D, Karle R, Iezzoni AF. 2000. QTL analysis of flower and fruit traits in sour cherry. Theor Appl Genet. 100(3–4):535–544. doi:10.1007/s001220050070.

Wang, Dechun. 1998. RFLP Mapping, QTL Identification, and Cytogenetic Analysis in Sour Cherry [Doctor of Philosophy]. [East Lansing, MI]: Michigan State University.

Wang J, Liu W, Zhu D, Hong P, Zhang S, Xiao S, Tan Y, Chen X, Xu L, Zong X, et al. 2020. Chromosome-scale genome assembly of sweet cherry (*Prunus avium* L.) cv. Tieton obtained using long-read and Hi-C sequencing. Hortic Res. 7(1):1–11. doi:10.1038/s41438-020-00343-8.

Wang J, Tian L, Madlung A, Lee H-S, Chen M, Lee JJ, Watson B, Kagochi T, Comai L, Chen ZJ. 2004. Stochastic and Epigenetic Changes of Gene Expression in *Arabidopsis* Polyploids. Genetics. 167(4):1961–1973. doi:10.1534/genetics.104.027896.

Willing H. 1960. Phenological and chemical studies on cherry fruit development. Arch Gartenb. 8:561–94.

Wilson J. 2019. Tart Cherry Pest Guide. Michigan State University Extension. [accessed 2023 Oct 21]. https://www.canr.msu.edu/fruit/uploads/files/2019CherryPestGuides.pdf.

Wöhner TW, Emeriewen OF, Wittenberg AHJ, Nijbroek K, Wang RP, Blom E-J, Schneiders H, Keilwagen J, Berner T, Hoff KJ, et al. 2023. The structure of the tetraploid sour cherry 'Schattenmorelle' (*Prunus cerasus* L.) genome reveals insights into its segmental allopolyploid nature. Front Plant Sci. 14. [accessed 2023 Dec 5]. https://www.frontiersin.org/articles/10.3389/fpls.2023.1284478.

Xiong Z, Gaeta RT, Pires JC. 2011. Homoeologous shuffling and chromosome compensation maintain genome balance in resynthesized allopolyploid *Brassica napus*. Proc Natl Acad Sci. 108(19):7908–7913. doi:10.1073/pnas.1014138108.

Yoo S-D, Gao Z, Cantini C, Loescher WH, Nocker S van. 2003. Fruit Ripening in Sour Cherry: Changes in Expression of Genes Encoding Expansins and other Cell-wall-modifying Enzymes. J Am Soc Hortic Sci. 128(1):16–22. doi:10.21273/JASHS.128.1.0016.

# CHAPTER 2: GENOME OF TETRAPLOID SOUR CHERRY (*PRUNUS CERASUS* L.) 'MONTMORENCY' IDENTIFIES THREE DISTINCT ANCESTRAL *PRUNUS* GENOMES

The work presented in this chapter is part of the final publication:

Goeckeritz C.Z., Rhoades K.E., Childs K.L., Iezzoni A.F., VanBuren R. and Hollender CA. Genome of tetraploid sour cherry (*Prunus cerasus* L.) 'Montmorency' identifies three distinct ancestral *Prunus* genomes. *Hortic Res* 2023;10:uhad097. doi.org/10.1093/uhad097

Author contributions:

CAH, AFI, and RV conceptualized the experiments. CZG performed genome assembly, annotation, subgenome assignment using orthologs, and divergence time estimate analyses. KER performed the k-mer hierarchical clustering, Ks analysis, synteny, *DAM* gene phylogenetic analyses, and *S*-allele alignments. KLC provided expertise and assistance with annotation and contributed code. CZG, KER, and AFI wrote the manuscript. All authors assisted with editing the manuscript.

**Abstract**

Sour cherry (*Prunus cerasus* L.) is a valuable fruit crop in the Rosaceae family and a hybrid between progenitors closely related to extant *Prunus fruticosa* (ground cherry) and *Prunus avium* (sweet cherry). Here we report a chromosome-scale genome assembly for sour cherry cultivar Montmorency, the predominant cultivar grown in the USA. We also generated a draft assembly of *P. fruticosa* to use alongside a published *P. avium* sequence for syntelog-based subgenome assignments for 'Montmorency' and provide compelling evidence *P. fruticosa* is also an allotetraploid. Using hierarchal k-mer clustering and phylogenomics, we show 'Montmorency' is trigenomic, containing two distinct subgenomes inherited from a *P. fruticosa*-like ancestor (A and A') and two copies of the same subgenome inherited from a *P. avium*-like ancestor (BB). The genome composition of 'Montmorency' is AA'BB and little-to-no recombination has occurred between progenitor subgenomes (A/A' and B). In *Prunus*, two known classes of genes are important to breeding strategies: the self-incompatibility loci (*S*-alleles), which determine compatible crosses, successful fertilization, and fruit set, and the *Dormancy Associated MADS-box* genes (*DAMs*), which strongly affect dormancy transitions and flowering time. The *S*-alleles and *DAMs* in 'Montmorency' and *P. fruticosa* were manually annotated and support subgenome assignments. Lastly, the hybridization event 'Montmorency' is descended from was estimated to have occurred less than 1.61 million years ago, making sour cherry a relatively recent allotetraploid. The 'Montmorency' genome highlights the evolutionary complexity of the genus *Prunus* and will inform future breeding strategies for sour cherry, comparative genomics in the Rosaceae, and questions regarding neopolyploidy.

**Summary**

The state of Michigan produces over 70 percent of sour cherries in the U.S., the majority of which are Montmorency, a 450 year old French landrace first brought to North America by European settlers. Michigan State University currently houses the only sour cherry breeding program in the U.S. When I started my Ph.D. there was still not a reference genome for sour cherry, so a group of researchers led by my collaborator Charity Goeckeritz worked to assemble, annotate, and validate a reference genome of Montmorency to assist with breeding efforts and sour cherry genomics research. I performed syntenic analyses on the three subgenomes of sour cherry and its extant progenitor sweet cherry to validate the collinearity of the subgenomes with each other and the progenitor genome. I used Ks analysis to determine that *P. fruticosa,* one of the extant progenitors of sour cherry, is likely an allopolyploid. I also performed k-mer hierarchical clustering analysis to separate the three subgenomes of sour cherry and used a set of dormancy-associated MADS box (DAM) genes to validate the ancestry of the three subgenomes and demonstrate the utility of the genome for sour cherry breeding and research. CZG, AIF, and I wrote the manuscript, and all authors participated in the editing process. CZG and I generated all figures. AIF, CAH, and RV conceptualized the work.

# CHAPTER 3: LANDRACE AND BRED ACCESSIONS OF ALLOTETRAPLOID SOUR CHERRY (*PRUNUS CERASUS* L.) REVEAL VARIATION IN SUBGENOME DOSAGE AND SUBGENOME EXPRESSION BIAS

The work presented in this chapter is being prepared for the publication:

Rhoades K.E.B., Goeckeritz C.Z., Bird K.A., Yocca A., Edger P.P., Iezzoni A.F. Landrace and bred accessions of allotetraploid sour cherry (*Prunus cerasus* L.) reveal variation in subgenome dosage and subgenome expression bias. *In prep.*

Author contributions:

KEBR, AIF, and PPE designed the experiments. KEBR completed all experiments and data analysis. CZG, KAB, AY, and PPE provided expertise and guidance on data analysis and interpretation. KEBR wrote the manuscript. CZG, AIF, and KEBR edited the manuscript.

**Abstract**

Subgenome dominance is a common, but not given, aspect of allopolyploidy that can occur when different genomes share space within the same nucleus. Subgenome dominance can present as preferential retention over generations of one subgenome through homoeologous exchange or biased fractionation, or through subgenome expression bias, where one subgenome is preferentially expressed over its counterpart despite equivalent gene dosage. The extent of subgenome dominance differs by clade, and variation in subgenome bias has been reported even among tissues of the same plant. Sour cherry (*Prunus cerasus*) is an allotetraploid fruit tree species resulting from an interspecific cross between ground cherry *P. fruticosa* and sweet cherry *P. avium*, although the actual ancestors may be an extant relative of either species. Recent work shows sour cherry cultivar Montmorency contains three subgenomes. A and A', each present in one copy, that are derived from a *P. fruticosa*-like ancestor and B, present in two copies, that is derived from a *P. avium*-like ancestor. This work investigates the subgenome dynamics of the three subgenomes of sour cherry in four diverse landrace and two bred accessions, including 'Montmorency'. We found evidence of 26 homoeologous exchange events and five whole-homoeolog replacements relative to 'Montmorency' in three of the six accessions. We also present evidence of subgenome expression bias favoring the A and A' subgenomes over the B subgenome, the magnitude of which differs between accessions and changes over the course of fruit development. Lastly, we illustrate the consequences of this dosage variation and expression bias for four previously-described genes associated with fruit softening in 'Montmorency'.

## Introduction

Allopolyploidy, the presence of three or more haploid sets of chromosomes derived from a cross between two different species, is a common phenomenon in plants (Barker et al. 2016). Newly-formed polyploids are observed to be genetically unstable, owing largely to the challenge of re-establishing stable meiosis after an allopolyploidy event (Mason and Wendel 2020). Over time, allopolyploids can achieve a diploid-like genome state that facilitates proper chromosome pairing through a combination of processes including genome rearrangement through homoeologous exchange (HE) and subgenome expression bias that eventually results in loss of genes from one subgenome (Li et al. 2021). While the strategies of diploidization are common among allopolyploids, the sequence of these processes vary by clade (The International Wheat Genome Sequencing Consortium (IWGSC) 2014; Cheng et al. 2018; Colle et al. 2019; Edger et al. 2019; Hufford et al. 2021; McRae et al. 2022; Zhuang et al. 2022; Sun et al. 2023).

Sour cherry (*Prunus cerasus* L., $2n = 2x = 32$), a perennial fruit crop native to Eastern Europe and Western Asia, is well-documented to exhibit meiotic irregularities that manifest as low fertility and poor fruit set (Hruby, K. 1950; Wang, Dechun 1998; Iezzoni et al. 2005; Akšić et al. 2016), suggesting that the species is still in the early stages of diploidization. As the species is most commonly propagated vegetatively, rather than through seed, landrace accessions that were maintained by vegetative propagation by humans would not be subject to escaped some of the natural selection pressure of poor fertility. Sour cherry is an allotetraploid resulting from an interspecific hybridization between an ancestor of the extant diploid sweet cherry (*P. avium* L., $2n = 2x = 16$) and an ancestor of the extant allotetraploid ground cherry (*P. fruticosa* Pall., $2n = 4x = 32$) where their native distributions overlap in Eastern Europe and Western Asia (Olden and Nybom 1968; Faust et al. 2011). Chloroplast data suggests this hybridization event occurred

multiple times, and that *P. fruticosa* was most commonly the seed parent (Iezzoni and Hancock 1996; Bird et al. 2022).

Studies of meiosis in sour cherry show varying rates of multivalent chromosome pairing depending on genotype, and occasionally the complete loss of chromosomes in telophase when they fail to pair or attach to spindle fibers (Hruby, K. 1950; Wang, Dechun 1998; Schuster and Wolfram 2004; Akšić et al. 2016). Sour cherry also exhibits disomic and tetrasomic patterns of inheritance, further supporting occasional pairing between homoeologs at meiosis (Beaver and Iezzoni 1993). This meiotic instability supports the finding of Goeckeritz et al. (2023) that sour cherry is a relatively young allopolyploid.

The first reference genome for sour cherry (cv. Montmorency) was recently released, providing the opportunity to compare subgenome dynamics in diverse sour cherry accessions (Goeckeritz et al., 2023). In this work, it was proposed that *P. fruticosa* is an allotetraploid that passed one copy of each of its subgenomes to sour cherry. As a result, the 'Montmorency' reference genome consists of three sets of eight chromosomes representing the two subgenomes inherited from the *P. fruticosa* progenitor (denoted as A and A') and the one subgenome inherited from the *P. avium* progenitor (denoted as B; Goeckeritz et al. 2023). Subgenome dosage in sour cherry may further be complicated by continued introgression that has been documented between sour cherry and it's progenitor species (Hillig and Iezzoni 1988). This reference genome provides a foundation for examining subgenome dominance, both structural and expression-based, of Montmorency and other sour cherry accessions.

In this study, we report (i) genome-wide subgenome dosage variation among four diverse landrace and two bred sour cherry accessions, (ii) patterns of subgenome expression bias in five tissues from those same accessions, and (iii) the consequences of this dosage variation and

expression bias for four previously-described genes associated with fruit softening in Montmorency. To aid in validation of our subgenome dosage results we use the previously characterized S-alleles in our material and the progenitor species of sour cherry. The fruit softening-associated genes studied, four expansins first characterized from 'Montmorency' fruit by Yoo et al. (2003), are located in areas of differential subgenome dosage among the accessions and therefore illustrate the consequences of subgenome dosage changes and subgenome expression bias during fruit development and ripening.

**Materials and Methods**

*Plant materials and tissue collection*

Plant materials are listed in Table 3.1, and tissue collection is outlined in Figure S1. All tissue collected was flash-frozen in liquid nitrogen and stored at -80C until extraction. To minimize the effects of circadian rhythm on our subgenome expression bias analyses, we collected each accession within the same three hour window each collection day. Young leaves were collected for DNA and RNA extraction. Whole flowers at "balloon" stage were collected for RNA extraction. For RNA extraction, developing fruit from each accession were collected weekly starting at anthesis, and the widest diameter of the fruit to create a fruit growth curve for each accession was recorded. Three stages of fruit development for RNA sequencing for each accession were selected based on the fruit growth curve (Supplemental Figure S1).

*DNA and RNA extraction and sequencing*

DNA was extracted using the Qiagen Plant DNeasy kit (www.qiagen.com) according to manufacturer's instructions. Libraries were prepared by the MSU Research Technology Support Facility (RTSF) genomics core using the Illumina TruSeq Nano DNA Library prep kit according to manufacturer's instructions (www.illumina.com). Short-read sequencing (paired-end, 150bp)

was also performed at the MSU RTSF genomics core on an Illumina HiSeq4000 to a depth of ~40x per sample.

RNA was extracted with a CTAB method based on Gasic et al. (Gasic et al. 2004). RNA libraries were prepared with the Illumina TruSeq Total mRNA kit according to manufacturer's instructions and sequenced 2x150bp on a HiSeq 4000 at the MSU RTSF to a target of 35 million reads per library.

*Assigning subgenome dosage*

Illumina 2x150bp DNA reads for each accession were trimmed for quality using Trimmomatic v0.38 and aligned to the 24 scaffolded chromosomes of the 'Montmorency' reference using BWA mem v0.7.17 on default settings (Li 2013 Mar; Bolger et al. 2014). The resulting sequence alignment file was sorted with SAMtools v1.15 and PCR duplicates were marked with GATK markduplicatesspark v4.0 (Van der Auwera and O'Connor 2020; Danecek et al. 2021)**.** SAMtools depth was then used to call sequence depth values for each base pair of the 'Montmorency' reference where the aligned read base quality was greater than 20 and the mapping quality greater than 20. These depth values were then averaged over 1 kb windows on each chromosome to obtain average sequence depth. Sites of HE were assigned based on reciprocal changes in sequence coverage between homoeologs. Sequence depth values averaged over 1 Mb windows for each chromosome for each accession were graphed in R using package ggplot2 to create Figure 3.1 and Figure S2 (Wickham 2016; R Core Team 2021). Sudden spikes and drops in chromosome coverage that were not mirrored in another homoeolog likely represent differences in repeat content between the accession and the reference assembly. Cases where dosage was consistently at 3x for one homoeolog and 0x for another across the whole chromosome, was categorized as a chromosome replacement of one homoeolog with another.

Karyotypes (Figure 3.2) were created to scale based on the findings in Figures 3.1 and S2 using Inkscape v1.0.1 (https://inkscape.org).

*Gene expression counts and dosage normalization to compare the A/A', A/B, and A'/B subgenomes*

RNA sequence reads were aligned to the 24 chromosome Montmorency genome using STAR v2.7.9a and quantified using StringTie v2.1.3 on expression estimation mode with multi-mapping correction (Dobin et al. 2013; Shumate et al. 2022). Stringtie abundance files for each sample were imported into R v4.2.2 and combined with the homoeolog dosage/pair assignment files using tidyverse v2.0.0 packages (Wickham et al. 2019). All homoeolog pairs with a combined transcripts per million (TPM) less than 10 were discarded. In order to avoid log2 fold change (log2FC) values of +/- infinity (R designates any value divided by zero as infinity), +0.01 was added to each TPM value of the remaining genes prior to log transformation. Gene expression counts for all three tissue replicates for each tissue type were combined, only including homoeolog pairs found to be expressed in all replicates, and average TPM for each gene was calculated. We then divided the average TPM for each gene by its dosage, e.g. for a gene with an average TPM of 30 and a dosage of 2, the TPM value used for homoeolog comparisons would be 15.

