Genome Synteny Has Been Conserved Among the Octoploid Progenitors of Cultivated Strawberry Over Millions of Years of Evolution

Michael A. Hardigan[1], Mitchell J. Feldmann[1], Anne Lorant[1], Kevin A. Bird[2], Randi Famula[1], Charlotte Acharya[1], Glenn Cole[1], Patrick P. Edger[2] and Steven J. Knapp[1]*

[1] Department of Plant Sciences, University of California, Davis, Davis, CA, United States, [2] Department of Horticulture, Michigan State University, East Lansing, MI, United States

Allo-octoploid cultivated strawberry (*Fragaria × ananassa*) originated through a combination of polyploid and homoploid hybridization, domestication of an interspecific hybrid lineage, and continued admixture of wild species over the last 300 years. While genes appear to flow freely between the octoploid progenitors, the genome structures and diversity of the octoploid species remain poorly understood. The complexity and absence of an octoploid genome frustrated early efforts to study chromosome evolution, resolve subgenomic structure, and develop a single coherent linkage group nomenclature. Here, we show that octoploid *Fragaria* species harbor millions of subgenome-specific DNA variants. Their diversity was sufficient to distinguish duplicated (homoeologous and paralogous) DNA sequences and develop 50K and 850K SNP genotyping arrays populated with co-dominant, disomic SNP markers distributed throughout the octoploid genome. Whole-genome shotgun genotyping of an interspecific segregating population yielded 1.9M genetically mapped subgenome variants in 5,521 haploblocks spanning 3,394 cM in *F. chiloensis* subsp. *lucida*, and 1.6M genetically mapped subgenome variants in 3,179 haploblocks spanning 2,017 cM in *F. × ananassa*. These studies provide a dense genomic framework of subgenome-specific DNA markers for seamlessly cross-referencing genetic and physical mapping information and unifying existing chromosome nomenclatures. Using comparative genomics, we show that geographically diverse wild octoploids are effectively diploidized, nearly completely collinear, and retain strong macro-synteny with diploid progenitor species. The preservation of genome structure among allo-octoploid taxa is a critical factor in the unique history of garden strawberry, where unimpeded gene flow supported its origin and domestication through repeated cycles of interspecific hybridization.

Keywords: *Fragaria*, strawberry, polyploidy, genome evolution, domestication

INTRODUCTION

Interspecific homoploid hybridization and polyploidy-inducing hybrid events have been creative forces in plant genome evolution and speciation, acting as catalysts for *de novo* reorganization of chromosome structure (Jiao et al., 2011; Yakimowski and Rieseberg, 2014; Soltis et al., 2014b; Soltis et al., 2014a; Vallejo-Marín et al., 2015; McKain et al., 2016; Soltis et al., 2016; Wendel et al., 2016; Alix et al., 2017; Mandáková et al., 2019). The cultivated strawberry (*Fragaria × ananassa* Duchesne ex Rozier) is unique among domesticated crop species because it arose through both processes. The chromosomes of octoploid garden strawberry (2n = 8x = 56) evolved through a combination of ancient polyploidy, and repeated homoploid hybridization in the last three centuries (Duchesne, 1766; Darrow, 1966). The presence of duplicated (homoeologous) chromosomes in plants frequently leads to meiotic anomalies and associated chromosomal rearrangements, e.g., translocations and inversions, that reduce or eliminate gene flow between the donors and their polyploid offspring (Soltis et al., 2014a; Soltis et al., 2014b; Alix et al., 2017; Latta et al., 2019; Mandáková et al., 2019). Similarly, meiotic mispairing in interspecific homoploid hybrids can lead to rearranged offspring chromosomes that differ from the chromosomes of one or both parents, resulting in reproductive barriers and hybrid speciation, as has been widely documented in sunflower (*Helianthus*) and other plants (Rieseberg, 1997; Burke et al., 2004; Abbott et al., 2010; Barb et al., 2014; Yakimowski and Rieseberg, 2014). However, reproductive barriers among octoploid *Fragaria* taxa remain essentially nonexistent, fueling the recurrence of interspecific homoploid hybridization in the origin, domestication, and modern-day breeding of *F. × ananassa*.