*Subgenome expression bias in 1:1 subgenome comparisons (A/A', A/B, A'/B)*

Using the dose-normalized average TPM values, log2FC was calculated as log2(TPM gene2/TPM gene1). In all results a negative log2FC value indicates that the first homoeolog in the pair is more highly expressed, and a positive log2FC value indicates that the second homoeolog in the pair is more highly expressed. Comparisons were done for the A/A' homoeolog pairs, A/B homoeolog pairs, and the A'/B homoeolog pairs set.

*Subgenome expression bias between the AA' and B homoeolog sets*

To compare expression between both *P. fruticosa*-derived homoeologs and the *P. avium*-derived homoeologs we selected gene pairs that were present in the A/A' homoelog pair set and then also present in the A/B homoeolog pair set. As described above, only homoeolog pairs that were expressed in all three tissue replicates were included. In cases where the homoeolog genotype was AA'BB, the TPM sum for the A + A' homoelogs and the TPM for B homoeologs were dose-normalized by dividing both by by 2. In cases where the homoeolog genotype was ABBB or A'BBB, the sum for the A + A' homoeologs was dose-normalized by dividing by 1 and the TPM for the B homoeolog was dose-normalized by dividing by 3. We then calculated the log2FC between the combined A/A' homoeologs' TPM and the B homoeolog TPM as log2FC(TPM geneB/TPM pairAA').

*Statistical testing of subgenome expression bias datasets*

Analysis of variance on log2FC values was performed for all tissues per accession and all accessions per tissue type, then Tukey's HSD was used to assign significance groups to each tissue or each accession.

*Comparison of commonly-biased homoeolog pairs*

Sets of homoeolog pairs with log2FC < -3.5 (AA'-biased) or log2FC > 3.5 (B-biased) for each tissue in each accession were compared across accessions using tools in the tidyverse R package (Wickham et al. 2019). Upset plots were created with R package UpSetR (Conway et al. 2017).

*Expansin location on chromosomes and phylogeny to determine progenitor relationships*

We identified the expansins from Yoo et al. (2003) in the Montmorency reference with a BLAST+ v2.7.1 search of the cds sequences (Camacho et al. 2009). Coding sequence for each

gene was taken from the *P. avium* cv. Tieton v2.0 (Wang et al. 2020), Montmorency v1.0, and Goeckeritz et al. *P. fruticosa* reference genomes as well as an outgroup non-*Prunus* ortholog (NCBI ID: 103447269) and aligned with MUSCLE v3.8.31 (Edgar 2004). Phylogenies were constructed with RAxML-NG v1.0 using the GTR+G algorithm and 500 bootstrap replicates to create a consensus tree with all branches supported in at least 80 percent of replicates (Kozlov et al. 2019). Tree figures were generated using the ETE Tree Viewer (etetoolkit.org). EXP4 AA and cds alignments were completed with MUSCLE v3.8.31 and figures were generated using R package ggmsa (Zhou et al. 2022).

**Table 3.1:** Sour cherry accessions included in this work, their geographic origins, and their *S*-allele genotypes. Nine ancestral *S*-alleles have previously been identified in the six sour cherry accessions included in this study (Sebolt et al. 2017). Five of these *S*-alleles (S1, S4, S6, S9, and S13) have only been identified previously in sweet cherry (Sonneveld et al. 2001; Sonneveld et al. 2003; Schuster 2017; Kivistik et al. 2022). Goeckeritz et al. (2023) identified S35 and S36a in *P. fruticosa*. S26 and S34 have only been identified in sour cherry, so they are presumably derived from *P. fruticosa*. Sweet cherry-derived *S*-alleles are in blue text, *P. fruticosa*-derived *S*-alleles are in red text.

| Accession | Origin | *S* allele genotype (citation) |
|---|---|---|
| Montmorency | ~ 500-year-old landrace from France | $S_6S_{13m}S_{35}S_{36a}$ (Tsukamoto et al. 2006) |
| English Morello[a] | Landrace - Germany and neighboring regions | $S_6S_{13'}S_{26}S_{36a}$ (Tsukamoto et al. 2006) |
| Oblačinska | Landrace - Serbia and neighboring regions | $S_{6m2}S_9S_{26}S_{36b2}$ (Sebolt et al. 2017) |
| Balaton[b] | Self-compatible sport of Pándy; Landrace - Hungary and neighboring regions | $S_{1'}S_4S_{35}S_{36b}$ (Hauck et al. 2006) |
| Tamaris | Bred cultivar – Russia | $S_{13'}S_{16}S_{34}S_{36b2}$ (Tsukamoto et al. 2010) |
| Érdi Jubileum | Bred cultivar – Hungary (Pándy × Nagy Angol) | $S_1S_6S_{13'}S_{36b}$ (Tsukamoto et al. 2010) |

[a]syn. Schattenmorelle
[b]syn. Újfehértói Fürtös

## Results

*Predicting chromosome replacements and sites of homoeologous exchange*

When short reads were aligned back to the Montmorency reference genome, differences in chromosome dosage and possible sites of HE were identified in three of the five accessions based on reciprocal changes in read coverage. Figures 3.1 and S2 show short-read coverage graphs used to determine dosage, and Figure 3.2 shows stylized "karyotypes" of final chromosome dosage and predicted HEs for each accession. As expected, the chromosome dosage for Montmorency is consistently 1A:1A':2B. Balaton and Oblacinska also exhibit a reference-level dosage of all homoeologs. In the remaining accessions (English Morello, Erdi Jubileum, and Tamaris) we identified a total of 26 HE events: eight in English Morello, eight in Erdi Jubileum, and 12 in Tamaris (Table S1A). All of these appear to be unique events, with the exception of exchanges between chromosomes 6A and 6A' and chromosomes 2A and 2A' that are common to both English Morello and Tamaris (Table S1B). The homoeolog dosage determined for the *S*-locus region of chromosome 6 for each accession is consistent with the previously reported *S*-alleles, including Erdi Jubileum, which has three *S*-alleles derived from *P. avium* (Table 3.1, Figure 3.2).

Twenty of the 26 HE events are between the A and A' subgenomes (Table S1A). Despite the high number of HE events in these three accessions, all accessions maintain an overall subgenome dosage ratio of 1 A:1 A':2 B (Figure S3). Erdi Jubileum and Tamaris, the only two bred cultivars, are the only two accessions to exhibit whole-chromosome replacement, with Erdi Jubileum having three copies each of chromosomes 3B, 6B, 7B, and 8B and Tamaris having two copies of chromosome 7A' and two copies of chromosome 7B. In both English Morello and Erdi Jubileum, chromosomes 8A and 8A' show ambiguous sequence coverage that made assigning

homoeolog dosage difficult (Figure S2b, c). Erdi Jubileum short read coverage indicates a ~3x dosage of chromosome 8B and a ~0.5x dosage each of chromosomes 8A and 8A'. English Morello short read coverage indicates a ~2x dosage of chromosome 8B, a ~1.5x dosage of chromosome 8A', and a ~0.5x dosage of chromosome 8A. Goeckeritz et al. (2023) noted some difficulty in assembling chromosomes 8A and 8A' due to possible sequence similarity and ambiguity of the Hi-C matrix, so in Figure 3.2 we label one copy of chromosome 8 as *A* for each accession to denote the uncertainty of homoeolog identity.

**Figure 3.1:** DNA sequence coverage graphs for Montmorency and Tamaris, generated by aligning 2x150bp short reads to the 'Montmorency' reference and averaging coverage over 1Mb windows. The 1x, 2x, and 3x dosage lines on the y axis correspond to 20x, 40x and 60x coverage, respectively. The predicted boundaries of the centromeric regions are denoted with black lines and were taken from Goeckeritz et al. (2023). Red, yellow and blue lines denote A, A' and B homoeolog coverage, respectively. The coverage graphs for the other four accessions are in Supplementary Figure S2.

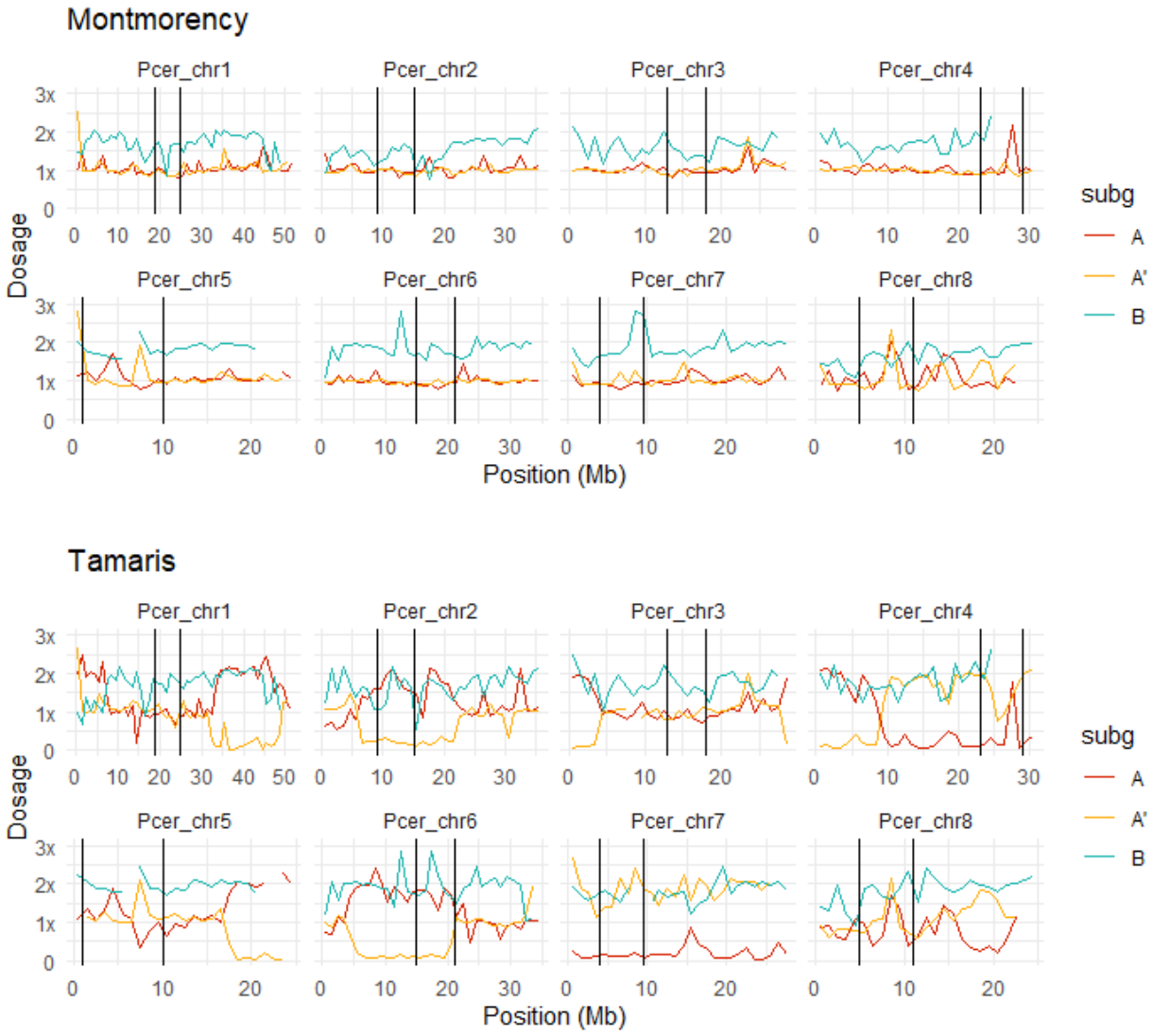
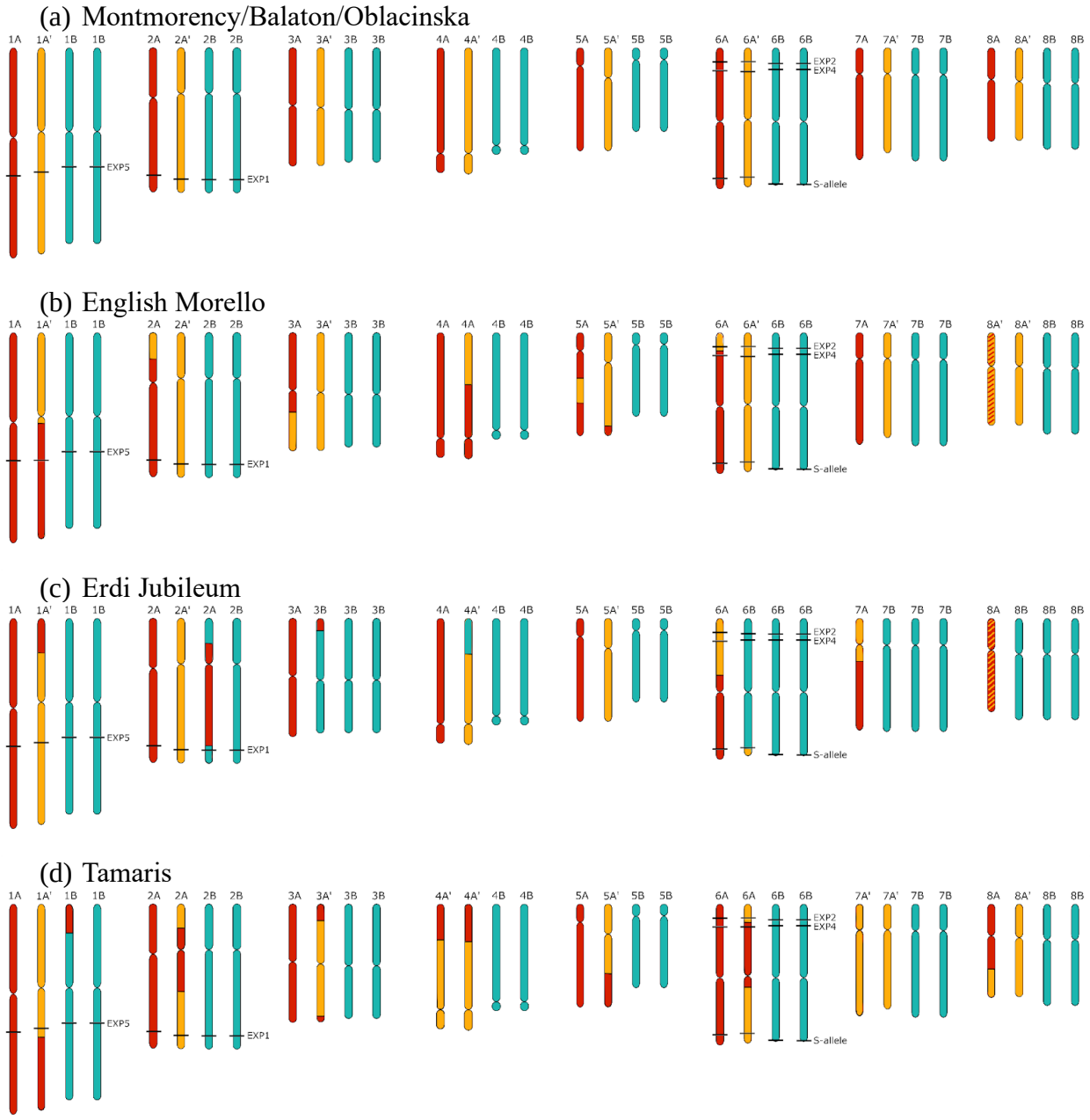


**Figure 3.2:** Karyotype illustration of subgenome content and homoeologous exchange sites for each accession based on the coverage shown in Fig. 1 and Fig. S2. The locations of the four expansins first characterized by Yoo et al. (2003) are marked, as well as the location of the *S*-locus for each homoeolog. The A, A' and B homoeologs are colored red, yellow and blue, respectively. Chromosome labels are based on the subgenome identity of the centromeric region. Homoeologs with ambiguous identity (English Morello chr. 8 and Erdi Jubileum chr. 8) are given striped coloring of the two subgenomes contributing to their makeup.



*Subgenome expression bias*

The pairwise comparisons of the expression of homoeologs, normalized for dosage, indicate the A and A' subgenomes appear fairly balanced while the A/B and A'/B comparisons show an overall expression bias towards the A and A' genomes respectively (Table 3.2, Figures S4, S5, S6). The overall median log2FC for A/B comparisons was -0.139, and the overall median log2FC for A'/B comparisons was -0.122 (Table 3.2, Table S2). This bias is consistent across tissues and accessions. When we treat the A and A' subgenomes as one, still normalizing for dosage, the expression bias is strongly in favor of AA', with an overall median log2FC of -0.234 (Table 3.2, Table S2). This is consistent across all tissues and accessions in this study (Figure 3.3). When comparing accessions by tissue type, fruit stages 1 and 2 show the most variation in the relative magnitude of the AA' bias (Figure 3.3). Within individual genotypes, the magnitude of AA' homoeolog bias changes significantly over the course of fruit development, peaking at different development stages depending on accession (Figure 3.3, Table S2). For example, Tamaris AA'/B homoeolog comparisons have a median log2FC of -0.66 in fruit stage 1, -0.627 in fruit stage 2, and -0.22 in fruit stage 3 (Figure 3.3, Table S2). Non-reference level homoeolog pairs resulting from homoeolog exchange or chromosome replacement did not show a significant difference in subgenome expression bias when compared to reference-level dosage homoeolog pairs.

When comparing which homoeolog pairs are significantly biased towards AA' between accessions, we found that the majority of biased homoeologs are unique to each genotype, but the next-largest group of homoeologs after those unique to each individual, between 20 and 27 pairs, is the group of AA'-biased homoeolog pairs shared among all accessions for that tissue (Figure 3.4). These homoeolog sets were not large enough to perform any meaningful gene

ontology enrichment analysis. Part of this may be attributable to the high log2FC threshold of +/- 3.5 for a homoeolog pair to qualify as biased. The Montmorency genome annotations for these homoeologs can be found in Table S3. Between zero and seven homoeolog pairs were found to be biased in favor of the B subgenome in any given sample (Table S2). Within that set we identified several genes cell wall-modifying enzymes that are biased in developing fruit (Table S4), but the set is also small enough that we cannot draw any meaningful conclusions about subgenome expression bias patterns in fruit development.

**Table 3.2:** Overview of subgenome expression bias for homoeolog pairs for four comparisons: A vs. A', A vs. B, A' vs. B and AA' vs B. Values represent medians across all tissues and accessions, and the median log2FC values for each set.

| Comparison (gene1/gene2) | Median number of pairs in analysis | Median log2FC of all pairs in analysis |
|---|---|---|
| A/A' | 5321 | -0.045 |
| A/B | 7050 | -0.139 |
| A'/B | 6551 | -0.122 |
| AA'/B | 6084 | -0.234 |

**Figure 3.3:** Dot plots of log2FC values for dosage-normalized AA' versusB homoeolog pair expression comparisons by tissue. AA' and B homoeolog pairs with a log2FC less than -3.5 or greater than 3.5, respectively, are colored according to the subgenome towards which they are biased. Median log2FC values for biased homoeolog groups and all homoeolog pairs are indicated in black. Significance groups based on analysis of variance and Tukey's HSD of log2FC values are indicated by the letter above each dotplot, with "a" representing the least amount of expression bias and subsequent letters denoting higher levels of significance in expression bias.

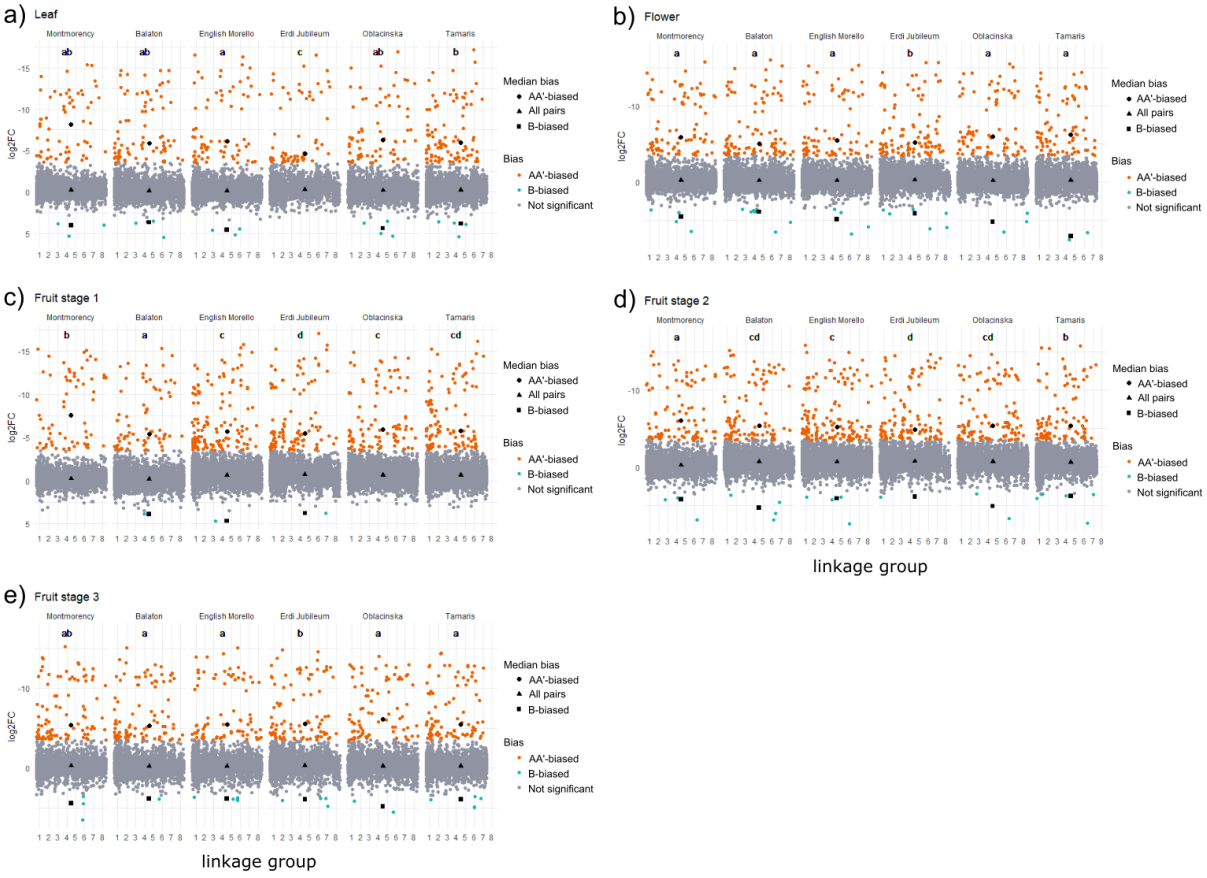


**Figure 3.3 (cont'd)**

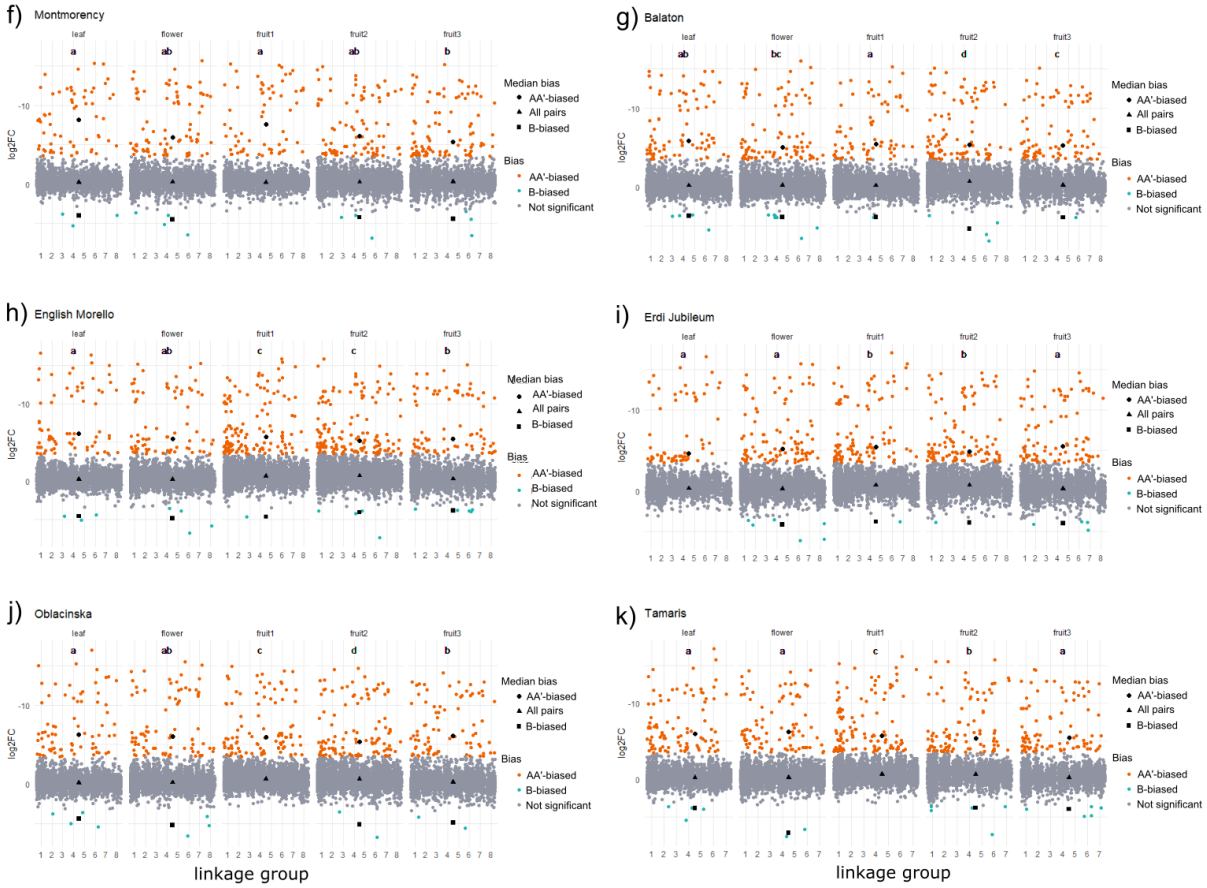
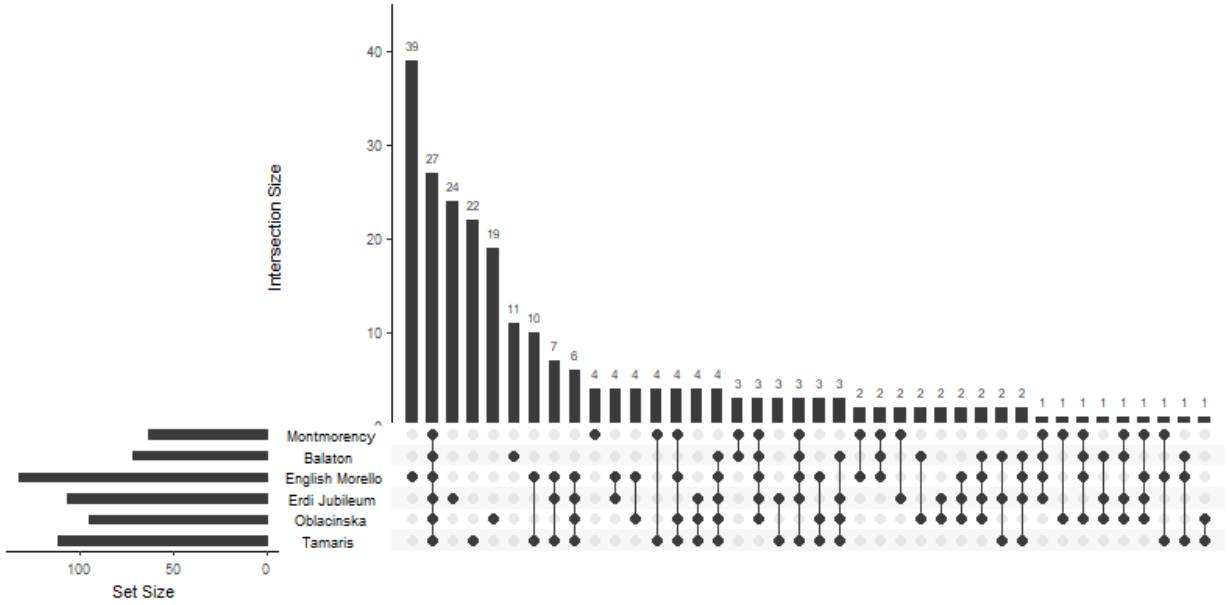


**Figure 3.4:** UpSet plots showing commonly AA'-biased homoeolog pairs between accessions for each tissue type for AA'/B subgenome comparison. The horizontal bars represent the total number of AA'-biased homoeolog pairs for each accession, and the vertical bars represent the number of AA'-biased homoeolog pairs shared among all accessions will filled-in black dots for that column.

(a) Fruit stage 1 AA'-biased




(b) Fruit stage 2 AA'-biased

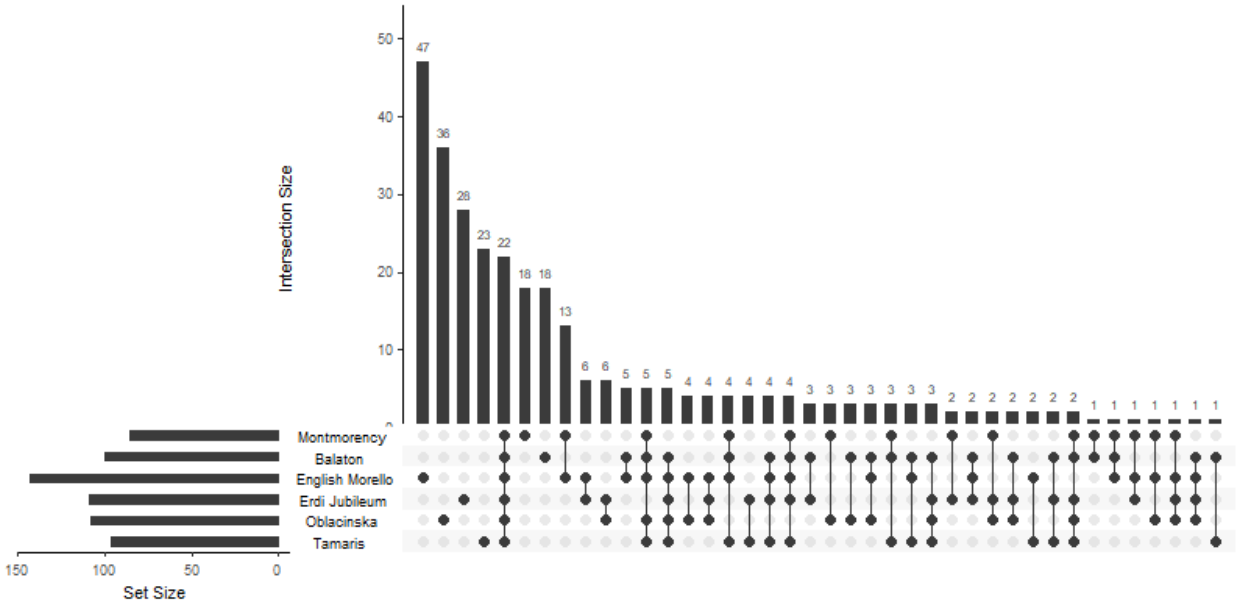
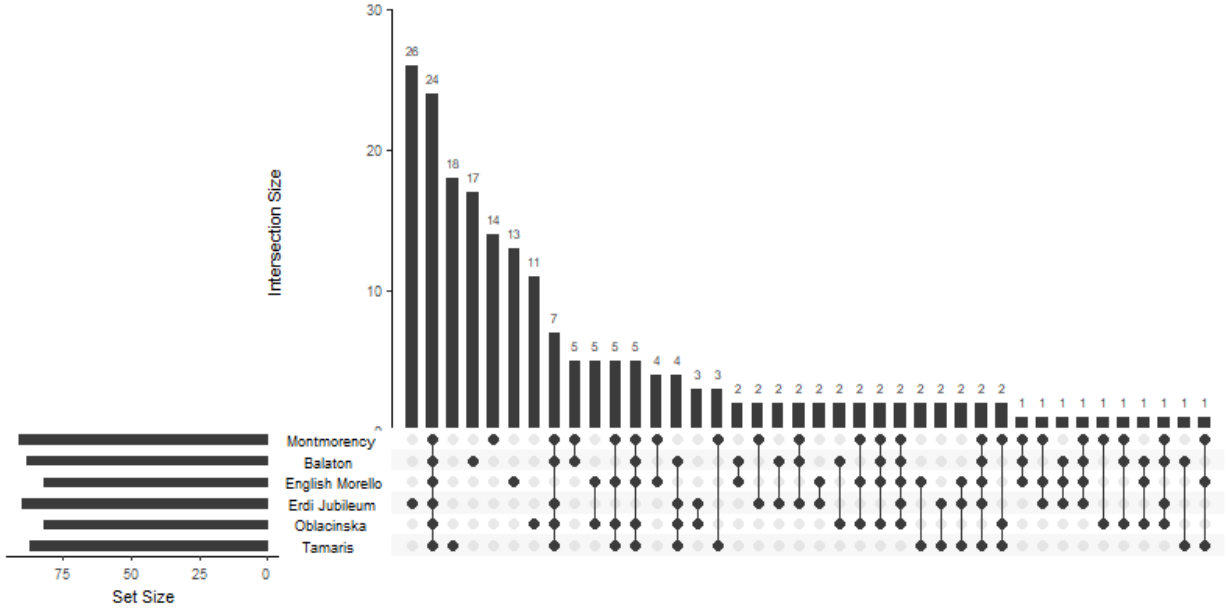


**Figure 3.4 (cont'd)**

(c) Fruit stage 3 AA'-biased




(d) Leaf AA'-biased

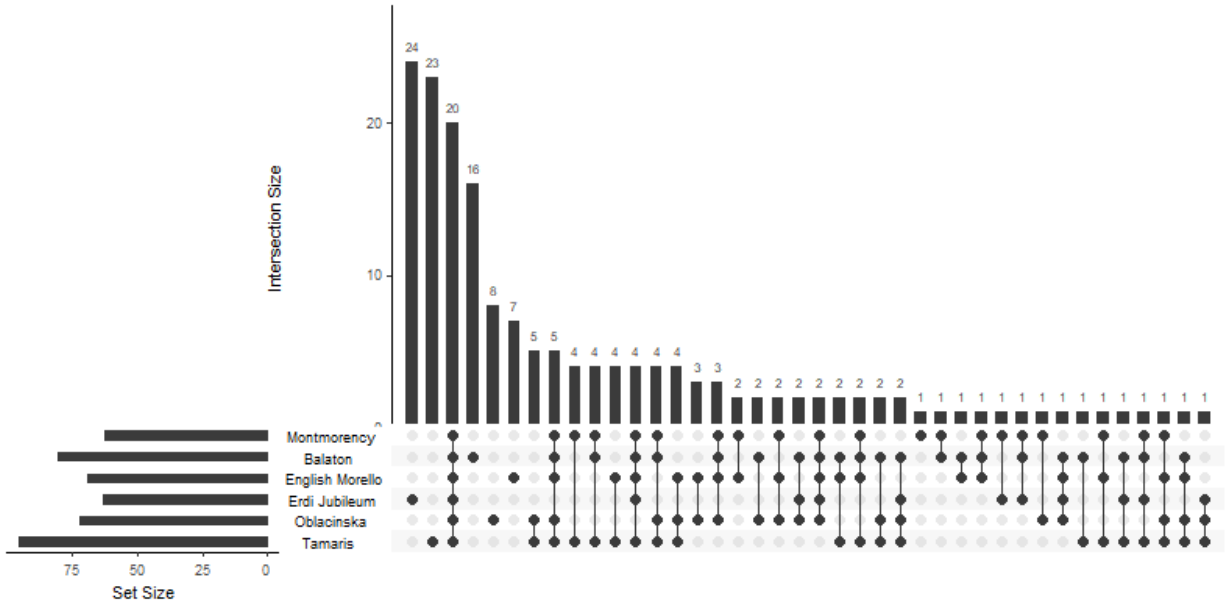
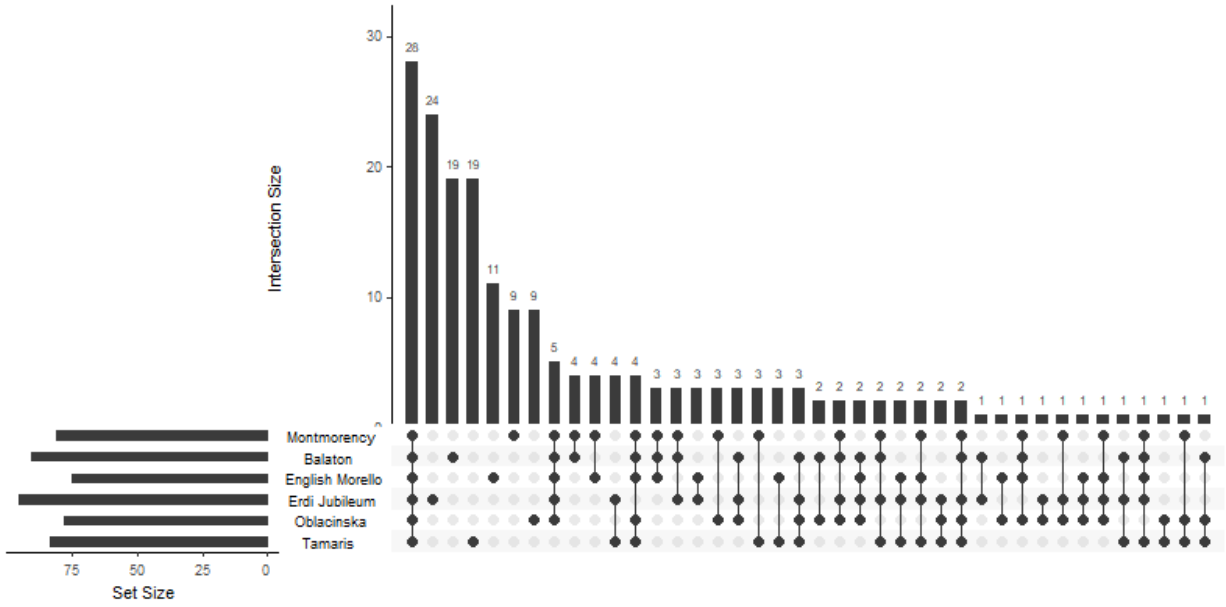


**Figure 3.4 (cont'd)**

(e) Whole flower AA'-biased



*Implications of polyploidy for expansins associated with fruit ripening*

A coding sequence phylogeny of the expansin homoeologs in Montmorency shows what we expect: homoeologs on the A and A' subgenomes cluster closest to their orthologs in *P. fruticosa* and homoeologs on the B subgenome cluster closest to their respective orthologs in *P. avium* (Figure 3.5). The sole exception is the EXP4 on chromosome 6A, which is equally distant to both progenitors. A closer examination of the coding sequence of the chromosome 6A EXP4 shows it has several unique variants not shared with either progenitor, so technically we cannot assign a progenitor to this homoeolog (Figure S6). EXP1, EXP2, and EXP5 exhibit the same pattern of expression in Montmorency fruit in our RNAseq experiment as in Yoo's original qPCR work (Figure S7), so we used RNAseq data from fruit stage 3 only to evaluate dosage and expression of the expansin homoeologs. EXP4 shows comparatively lower expression overall, which we attribute to having missed peak expression with our fruit stage 3 sampling, as Yoo et al. shows EXP4 expression dropping off at the very end of fruit development (2003).

The expansins serve as a useful demonstration of how changes in homoeolog dosage and the presence of subgenome expression bias can affect genes potentially of interest to breeders. Figure 3.6 and Figure S8 show DNA coverage, RNA coverage and TPM counts for each homoeolog for each accession. As expected, homoeologs predicted to be missing due to dosage changes show negligible expression, and the homoeologs with higher predicted dosage have higher expression. The chromosome 6A' EXP2 homoeolog shows unexpectedly low expression for its dosage (Figure 3.6a). There are no variants within the locus itself or its promoter region that would explain lowered expression, so we believe this may be attributable to subgenome expression bias. Indeed, EXP2 is significantly biased towards the B subgenome in fruit stage 3 of both English Morello and Erdi Jubileum (Table S4).

**Figure 3.5:** Phylogeny of the coding sequences for the homoeologs of four Montmorency expansin genes first characterized in Yoo et al. (2003) and their corresponding orthologs in sweet cherry (Tieton v2.0 reference genome, Wang et al. 2020) and *P. fruticosa* (Goeckeritz et al., 2023). The subgenome location of each expansin homoeolog (A, A' or B) is identified and color coded. An expansin from apple (*Malus domestica*) is included as an outgroup (NCBI # 103447269).

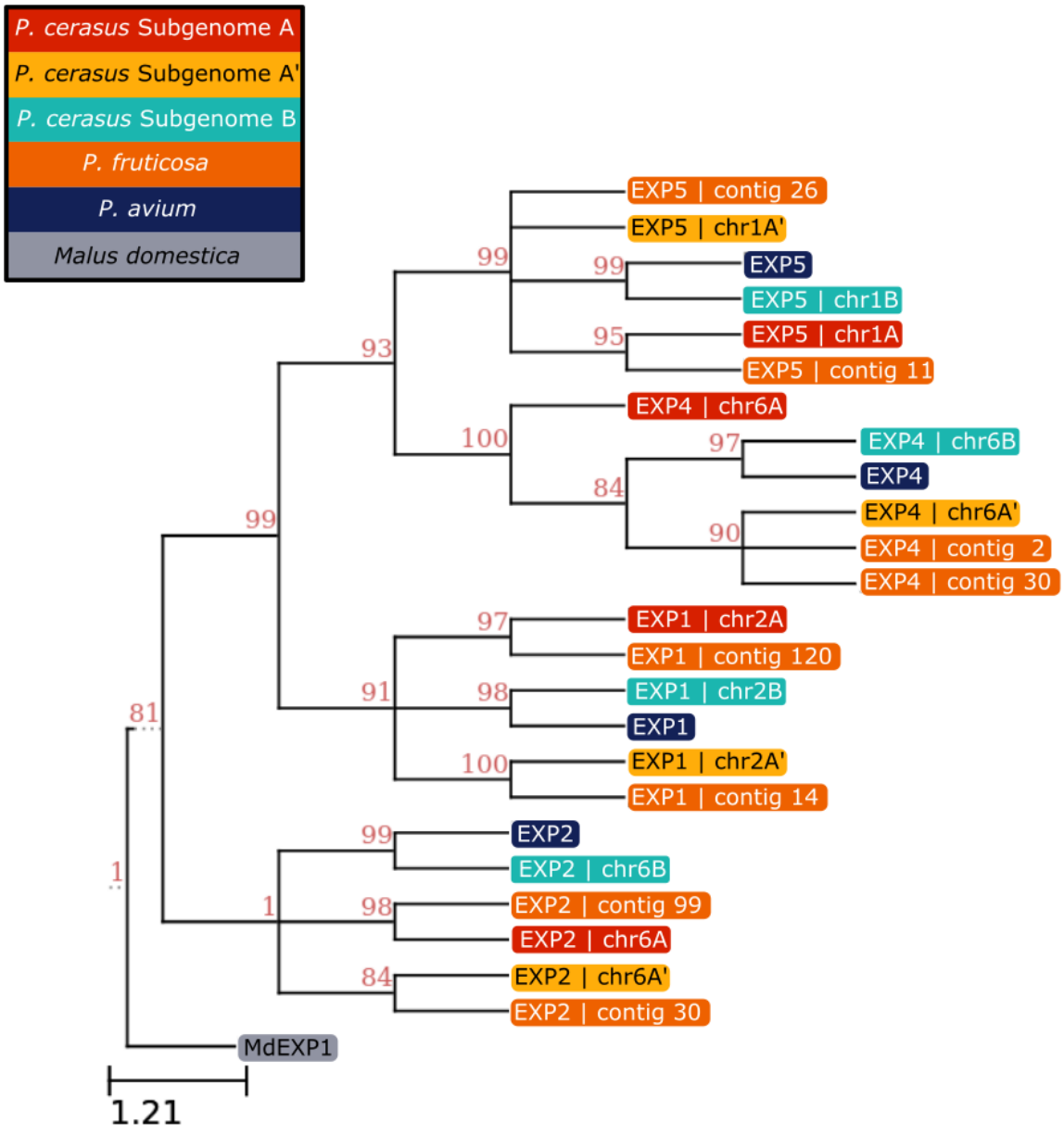
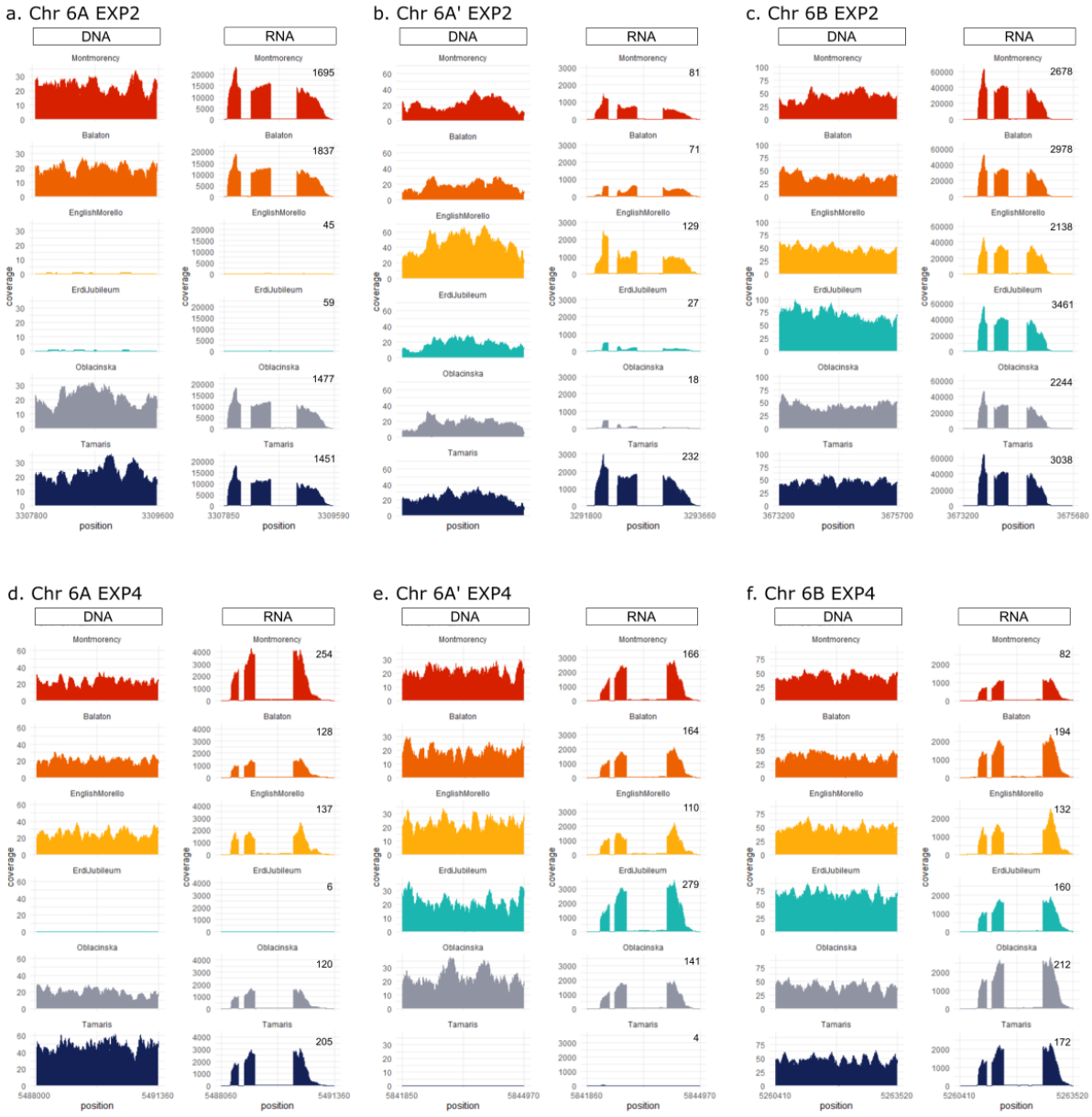


**Figure 3.6:** Alignment and coverage of DNA short reads and RNA-seq short reads for each sour cherry accession to the EXP2 and EXP4 gene regions (a-c and d-f, respectively) from the 'Montmorency' reference genome. RNA sequence data is from fruit in development stage 3.



**Discussion**

Here we present six accessions of a perennial allopolyploid species, three of which are landraces with no evidence of homoeologous exchange. The Montmorency reference genome we used showed no "obvious" evidence of HE between the ancestral subgenomes, based on the combined results from PacBio long read sequencing and Hi-C sequencing used in assembly, and kmer analysis of the assembled subgenomes. Because of this the authors believe Montmorency may itself be the direct result of an interspecific hybridization, which makes it an excellent reference against which to compare other material (Goeckeritz et al. 2023). The lack of HE shown here in Balaton and Oblacinska then raises the question of whether these two accessions are themselves first-generation interspecific hybrids, formed in the wild and preserved clonally by humans. It is possible that, with only short-read sequencing of our material, we are failing to detect HE events in Balaton and Oblacinska that might have been picked up by high quality long read sequencing. If the HE events in these accessions were balanced between the subgenomes, then no reciprocal changes in dosage would have appeared on the coverage graphs. It is also possible, although it seems highly unlikely based on the meiotic evidence, that Balaton, Oblacinska, and Montmorency are all the result of uncharacteristically stable meiosis that resulted in consistent 1A:1A':2B homoeolog dosage. Unlike the other three landraces examined, English Morello has imbalanced exchanges between the A and A' subgenomes that were easily identified with short-read DNA alignment to the 'Montmorency' reference. English Morello is a clonal selection from the same landrace as Schattenmorelle, for which a reference genome preprint was released earlier this year (Wöhner et al. 2023). The creators of the Schattenmorelle genome collapsed both *P. fruticosa*-derived subgenomes into one, so they would not have detected HE events between A and A' as we have. Wöhner et al. used a series of sequence

similarity comparisons to a wild *P. fruticosa* ecotype (Wöhner et al. 2021) and the domesticated *P. avium* 'Tieton' reference genome (Wang et al. 2020) to assign 28 regions of Schattenmorelle as HE events between the *P. fruticosa*- and *P. avium*-derived subgenomes. Given that the exact progenitors of sour cherry are unknown but were almost certainly undomesticated, further examination of English Morello/Schattenmorelle with comparisons to more sour cherries and more wild sweet and ground cherry germplasm would provide further clarity on its genome content.

Two bred sour cherry cultivars, resulting from known crosses where at least one parent is a landrace, exhibit HE events and whole-chromosome replacements. Previous research shows that neopolyploids tend to exhibit the highest rates of irregular chromosome pairing in the first generations following polyploidization, and that HEs are part of an effort to regain stable meiosis through diploidization (Li et al. 2021). Tamaris has 11 exchanges between the A and A' subgenomes and 1 exchange between A and B. Curiously, two of these A/A' exchanges, the top of chromosome 2 and the bottom of chromosome 5, appear to be in common with that in English Morello, suggesting the two accessions may have shared ancestry. The 3x dosage of chromosomes 3B, 6B, 7B, and 8B in Erdi Jubileum could be the result of irregular chromosome pairing at meiosis, or the result of a *P. avium* introgression further back in its pedigree. Erdi Jubileum is progeny of a Pandy *x* Eugenia cross, and Balaton (Újfehértói Fürtös) is a sport of Pandy; therefore there is some ability to liken these accessions as a pseudo-parent-offspring set. Thus, the question of how Erdi Jubileum's irregularities were inherited remains to be answered since Balaton shows no evidence of HE and the Eugenia genotype has since been lost, so. Both Tamaris and Erdi Jubileum have lower fruit set than Montmorency, which may be partially

explained by the challenge of achieving regular meiotic pairing with their current homoeolog configurations.

The overall frequency of HE events identified in English Morello, Erdi Jubileum, and Tamaris is similar to what has been observed in S6-S12 generation synthetic allotetraploid wheat (Zhang et al. 2020). Yet, despite the HEs and whole-chromosome replacements, the observed consistency of overall subgenome ratios (1A:1A':2B) (Figure S3) is a phenomenon that has also been observed in *Brassica napus,* and was hypothesized to be enforced by dosage-sensitive gene networks that have deleterious effects when unbalanced (Xiong et al. 2011). Xiong et al. (2011) also identified an inverse correlation between the number of HE in an individual and that individual's ability to set seed, providing further support to our explanation of Tamaris and Erdi Jubileum's lower fruit set.

The high number of HE events between the A and A' subgenomes supports the hypothesis these two subgenomes preferentially pair at meiosis. Previous results indicated that both and A and A' were inherited from the *P. fruticosa*-like progenitor (Goeckeritz et al. 2023), and this result underlines the role of sequence similarity in chromosome pairing at meiosis. Based on their common origin, the A and A' homoeologs are expected to have greater sequence similarity to each other than to the *P. avium*-derived B subgenome. The higher number of exchanges between A and A' indicates preferential pairing within sour cherry, and the exchanges may be first steps towards homogenization of the A and A' subgenomes in sour cherry. The exchanges occurring between A and B or A' and B homoeologs and whole-chromosome replacements correspond with the myriad observations of aberrant chromosome pairing at meiosis in sour cherry (Hruby, K. 1950; Wang, Dechun 1998; Iezzoni et al. 2005; Akšić et al. 2016). As sour

cherry can readily cross with both of its extant progenitors, we also cannot rule out one or more introgressions that could have further interrupted diploidization in different sour cherry lineages.

Subgenome expression bias is consistently in favor of the A and A' subgenomes in sour cherry, and the magnitude of this bias increases during fruit development. While we are cautious to make any generalizations with sour cherry, the consistency of subgenome expression bias across accessions does replicate results in other clades that show expression bias establishes immediately after polyploidization and tends to persist through generations (Edger et al. 2017; Bird et al. 2021; Vasudevan et al. 2023). We initially looked at 1:1 subgenome expression comparisons and observed consistent bias favoring A and A' over the B subgenome at magnitudes similar to other subgenome expression bias publications (overall median A/B log2FC is -0.139; overall median A'/B log2FC is -0.122) (Bird et al. 2021; Benson et al. 2023). But similar to recent work in synthetic hexaploid wheat (Vasudevan et al. 2023), we found that treating the two subgenomes inherited from the tetraploid ancestor (A and A') as one unit and comparing them with the subgenome inherited from the diploid ancestor (B) suggested almost twice as much subgenome bias (overall median AA'/B log2FC is -0.234). Similarly in hexaploid wheat (Vasudevan et al. 2023) the two subgenomes that dominate are those that were inherited from the allotetraploid progenitor. The authors hypothesize that the strong AB dominance over the D subgenome in newly-synthesized polyploid lines may be an effort to balance gene expression, as the parental DD line exhibits much higher gene expression than the parental AABB line, and established hexaploid wheat lines do not exhibit the same repression of D genome expression. Because the ancestry of our landrace accessions is unknown, testing whether this is true in sour cherry would require resynthesizing the interspecific hybrid and collecting

expression data for it as well as both parental lines. However, recreating sour cherry is challenging as it requires an unreduced gamete from the diploid sweet cherry parent.

Determining subgenome dosage allowed for evaluation of subgenome expression bias. The S-alleles, which have known origins in sour cherry's progenitors, allowed us to partially validate our subgenome assignments. The expansins, provide a demonstration of why knowledge of subgenome dosage is critical to interpreting gene expression results. The A-subgenome homoeolog of EXP2 is missing in both English Morello and Erdi Jubileum, and in fruit stage 3 of both these accessions expression of the remaining homoeologs is strongly biased in favor of the B subgenome. Without the context of subgenome dosage we might have falsely categorized a greater number of homoeolog pairs as biased towards the B subgenome, but since we know the dosage information we can have greater confidence in our results. In the remaining accessions with 1:1:2 dosage of EXP2, Montmorency, Oblacinska, and Balaton all exhibit bias against the A' allele in the A/A' and A/B homoeolog comparisons in fruit stage 3. Tamaris shows differential expression between the A and A' and A' and B EXP2 homoeologs as well, but it does not meet our log2FC threshold of 3.5 (A/A' log2FC: -2.762; A'/B log2FC: 2.801). Other B-biased homoeologs in developing fruit include pectin methylesterase and pectinesterase inhibitors, which may also play a role in fruit softening processes (Table S4).

Going forward, additional analyses of sour cherry may provide more context for the establishment of subgenome dynamics. Previous work characterizing sour cherry S-haplotypes has already shown that TE activity occurred post-polyploidy event and is directly affecting GSI in sour cherry (Tsukamoto et al. 2006)**.** A burst of TE activity is thought to be common to neopolyploids (Alger and Edger 2020)**.** DNA methylation patterns, which are often associated with TE content, also play a role in the determination of subgenome dominance in some

polyploids (Edger et al. 2017; Li et al. 2019; Bird et al. 2021). As mentioned above, resynthesizing sour cherry by crossing a known ground cherry and sweet cherry would aid in evaluating these possible mechanisms of subgenome dominance and expression bias in this species. Information about parental TE content, DNA methylation, and gene expression levels would provide a baseline that could be used to compare resynthesized sour cherry and subsequent generations, and provide clues to how the subgenomes interact and gene expression changes translate into structural changes. Examining pedigrees derived from landraces would also provide insight into how subgenome dynamics are passed down through generations, and what they ultimately "settle" into several generations out from the polyploidy event.

In conclusion, we present sour cherry as a neopolyploid with highly variable HE activity and consistent subgenome expression bias. The HE activity we observed in English Morello, Tamaris, and Erdi Jubileum corresponds with what we expect from a neopolyploid, while the lack of evidence for HE in Montmorency, Balaton, and Oblacinska certainly warrants further investigation into the origins of these three landraces. The subgenome expression bias common across all accessions and tissues shown here supports the growing hypothesis that if subgenome expression bias exists, it tends to fall in a consistent direction across individuals within a species (Edger et al. 2017; Bird et al. 2021; Vasudevan et al. 2023). These results have implications both for the evolutionary genomics of polyploidy and for sour cherry breeders seeking to better-characterize their germplasm.

**Supplementary tables and figures**

**Figure S1:** Tissues collected for RNA sequencing illustrated for 'Montmorency': (a) leaf, (b) whole flower, ( c) fruit stage 1 (whole fruit), (d) fruit stage 2 (exocarp and mesocarp), and (e) fruit stage 3 (exocarp and mesocarp). (f) A sigmoidal fruit growth curve was plotted for each of the six accessions to enable the sequencing of fruit at equivalent stages.

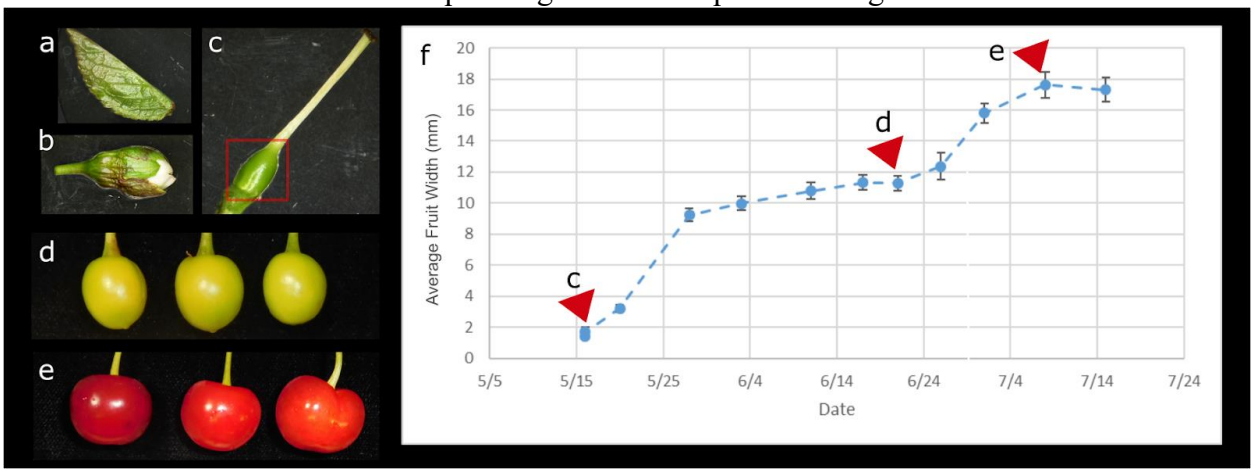



**Figure S2:** DNA sequence coverage graphs for Balaton, English Morello, Erdi Jubileum, and Oblacinska generated by aligning 2x150bp short reads to the Montmorency reference and averaging coverage over 1Mb windows. The 1x, 2x, and 3x dosage lines on the y axis correspond to 20x, 40x and 60x coverage, respectively. The predicted boundaries of the centromeric regions are denoted with black lines and were taken from Goeckeritz et al. (2023). Red, yellow and blue lines denote A, A' and B homoeolog coverage, respectively.

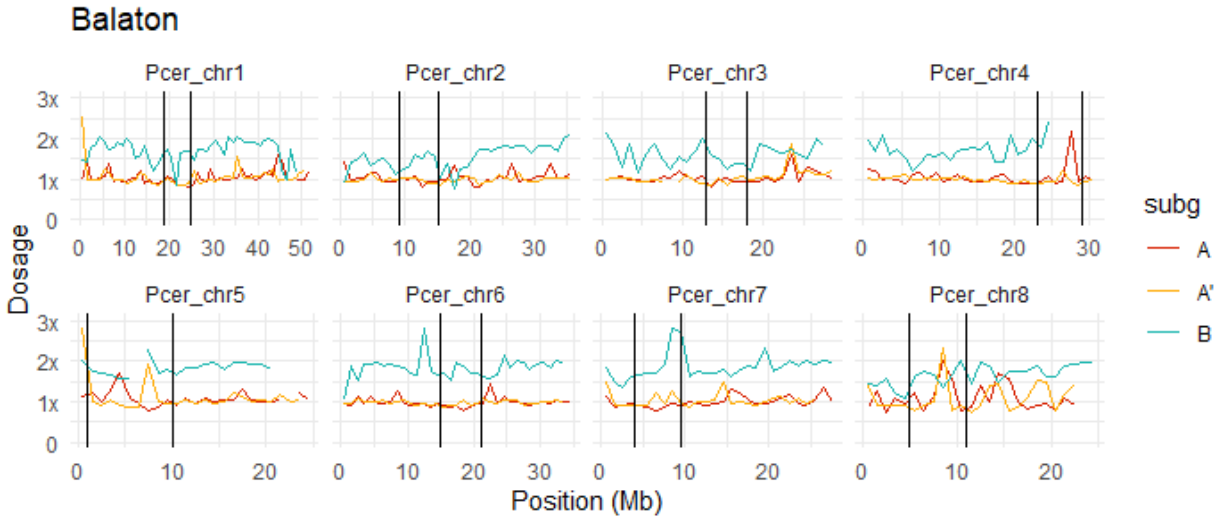


**Figure S2 (cont'd)**

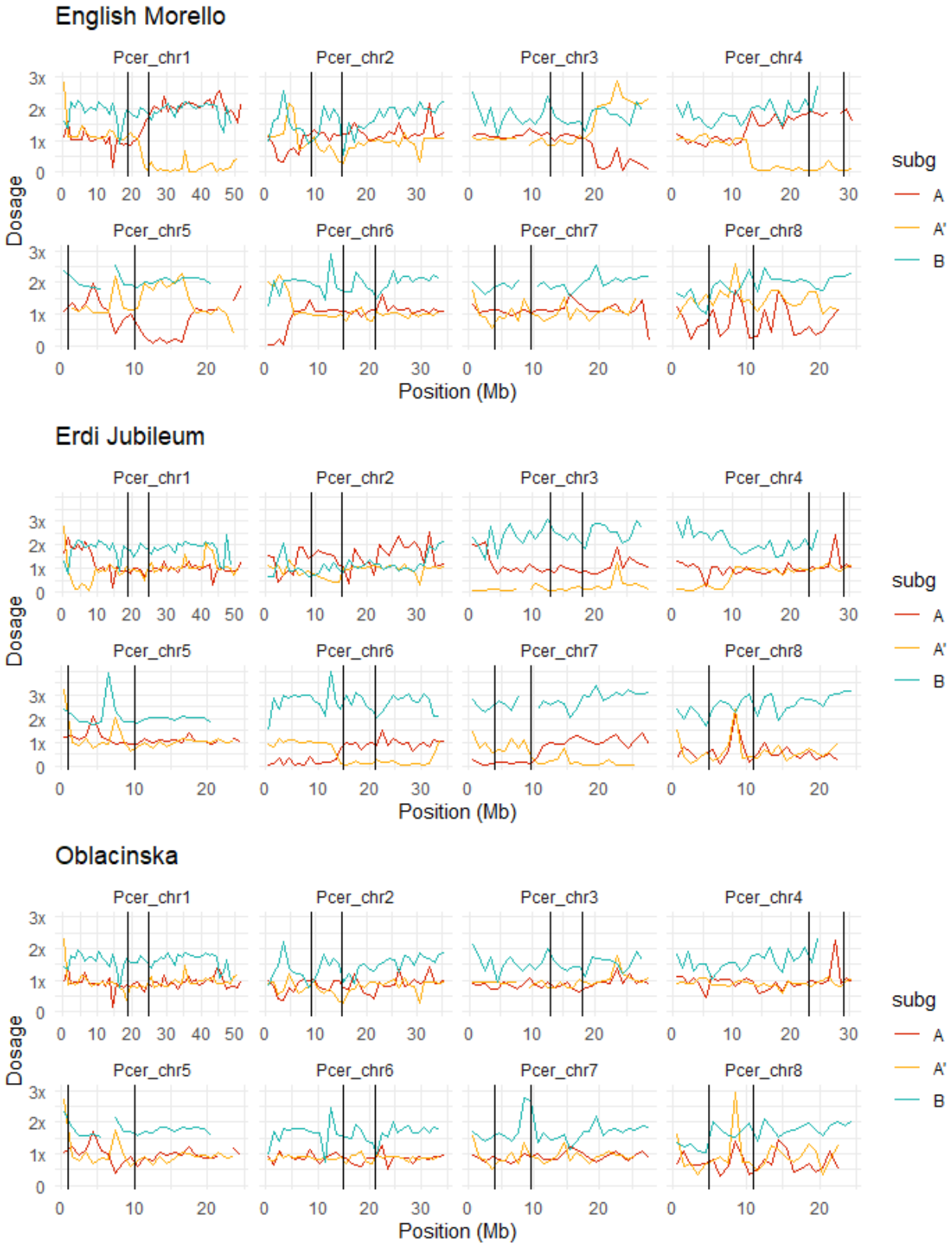


**Table S1:** Homoeologous exchange events identified for the 6 accessions. A) Total number of homoeologous exchange events for each of the three possible subgenome exchanges. B) Physical locations of all homoeologous exchange events organized by chromosome location (kbp). Exchange location is given in kilobases for each homoeolog in the exchange.

**A**

| **total number of exchanges** | | |
|---|---|---|
| **A / A'** | **A / B** | **A' / B** |
| 20 | 4 | 2 |

**B**

| **chromosome** | **accession** | **exchange partners** | **exchange location (kb)** |
|---|---|---|---|
| 1 | Erdi Jublieum | A / A’ | 8752.5 / 8619.5 |
| 1 | English Morello | A / A' | 22,665.5 / 22,650.5 |
| 1 | Tamaris | A / B | 7223.5 / 7018.5 |
| 1 | Tamaris | A / A' | 33,212.5 / 32,203.5 |
| 2 | Erdi Jublieum | A / B | 5868.5 / 6182.5 |
| 2 | Erdi Jublieum | A / B | 30,784.5 / 31,822.5 |
| 2 | English Morello | A / A' | 6576.5 / 6015.5 |
| 2 | Tamaris | A / A' | 5926.5 / 6015.5 |
| 2 | Tamaris | A / A' | 21,824.5 / 21,857.5 |
| 3 | Erdi Jublieum | A / B | 3216.5 / 3018.5 |
| 3 | English Morello | A / A' | 19,803.5 / 19,710.5 |
| 3 | Tamaris | A / A' | 4153.5 / 4128.5 |
| 3 | Tamaris | A / A' | 28,142.5 / 28,007.5 |
| 4 | Erdi Jublieum | A’ / B | 9585.5 / 8904.5 |
| 4 | English Morello | A / A' | 12,793.5 / 12,930.5 |
| 4 | Tamaris | A / A' | 8941.5 / 9397.5 |
| 5 | English Morello | A / A' | 11,339.5 / 11,049.5 |
| 5 | English Morello | A / A' | 17,590.5 / 17,310.5 |
| 5 | English Morello | A / A' | 24,128.5 / 23,341.5 |
| 5 | Tamaris | A / A' | 17,592.5 / 17,311.5 |
| 6 | Erdi Jublieum | A / A’ | 14,213.5 / 13,672.5 |
| 6 | Erdi Jublieum | A’ / B | 32,680.5 / 31,921.5 |
| 6 | English Morello | A / A' | 4504.5 / 4542.5 |
| 6 | Tamaris | A / A' | 4504.5 / 4543.5 |
| 6 | Tamaris | A / A' | 20,760.5 / 20,765.5 |
| 7 | Erdi Jublieum | A / A’ | 10,736.5 / 9404.5 |
| 8 | Tamaris | A / A' | 16,179.5 / 17,971.5 |

**Figure S3:** Histogram showing total percentages of predicted subgenome content for each of the six sour cherry accessions relative to the 'Montmorency' reference.

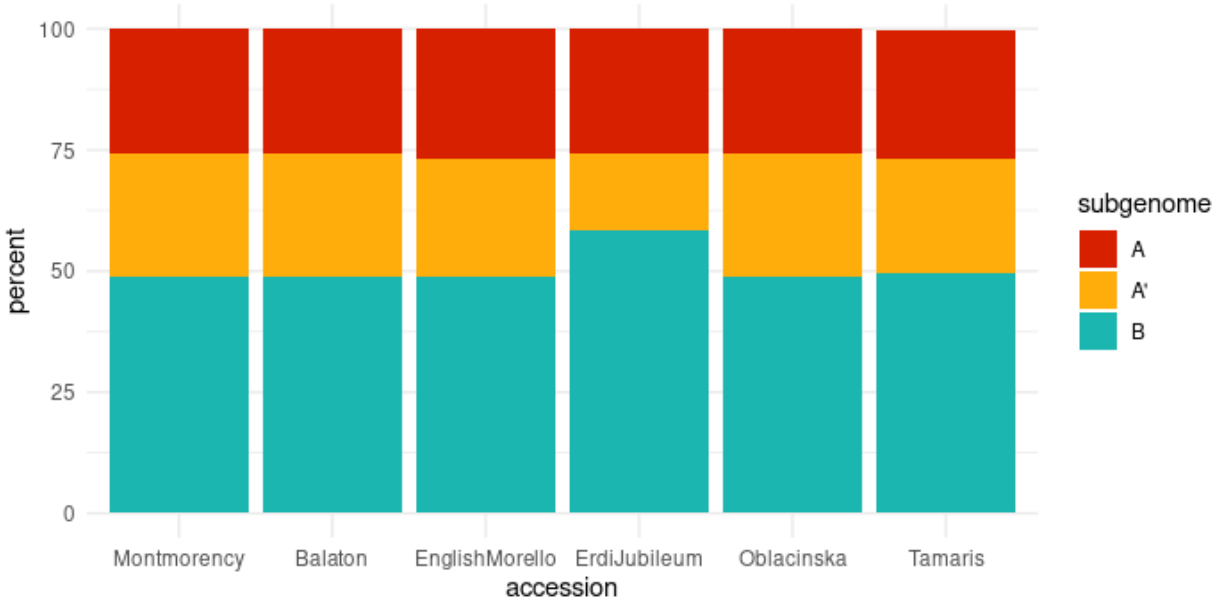


**Figure S4:** Dot plots of log2FC values for A/A', A/B and A'/B homoeolog pair comparisons by tissue for the six sour cherry accessions. Homoeolog pairs with a log2FC less than -3.5 or greater than 3.5 are colored according to the subgenome towards which they are biased. Median log2FC values for biased homoeolog groups and all homoeolog pairs are indicated in black. Expression data is normalized to homoeolog dose and significance groups based on Tukey's HSD of log2FC values are indicated by the letter above each dotplot.

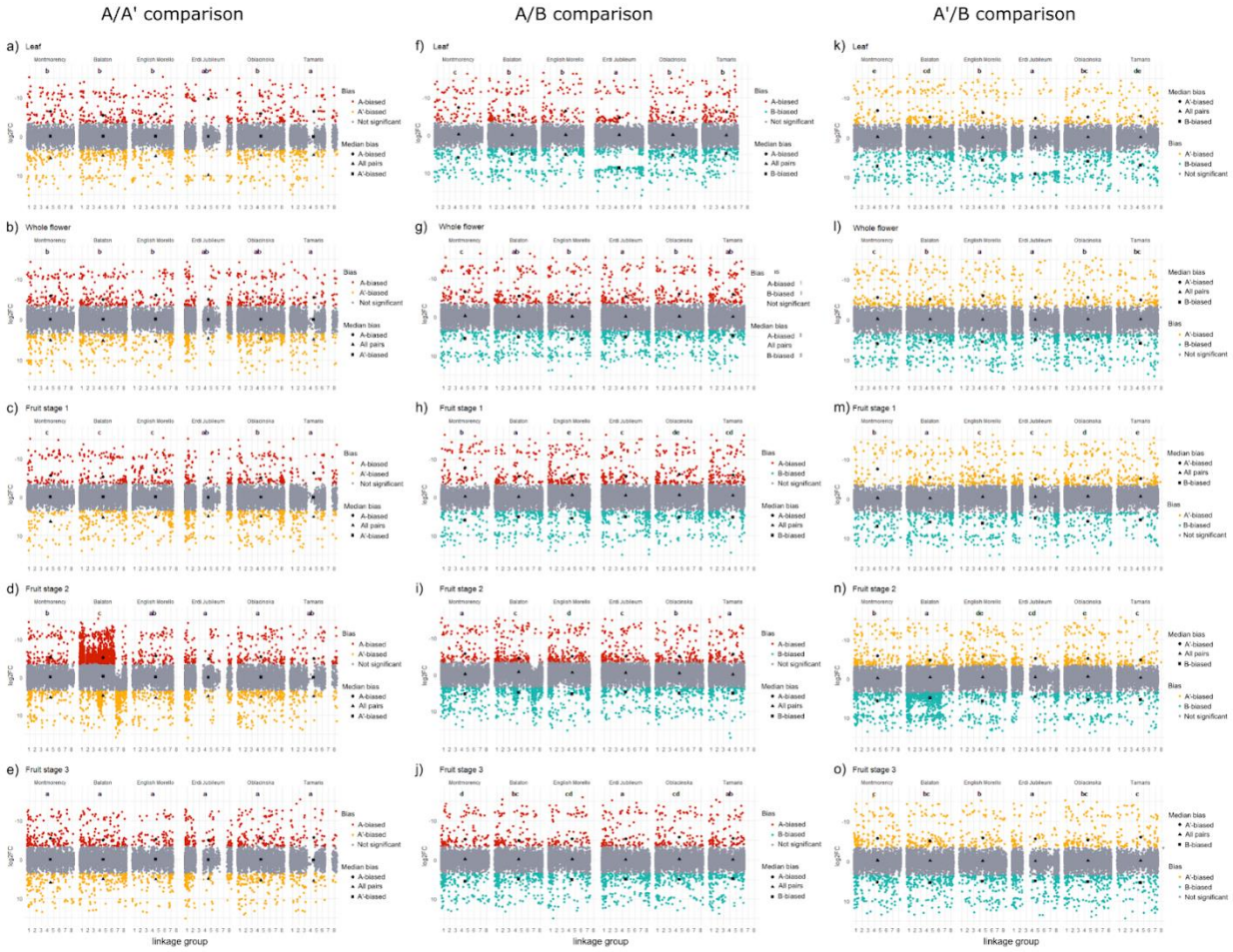


**Figure S5:** Dot plots of log2FC values for A/A', A/B and A'/B homoeolog pair comparisons by accession for the five tissues sampled. Homoeolog pairs with a log2FC less than -3.5 or greater than 3.5 are colored according to the subgenome towards which they are biased. Median log2FC values for biased homoeolog groups and all homoeolog pairs are indicated in black. Expression data is normalized to homoeolog dose and significance groups based on Tukey's HSD of log2FC values are indicated by the letter above each dot plot with "a" representing the least amount of expression bias and subsequent letters denoting higher levels of significance in expression bias.

**Table S2:** Homoeolog pair counts, with total pairs labeled significantly biased in our analyses and median log2FC values for each subgenome comparison, divided into individual tissue types and accessions. Link to table for the sake of space.

**Table S3:** Shared AA'-biased homoeolog pairs with annotation information from the Montmorency reference genome annotation (Goeckeritz et al. 2023). Link to table for sake of space.

**Table S4:** All B-biased homoeolog pairs with annotation information from the Montmorency reference genome annotation (Goeckeritz et al. 2023). Link to table for sake of space.

**Figure S6:** Amino acid (a) and coding sequence (b) alignments for all three Montmorency expansin 4 (EXP4) homoeologs, along with the sweet cherry EXP4 ortholog and two *P. fruticosa* orthologs. Non-reference amino acids or bases are highlighted in color.
(a)

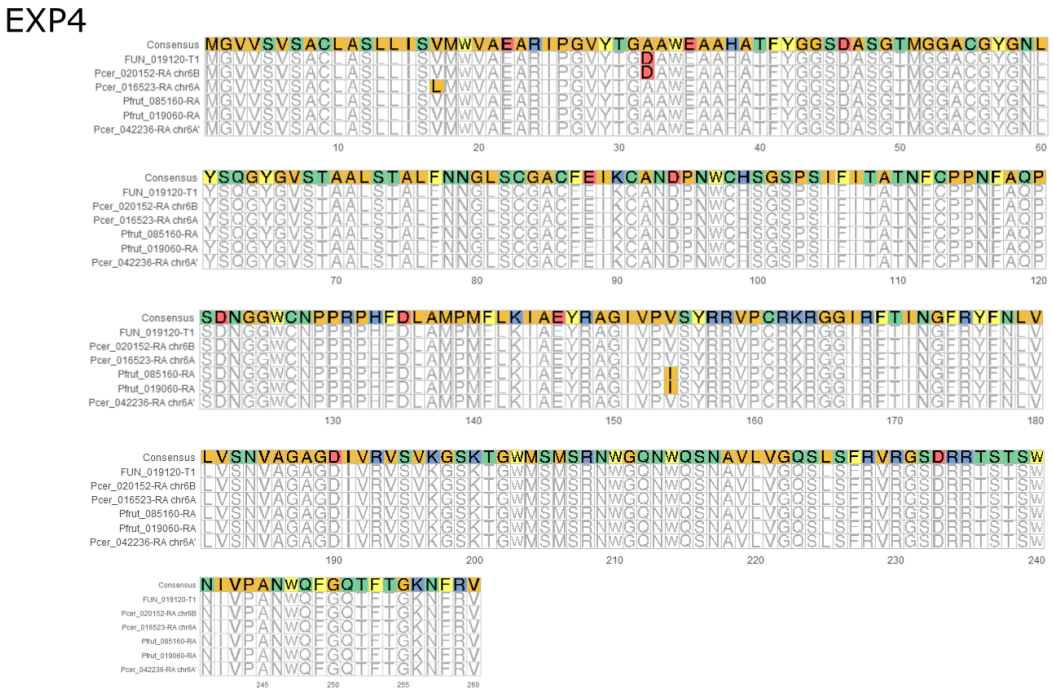


**Figure S6 (cont'd)**

(b)

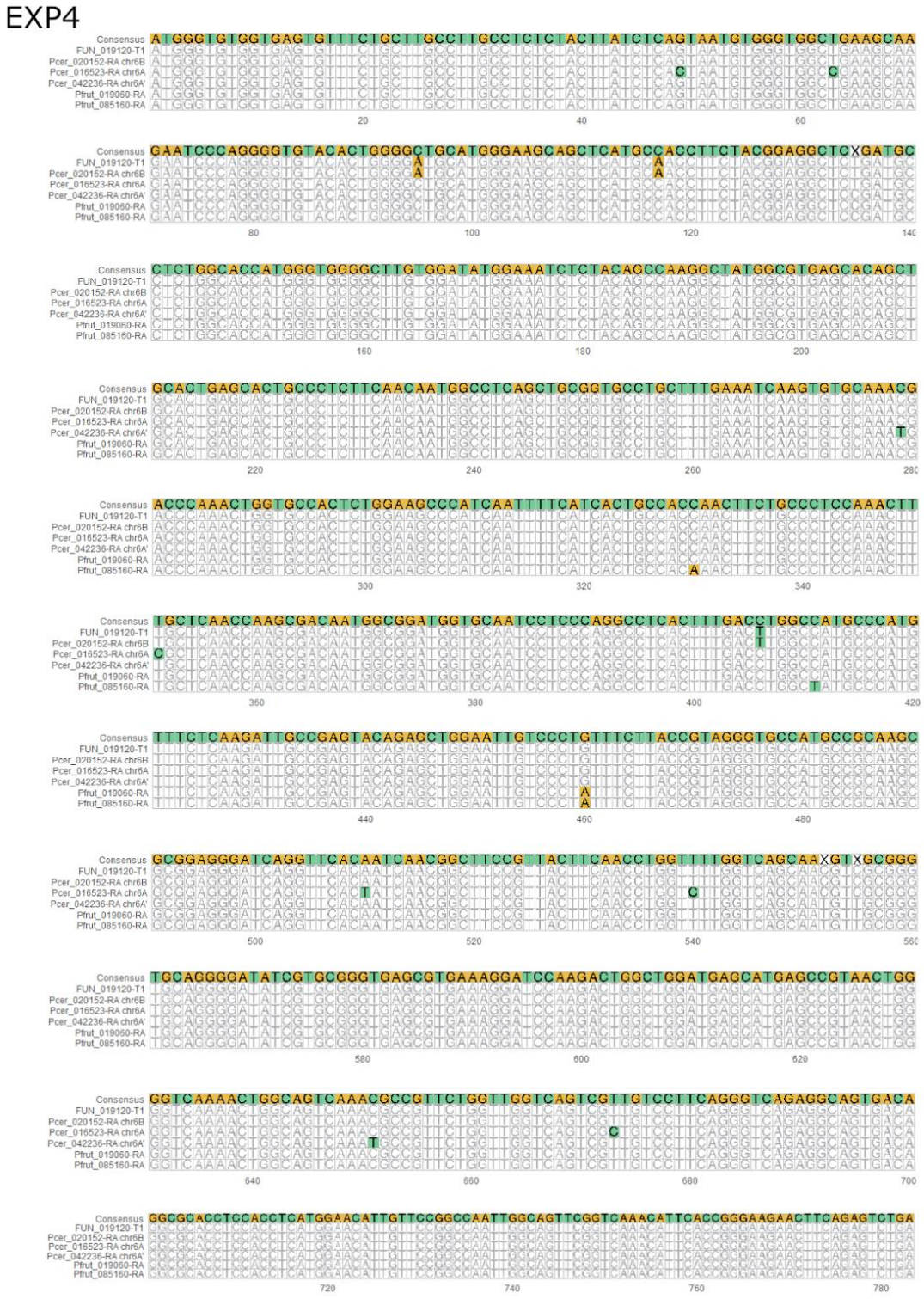


**Figure S7:** Heatmap showing relative gene expression in 'Montmorency' for each expansin homoeolog at three points during fruit development (stage 1, 2 and 3, See Fig. S1). Results from each of the three biological replicates are shown. Heatmap colors reflect transcript counts that were transformed with a variance stabilizing transformation.

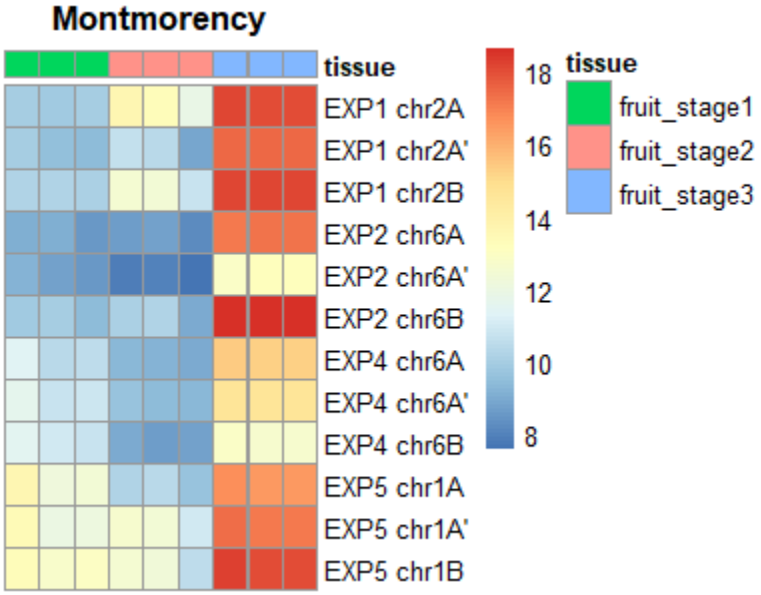


**Figure S8:** Alignment and coverage of DNA short reads and RNA-seq short reads for each sour cherry accession to the EXP1 and EXP5 gene regions from the 'Montmorency' reference genome (a-c and d-f, respectively). RNA sequence data is from fruit in development stage 3.

**Table S6:** Expansion homoeologs with gene IDs and TPM counts for fruit stage 3 for the six sour cherry accessions. Results are graphically presented in Figures 6 and S10 for EXP2/EXP4 and EXP1/EXP5, respectively.

| gene | chromosome | geneID | Montmorency | Balaton | English Morello | Erdi Jubileum | Oblacinska | Tamaris |
|---|---|---|---|---|---|---|---|---|
| EXP1 | chr2A' | Pcer_070675 | 1305 | 1081 | 759 | 836 | 1323 | 1268 |
| | chr2A | Pcer_075440 | 1959 | 1509 | 1013 | 1275 | 2079 | 1625 |
| | chr2B | Pcer_052272 | 3344 | 2614 | 2023 | 2428 | 3655 | 2759 |
| EXP2 | chr6A | Pcer_016199 | 1695 | 1837 | 45 | 59 | 1477 | 1451 |
| | chr6A' | Pcer_041884 | 81 | 71 | 129 | 27 | 18 | 232 |
| | chr6B | Pcer_019901 | 2678 | 2978 | 2138 | 3461 | 2244 | 3038 |
| EXP4 | chr6A | Pcer_016523 | 254 | 128 | 137 | 6 | 120 | 205 |
| | chr6A' | Pcer_042236 | 166 | 164 | 110 | 279 | 141 | 4 |
| | chr6B | Pcer_020152 | 82 | 194 | 132 | 160 | 212 | 172 |
| EXP5 | chr1A | Pcer_002587 | 762 | 839 | 775 | 633 | 919 | 1152 |
| | chr1A' | Pcer_007821 | 1286 | 1087 | 137 | 1205 | 1648 | 1043 |
| | chr1B | Pcer_013136 | 2882 | 2435 | 1516 | 2700 | 3183 | 3157 |

## REFERENCES


Akšić MF, Cerović R, Ercişli S, Jensen M. 2016. Microsporogenesis and meiotic abnormalities in different 'Oblačinska' sour cherry (Prunus cerasus L.) clones. Flora. 219:25–34. doi:10.1016/J.FLORA.2015.12.009.

Alger EI, Edger PP. 2020. One subgenome to rule them all: underlying mechanisms of subgenome dominance. Curr Opin Plant Biol. 54:108–113. doi:10.1016/j.pbi.2020.03.004.

Barker MS, Arrigo N, Baniaga AE, Li Z, Levin DA. 2016. On the relative abundance of autopolyploids and allopolyploids. New Phytol. 210(2):391–398.

Beaver JA, Iezzoni AF. 1993. Allyzyme Inheritance in Tetraploid Sour Cherry (Prunus cerasus L.). J Am Soc Hortic Sci. 118(6):873–877.

Benson CW, Sheltra MR, Maughan PJ, Jellen EN, Robbins MD, Bushman BS, Patterson EL, Hall ND, Huff DR. 2023. Homoeologous evolution of the allotetraploid genome of Poa annua L. BMC Genomics. 24(1):350. doi:10.1186/s12864-023-09456-5.

Bird KA, Jacobs M, Sebolt A, Rhoades K, Alger EI, Colle M, Alekman ML, Bies PK, Cario AJ, Chigurupati RS, et al. 2022. Parental origins of the cultivated tetraploid sour cherry (Prunus cerasus L.). PLANTS PEOPLE PLANET. 4(5):444–450. doi:10.1002/ppp3.10267.

Bird KA, Niederhuth CE, Ou S, Gehan M, Pires JC, Xiong Z, VanBuren R, Edger PP. 2021. Replaying the evolutionary tape to investigate subgenome dominance in allopolyploid Brassica napus. New Phytol. 230(1):354–371. doi:10.1111/nph.17137.

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

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. BMC Bioinformatics. 10:421. doi:10.1186/1471-2105-10-421.

Cheng F, Wu J, Cai X, Liang J, Freeling M, Wang X. 2018. Gene retention, fractionation and subgenome differences in polyploid plants. Nat Plants. 4(5):258–268. doi:10.1038/s41477-018-0136-7.

Colle M, Leisner CP, Wai CM, Ou S, Bird KA, Wang J, Wisecaver JH, Yocca AE, Alger EI, Tang H, et al. 2019. Haplotype-phased genome and evolution of phytonutrient pathways of tetraploid blueberry. GigaScience. 8(3). doi:10.1093/gigascience/giz012. [accessed 2020 Oct 23]. https://academic.oup.com/gigascience/article/doi/10.1093/gigascience/giz012/5304886.

Conway JR, Lex A, Gehlenborg N. 2017. UpSetR: an R package for the visualization of intersecting sets and their properties. Bioinformatics. 33(18):2938–2940. doi:10.1093/bioinformatics/btx364.

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, et al. 2021. Twelve years of SAMtools and BCFtools. GigaScience. 10(2):giab008. doi:10.1093/gigascience/giab008.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. Bioinforma Oxf Engl. 29(1):15–21. doi:10.1093/bioinformatics/bts635.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32(5):1792–1797. doi:10.1093/nar/gkh340.

Edger PP, Poorten TJ, VanBuren R, Hardigan MA, Colle M, McKain MR, Smith RD, Teresi SJ, Nelson ADL, Wai CM, et al. 2019. Origin and evolution of the octoploid strawberry genome. Nat Genet. 51(3):541–547. doi:10.1038/s41588-019-0356-4.

Edger PP, Smith R, McKain MR, Cooley AM, Vallejo-Marin M, Yuan Y, Bewick AJ, Ji L, Platts AE, Bowman MJ, et al. 2017. Subgenome Dominance in an Interspecific Hybrid, Synthetic Allopolyploid, and a 140-Year-Old Naturally Established Neo-Allopolyploid Monkeyflower. Plant Cell. 29(9):2150–2167. doi:10.1105/tpc.17.00010.

Faust M, Timon B, Surányi D, Nyujtó F, Gradziel TM. 2011. Origin and dissemination of Prunus crops: peach, cherry, apricot, plum and almond. Orig Dissem Prunus Crops Peach Cherry Apric Plum Almond. https://www.cabdirect.org/cabdirect/abstract/20123346667.

Gasic K, Hernandez A, Korban SS. 2004. RNA extraction from different apple tissues rich in polyphenols and polysaccharides for cDNA library construction. Plant Mol Biol Report. 22(4). doi:10.1007/BF02772687.

Goeckeritz CZ, Rhoades KE, Childs KL, Iezzoni AF, VanBuren R, Hollender CA. 2023. Genome of tetraploid sour cherry (Prunus cerasus L.) ‘Montmorency’ identifies three distinct ancestral Prunus genomes. Hortic Res. 10(7):uhad097. doi:10.1093/hr/uhad097.

Hillig KW, Iezzoni AF. 1988. Multivariate Analysis of a Sour Cherry Germplasm Collection. J Am Soc Hortic Sci. 113(6):928–934. doi:10.21273/JASHS.113.6.928.

Hruby, K. 1950. The cytology of tetraploid cherries. Stud Bot Cechoslov-. 11(3):87–97.

Hufford MB, Seetharam AS, Woodhouse MR, Chougule KM, Ou S, Liu J, Ricci WA, Guo T, Olson A, Qiu Y, et al. 2021. De novo assembly, annotation, and comparative analysis of 26 diverse maize genomes. Science. 373(6555):655–662. doi:10.1126/science.abg5289.

Iezzoni AF, Hancock AM. 1996. Chloroplast DNA variation in sour cherry. Acta Hortic.(410):115–120. doi:10.17660/ActaHortic.1996.410.17.

Iezzoni AF, Sebolt AM, Wang D. 2005. Sour cherry breeding program at Michigan state university. Acta Hortic. 667:131–134. doi:10.17660/ActaHortic.2005.667.18.

Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. 2019. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Wren J, editor. Bioinformatics. 35(21):4453–4455. doi:10.1093/bioinformatics/btz305.

Li H. 2013 Mar. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. http://arxiv.org/abs/1303.3997.

Li N, Xu C, Zhang A, Lv R, Meng X, Lin X, Gong L, Wendel JF, Liu B. 2019. DNA methylation repatterning accompanying hybridization, whole genome doubling and homoeolog exchange in nascent segmental rice allotetraploids. New Phytol. 223(2):979–992. doi:10.1111/nph.15820.

Li Z, McKibben MTW, Finch GS, Blischak PD, Sutherland BL, Barker MS. 2021. Patterns and Processes of Diploidization in Land Plants. Annu Rev Plant Biol. 72(1):387–410. doi:10.1146/annurev-arplant-050718-100344.

Mason AS, Wendel JF. 2020. Homoeologous Exchanges, Segmental Allopolyploidy, and Polyploid Genome Evolution. Front Genet. 11. [accessed 2023 Jan 24]. https://www.frontiersin.org/articles/10.3389/fgene.2020.01014.

McRae L, Beric A, Conant GC. 2022. Hybridization order is not the driving factor behind biases in duplicate gene losses among the hexaploid Solanaceae. Proc R Soc B Biol Sci. 289(1985):20221810. doi:10.1098/rspb.2022.1810.

Olden EJ, Nybom N. 1968. On the Origin of Prunus Cerasus L. Hereditas. 59(2-3):327–345. doi:10.1111/J.1601-5223.1968.TB02181.X.

R Core Team. 2021. R: A language and environment for statistical computing. https://www.R-project.org/.

Schuster M, Wolfram B. 2004. Results of Sour Cherry Breeding in Dresden-Pillnitz. Acta Hortic.(663):911–914. doi:10.17660/ActaHortic.2004.663.167.

Shumate A, Wong B, Pertea G, Pertea M. 2022. Improved transcriptome assembly using a hybrid of long and short reads with StringTie. PLOS Comput Biol. 18(6):e1009730. doi:10.1371/journal.pcbi.1009730.

Sun Y, Liu Y, Shi J, Wang L, Liang C, Yang J, Chen J, Chen M. 2023. Biased mutations and gene losses underlying diploidization of the tetraploid broomcorn millet genome. Plant J. 113(4):787–801. doi:10.1111/tpj.16085.

The International Wheat Genome Sequencing Consortium (IWGSC). 2014. A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome. Science. 345(6194):1251788. doi:10.1126/science.1251788.

Tsukamoto T, Hauck NR, Tao R, Jiang N, Iezzoni AF. 2006. Molecular characterization of three non-functional S-haplotypes in sour cherry (Prunus cerasus). Plant Mol Biol. 62(3):371–383. doi:10.1007/s11103-006-9026-x.

Van der Auwera G, O'Connor B. 2020. Genomics in the Cloud: Using Docker, GATK, and WDL in Terra. O'Reilly Media.

Vasudevan A, Lévesque-Lemay M, Edwards T, Cloutier S. 2023. Global transcriptome analysis of allopolyploidization reveals large-scale repression of the D-subgenome in synthetic hexaploid wheat. Commun Biol. 6(1):1–14. doi:10.1038/s42003-023-04781-7.

Wang, Dechun. 1998. RFLP Mapping, QTL Identification, and Cytogenetic Analysis in Sour Cherry [Doctor of Philosophy]. [East Lansing, MI]: Michigan State University.

Wang J, Liu W, Zhu D, Hong P, Zhang S, Xiao S, Tan Y, Chen X, Xu L, Zong X, et al. 2020. Chromosome-scale genome assembly of sweet cherry ( Prunus avium L.) cv. Tieton obtained using long-read and Hi-C sequencing. Hortic Res. 7(1):1–11. doi:10.1038/s41438-020-00343-8.

Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. 2nd ed. 2016. Cham: Springer International Publishing : Imprint: Springer (Use R!).

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, et al. 2019. Welcome to the Tidyverse. J Open Source Softw. 4(43):1686. doi:10.21105/joss.01686.

Wöhner TW, Emeriewen OF, Wittenberg AHJ, Nijbroek K, Wang RP, Blom E-J, Schneiders H, Keilwagen J, Berner T, Hoff KJ, et al. 2023. The structure of the tetraploid sour cherry 'Schattenmorelle' (*Prunus cerasus* L.) genome reveals insights into its segmental allopolyploid nature. Front Plant Sci. 14. [accessed 2023 Dec 5]. https://www.frontiersin.org/articles/10.3389/fpls.2023.1284478.

Wöhner TW, Emeriewen OF, Wittenberg AHJ, Schneiders H, Vrijenhoek I, Halász J, Hrotkó K, Hoff KJ, Gabriel L, Lempe J, et al. 2021. The draft chromosome-level genome assembly of tetraploid ground cherry (Prunus fruticosa Pall.) from long reads. Genomics. 113(6):4173–4183. doi:10.1016/j.ygeno.2021.11.002.

Xiong Z, Gaeta RT, Pires JC. 2011. Homoeologous shuffling and chromosome compensation maintain genome balance in resynthesized allopolyploid Brassica napus. Proc Natl Acad Sci. 108(19):7908–7913. doi:10.1073/pnas.1014138108.

Zhang Z, Gou X, Xun H, Bian Y, Ma X, Li J, Li N, Gong L, Feldman M, Liu B, et al. 2020. Homoeologous exchanges occur through intragenic recombination generating novel transcripts and proteins in wheat and other polyploids. Proc Natl Acad Sci. 117(25):14561–14571. doi:10.1073/pnas.2003505117.

Zhou L, Feng T, Xu S, Gao F, Lam TT, Wang Q, Wu T, Huang H, Zhan L, Li L, et al. 2022. ggmsa: a visual exploration tool for multiple sequence alignment and associated data. Brief Bioinform. 23(4):bbac222. doi:10.1093/bib/bbac222.

Zhuang Y, Wang X, Li X, Hu J, Fan L, Landis JB, Cannon SB, Grimwood J, Schmutz J, Jackson SA, et al. 2022. Phylogenomics of the genus Glycine sheds light on polyploid evolution and life-strategy transition. Nat Plants. 8(3):233–244. doi:10.1038/s41477-022-01102-4.

# CHAPTER 4: FUTURE DIRECTIONS

In this work I present new insights into sour cherry as an allopolyploid system. I established that sour cherry exhibits variation in subgenome dosage and that there is evidence of homoeologous exchange (HE) events between its subgenomes. However, three of the six accessions I examined show no evidence of HE, for reasons that are still unclear. We are reasonably confident in a lack of HE events in 'Montmorency', as the separation of chromosomes with HiC sequencing and separation of subgenomes with kmer clustering analysis were clear except in the case of chromosomes 8A and 8A'. Resequencing 'Montmorency' with PacBio HiFi or Oxford Nanopore long reads could further clarify how similar chromosomes 8A and 8A' are and would provide greater clarity on whether 'Montmorency' has any HE events that we were simply unable to detect with the technology available to us at the time. Short-read alignment analysis may not have been sufficient to detect HE events in 'Balaton' or 'Oblacinska' if the overall homoeolog dosage remained balanced, so long-read sequencing of these individuals would aid in further evaluation of subgenome structure as well. If these individuals truly do not have any HE events amongst their subgenomes, it becomes a question of how this is the case for three landraces of sour cherry given that even one generation from the initial interspecific hybridization event would have been expected to result in cross over events, especially between the A and A' subgenomes. One explanation, albeit highly unlikely, is that 'Montmorency,' 'Balaton,' and 'Oblacinska' may all be first-generation interspecific hybrids that were clonally propagated by humans for centuries..

Preferential pairing of the A/A' homoeologs and the B homoelogs is supported by the high proportion of HE events between the A and A' homoeologs in 'English Morello', 'Erdi Jubileum', and 'Tamaris'. The presence of a lower proportion of HE events between A and B and A' and B homoeologs is in turn supported by the irregular pairing frequently observed at meiosis

in sour cherry (Hruby, K. 1950; Wang, Dechun 1998; Akšić et al. 2016). This irregular meiosis has long been hypothesized to play a role in the poor fruit set common to sour cherry, as chromosomes that fail to pair at metaphase are often left behind in telophase and lost altogether during cell division resulting in aneuploid gametes.

We have yet to find any compelling evidence of a relationship between subgenome dosage and successful fruit set. 'Balaton' and 'Oblacinska' both have the same subgenome dosage as 'Montmorency,' but neither achieves the same level of fruit set. Poor fruit set is the biggest breeding challenge in sour cherry, but it would appear based on this work that subgenome dosage alone is not the driving factor behind successful fertility and fruit set. Within the MSU breeding germplasm there are several sour cherry individuals with known pedigrees that exhibit remarkably high fruit set (all over 75 percent in summer 2020), and genomic examination of these individuals would be an excellent next step for determining the mechanisms of successful fruit set in sour cherry. A new potential starting point with that population would be to go back to a QTL for pollen germination identified on 'Erdi Botermo' chromosome 1 that was found to explain 17 percent of phenotypic variance (Wang et al. 2000).

With regards to subgenome expression bias, I found a consistent bias in favor of the A and A' subgenomes over the B subgenome in all tissues of all accessions studied, but we do not have the parental gene expression data that would provide greater context as to how the relative expression of these subgenomes differs from the baseline set in the progenitor species. We also do not have DNA methylation data or TE analysis for parents or progeny to give greater context about how subgenome expression bias is affected by these two characteristics in *Prunus*. DNA methylation and TE activity are often studied together, as genomes tend to methylate TEs to prevent their activity and polyploidy events commonly result in rapid changes in both DNA

methylation and TE activity (Alger and Edger 2020). A brief overview of the research on DNA methylation and TE activity in polyploids is given in Chapter 1 of this dissertation. DNA methylation has been found to be both negatively (Edger et al. 2017) and positively (Bird et al. 2021) correlated with subgenome expression dominance, depending on the clade. We already have evidence of a TE insertion in nonfunctional haplotypes of the *S*-locus of sour cherry (Tsukamoto et al. 2010), illustrating the potential direct application of TE research to sour cherry breeding efforts. The short-read data we have already collected could be used to begin to identify non-reference TE insertions using a software such as SPLITREADER (Quadrana et al. 2016) and determine if TE insertions are more common in one subgenome over another.

Using populations with known parents would allow for characterization of DNA methylation and TE content over generations, and resynthesizing sour cherry with known progenitors would provide an idea of how subgenome structure and subgenome expression bias are affected by a polyploidy event. Additional insight for the evolution of sour cherry subgenomes would be gained by studying full and half-sib populations of sour cherry that already exist at the MSU research station in Clarksville, alongside their known parents. Resynthesizing sour cherry from known *P. avium* and *P. fruticosa* parents would allow for the quantification of parental gene expression levels that commonly aids evaluation of subgenome expression bias in allopolyploids (Edger et al. 2017; Bird et al. 2021; Vasudevan et al. 2023), although this obviously requires a greater time investment in sour cherry than it does in annual plants. Sweet cherry has been observed to produce unreduced (diploid) pollen at a low rate (Iezzoni and Hancock 1984), so a high number of pollination attempts would be necessary to achieve a successful resynthesis. It should also be noted that while *P. avium* and *P. fruticosa* are regarded as the extant progenitors of sour cherry, it is more precise to describe the progenitors of

sour cherry as *P. avium*-like and *P. fruticosa*-like. Any resynthesis experiments would therefore only be a best approximation of the initial hybridization events that created sour cherry.

There is a substantial amount of sequence data that I collected and did not end up using for this dissertation, which has been deposited into NCBI under BioProject number PRJNA1047034 (Table 4.1). These datasets fall into three broad categories: (1) additional sour cherry RNAseq data, (2) *P. avium, P. fruticosa,* and additional sour cherry DNA and RNA sequence datasets, and (3) a three generation introgression lineage of sour cherry with the wild cherry species *P. canescens.*

In the first category, I have RNA sequence of developing shoot internode for the following sour cherry accessions: 'Montmorency,' 'Balaton,' 'Erdi Jubileum,' 'M172,' and 'Tamaris.' Figure 4.1 shows the part of the shoot internode we sampled, and the RNA extraction and sequencing was completed using the same methods outlined in Chapter 3. We also have RNAseq data for 'Montmorency' stage 3 fruit skin and 'Montmorency' stage 3 fruit flesh, originally collected in order to identify the MYB10 transcription factor that is very likely responsible for red coloring in sour cherry (Lin-Wang et al. 2010; Stegmeir et al. 2015; Jin et al. 2016; Castillejo et al. 2020).

The second and third category accessions are listed in Table 4.1. I have DNA short-read sequence and RNA sequence data (leaf, whole flower, fruit stage 1, fruit stage 2, and fruit stage 3) of 17 additional cherry accessions. All DNA sequencing was completed to a depth of about ~40x and we targeted 35 million reads for each RNAseq library. The second category material consists of a sweet cherry cultivar, landrace and wild sweet cherry accessions, hybrids between wild and landrace sweet cherries, two additional sour cherry accessions, and three wild accessions of *P. fruticosa* (one with only DNA sequence). The RNAseq data of developing fruit

alone would be useful for a researcher interested in studying differential gene expression of developing fruit in *P. avium*, *P. fruticosa*, and/or sour cherry. The *P. fruticosa* sequence data, together with the two currently-existing reference genomes for the species (Wöhner et al. 2021; Goeckeritz et al. 2023) could be used for further characterization of the subgenomes in this probable allopolyploid.

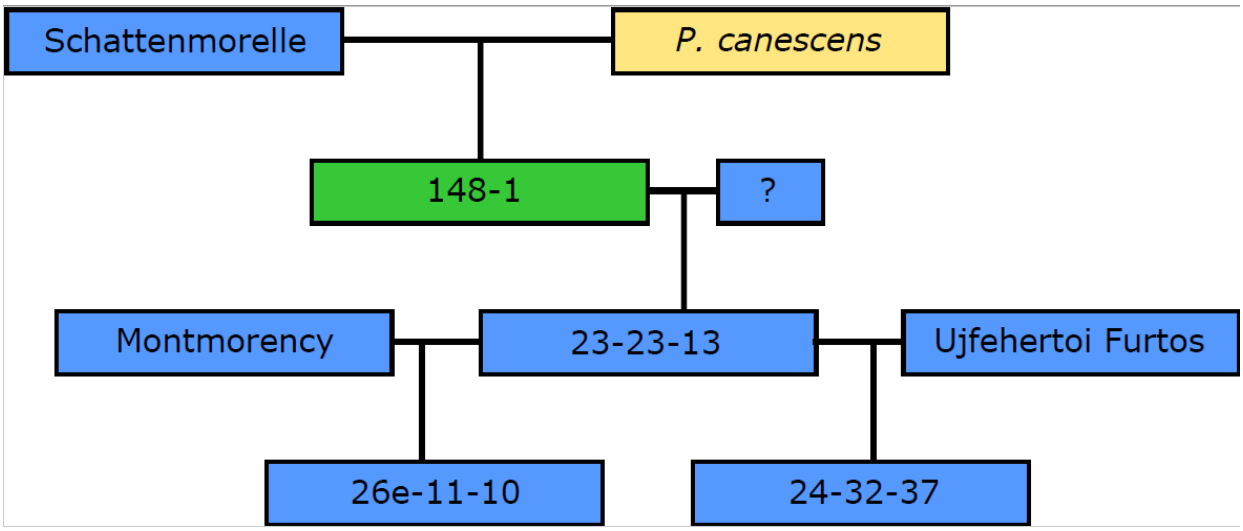
The third category dataset is short read DNA and RNAseq libraries (leaf, whole flower, fruit stage 1, fruit stage 2, fruit stage 3) of a *P. canescens* introgression lineage of sour cherry that includes a triploid individual (148-1) and two generations of its progeny (Table 4.1, Figure 4.2). The *P. canescens* introgression lineage was created with the goal of introgressing cherry leaf spot resistance from the wild diploid Asian cherry species *P. canescens* into sour cherry. The resistance was successfully introgressed, and some progeny in this population also show unexpectedly high fruit firmness which may be associated with a fruit firmness QTL on sweet cherry chromosome 4 that is near the cherry leaf spot resistance locus (Cai et al. 2019). Beyond the inheritance of the chromosome 4 *P. canescens* cherry leaf spot resistance, we do not know how much of the *P. canescens* genome persists in these progeny or how subgenome dosage and expression bias may be affected by this addition of another subgenome. Analysis of the patterns of subgenome retention and bias will provide insight into these mechanisms in a unique interspecific lineage that represents three ploidy levels (2x, 3x and 4x).

**Table 4.1:** Additional plant materials for which DNA and RNA sequencing was completed but which were not included in this dissertation.

| Accession | Species | Notes |
|---|---|---|
| Bing | *P. avium* | cultivar |
| Krupnoplodnaya | *P. avium* | landrace |
| Emperor Francis | *P. avium* | landrace |
| Schneiders | *P. avium* | landrace |
| NY54 | *P. avium* | wild |
| P20 clone | *P. avium* | NY54 *x* Cristobalina |
| MSU 23-01-14 | *P. avium* | NY54 *x* Emperor Francis |
| M172 | *P. cerasus* | Pandy *x* Eugenia, full sib to Erdi Jubileum |
| North Star | *P. cerasus* | cultivar |
| P.fruticosa12-02 | *P. fruticosa* | wild; used for *P. fruticosa* draft reference |
| P.fruticosa11-35 | *P. fruticosa* | wild |
| P.fruticosa11-59 | *P. fruticosa* | wild; DNA sequence only |
| P.canescens | *P. canescens* | wild |
| 148-1 | Schattenmorelle × *P. canescens* hybrid | triploid |
| 23-23-13 | *P. cerasus* × *P. canescens* introgression lineage | 148-1 *x* unknown |
| 24-32-37 | *P. cerasus* × *P. canescens* introgression lineage | Balaton × 23-23-13 |
| 26e-11-10 | *P. cerasus* × *P. canescens* introgression lineage | Montmorency × 23-23-13 |

**Figure 4.1:** Developing cherry shoots with red circles indicating the tissue collected for RNA sequencing for sour cherry accessions Montmorency, Balaton, Erdi Jubileum, M172, and Tamaris. Shoots were chosen that had not yet set terminal bud and the 2nd internode below the shoot apical meristem was sampled. We collected three shoot internodes per sample and sequenced three samples per genotype. RNA extraction and sequencing were performed according to the methods detailed in Chapter 3.

**Figure 4.2.** Pedigree diagram of the *P. canescens* introgression lineage. Individuals in blue are tetraploid, individuals in yellow are diploid, and the individual in green is a triploid.



## REFERENCES


Akšić MF, Cerović R, Ercişli S, Jensen M. 2016. Microsporogenesis and meiotic abnormalities in different 'Oblačinska' sour cherry (*Prunus cerasus* L.) clones. Flora. 219:25–34. doi:10.1016/J.FLORA.2015.12.009.

Alger EI, Edger PP. 2020. One subgenome to rule them all: underlying mechanisms of subgenome dominance. Curr Opin Plant Biol. 54:108–113. doi:10.1016/j.pbi.2020.03.004.

Bird KA, Niederhuth CE, Ou S, Gehan M, Pires JC, Xiong Z, VanBuren R, Edger PP. 2021. Replaying the evolutionary tape to investigate subgenome dominance in allopolyploid *Brassica napus*. New Phytol. 230(1):354–371. doi:10.1111/nph.17137.

Cai L, Quero-García J, Barreneche T, Dirlewanger E, Saski C, Iezzoni A. 2019. A fruit firmness QTL identified on linkage group 4 in sweet cherry (*Prunus avium* L.) is associated with domesticated and bred germplasm. Sci Rep. 9(1):5008. doi:10.1038/s41598-019-41484-8.

Castillejo C, Waurich V, Wagner H, Ramos R, Oiza N, Muñoz P, Triviño JC, Caruana J, Liu Z, Cobo N, et al. 2020. Allelic Variation of MYB10 Is the Major Force Controlling Natural Variation in Skin and Flesh Color in Strawberry (*Fragaria* spp.) Fruit. Plant Cell. 32(12):3723–3749. doi:10.1105/tpc.20.00474.

Edger PP, Smith R, McKain MR, Cooley AM, Vallejo-Marin M, Yuan Y, Bewick AJ, Ji L, Platts AE, Bowman MJ, et al. 2017. Subgenome Dominance in an Interspecific Hybrid, Synthetic Allopolyploid, and a 140-Year-Old Naturally Established Neo-Allopolyploid Monkeyflower. Plant Cell. 29(9):2150–2167. doi:10.1105/tpc.17.00010.

Goeckeritz CZ, Rhoades KE, Childs KL, Iezzoni AF, VanBuren R, Hollender CA. 2023. Genome of tetraploid sour cherry (*Prunus cerasus* L.) 'Montmorency' identifies three distinct ancestral Prunus genomes. Hortic Res. 10(7):uhad097. doi:10.1093/hr/uhad097.

Hruby, K. 1950. The cytology of tetraploid cherries. Stud Bot Cechoslov-. 11(3):87–97.

Iezzoni AF, Hancock AM. 1984. A Comparison of Pollen Size in Sweet and Sour Cherry. HortScience. 19(4):560–562. doi:10.21273/HORTSCI.19.4.560.

Jin W, Wang H, Li M, Wang J, Yang Y, Zhang X, Yan G, Zhang H, Liu J, Zhang K. 2016. The R2R3 MYB transcription factor *PavMYB10.1* involves in anthocyanin biosynthesis and determines fruit skin colour in sweet cherry (*Prunus avium* L.). Plant Biotechnol J. 14(11):2120–2133. doi:10.1111/pbi.12568.

Lin-Wang K, Bolitho K, Grafton K, Kortstee A, Karunairetnam S, McGhie TK, Espley RV, Hellens RP, Allan AC. 2010. An R2R3 MYB transcription factor associated with regulation of the anthocyanin biosynthetic pathway in *Rosaceae.* BMC Plant Biol. 10(1):50. doi:10.1186/1471-2229-10-50.

Quadrana L, Bortolini Silveira A, Mayhew GF, LeBlanc C, Martienssen RA, Jeddeloh JA, Colot V. 2016. The *Arabidopsis thaliana* mobilome and its impact at the species level. Zilberman D, editor. eLife. 5:e15716. doi:10.7554/eLife.15716.

Stegmeir T, Cai L, Basundari FRA, Sebolt AM, Iezzoni AF. 2015. A DNA test for fruit flesh color in tetraploid sour cherry (Prunus cerasus L.). Mol Breed. 35(7):149. doi:10.1007/s11032-015-0337-y.

Tsukamoto T, Hauck NR, Tao R, Jiang N, Iezzoni AF. 2010. Molecular and genetic analyses of four nonfunctional S haplotype variants derived from a common ancestral S haplotype identified in sour cherry (Prunus cerasus L.). Genetics. 184(2):411–27. doi:10.1534/genetics.109.109728.

Vasudevan A, Lévesque-Lemay M, Edwards T, Cloutier S. 2023. Global transcriptome analysis of allopolyploidization reveals large-scale repression of the D-subgenome in synthetic hexaploid wheat. Commun Biol. 6(1):1–14. doi:10.1038/s42003-023-04781-7.

Wang D, Karle R, Iezzoni AF. 2000. QTL analysis of flower and fruit traits in sour cherry. Theor Appl Genet. 100(3–4):535–544. doi:10.1007/s001220050070.

Wang, Dechun. 1998. RFLP Mapping, QTL Identification, and Cytogenetic Analysis in Sour Cherry [Doctor of Philosophy]. [East Lansing, MI]: Michigan State University.

Wöhner TW, Emeriewen OF, Wittenberg AHJ, Schneiders H, Vrijenhoek I, Halász J, Hrotkó K, Hoff KJ, Gabriel L, Lempe J, et al. 2021. The draft chromosome-level genome assembly of tetraploid ground cherry (Prunus fruticosa Pall.) from long reads. Genomics. 113(6):4173–4183. doi:10.1016/j.ygeno.2021.11.002.