The modern *F. × ananassa* lineage traces its origin to extinct cultivars developed in western Europe in the 1700s. These cultivars were interspecific hybrids of non-sympatric wild octoploids from the New World: *F. chiloensis* subsp. *chiloensis* from South America and *F. virginiana* subsp. *virginiana* from North America (Darrow, 1966). Repeated introgression of alleles from diverse subspecific ecotypes of *F. virginiana* and *F. chiloensis* defined the later generations, coinciding artificial selection of horticulturally important traits among hybrid descendants in Europe and North America. Modern cultivars have emerged from 250 years of global migration and breeding within this admixed population (Darrow, 1966; Hardigan et al., 2018). Because alleles have been introgressed from up to eight octoploid subspecies, the genomes of modern *F. × ananassa* individuals are mosaics of their wild ancestors (Liston et al., 2014; Hardigan et al., 2018). Since the discovery of artificial hybrids at the Gardens of Versaille (Duchesne, 1766), natural interspecific hybrids (*F. × ananassa* subsp. *cuneifolia*) were discovered in zones of sympatry between *F. chiloensis* subsp. *pacifica* and *F. virginiana* subsp. *platypetala* in western North America (Hancock and Bringhurst, 1979; Luby et al., 1992; Staudt, 1999; Salamone et al., 2013). Neither cultivated *F. × ananassa* or wild *F. × ananassa* subsp. *cuneifolia* is reproductively isolated from their octoploid progenitors. Thus, genes appear to flow freely between the wild octoploid progenitors, and between the hybrids and their progenitors. While some genomic rearrangements have been identified between homoeologous chromosomes, and relative to the diploid species (Tennessen et al., 2014; van Dijk et al., 2014), the apparent absence of reproductive isolation implies that homoploid and polyploid hybridization events have not produced significant chromosome rearrangements among octoploid taxa. We hypothesized that the octoploids carry nearly collinear chromosomes tracing to the most recent common ancestor, despite one million years of evolution which produced multiple recognized species and subspecies.

The octoploid strawberry genome has been described as "notoriously complex" and an "extreme example of difficulty" for study (Folta and Davis, 2006; Hirsch and Buell, 2013; Hirakawa et al., 2014; Koskela et al., 2016). While GBS, GWAS, and NGS-reliant applications are relatively straightforward in organisms with well-characterized reference genomes, such approaches were previously difficult or intractable in octoploid strawberry (Liston et al., 2014; Tennessen et al., 2014; Bassil et al., 2015; Vining et al., 2017). Genetic studies in octoploid strawberry previously relied on the genome of woodland strawberry (*F. vesca*), an extant relative to one of four diploid subgenomes contained in *F. × ananassa* (Shulaev et al., 2011; Edger et al., 2017). For nearly a decade, the *F. vesca* genome was the only framework available for DNA variant discovery, gene discovery, genetic mapping, and genome-wide association studies in octoploid strawberry (Tennessen et al., 2014; Bassil et al., 2015; Davik et al., 2015; Vining et al., 2017; Pincot et al., 2018). The development of a chromosome-scale reference genome for *F. × ananassa* (Edger et al., 2019b) provided the physical framework needed to overcome previous barriers, and explore the organization and evolution of its progenitor genomes.

Here, we report the first study of chromosome evolution and genome structure in octoploid *Fragaria* using an octoploid genome-guided approach to DNA variant discovery and comparative mapping. We demonstrate the ability to differentiate duplicated (homoeologous) octoploid sequences using both NGS and array-based genotyping technologies when applied in conjunction with an octoploid reference genome. In doing so, we overcome a long-standing technical hurdle that has impeded efforts to study strawberry subgenome diversity and chromosome evolution. We estimated strawberry subgenomic diversity by whole-genome shotgun (WGS) sequencing of 93 genealogically and phylogenetically diverse *F. × ananassa*, *F. chiloensis*, and *F. virginiana* individuals. The frequency of unique WGS sequence alignment to the octoploid strawberry genome was characteristic of many diploid plant species (Hamilton and Buell, 2012; Lee and Schatz, 2012; Schatz et al., 2012; Treangen and Salzberg, 2012), and permitted the identification of millions of subgenome-specific DNA variants, effectively distinguishing homologous and homoeologous DNA sequences on every chromosome. Using the genetic diversity of *F. × ananassa*, we developed publicly available 50K and 850K SNP arrays populated with subgenome anchored marker probes for octoploid genetic mapping and forward genetic studies. We then performed high-density

genetic mapping of five octoploids representing $F. \times ananassa$ and both its octoploid progenitor species using a combination of WGS-based and array-based genotyping. Telomere-to-telomere genetic mapping of nearly every chromosome was enabled by the conserved disomic segregation observed in populations derived from wild species and $F. \times ananassa$, underscoring the effective diploidization and meiotic stability of octoploid *Fragaria*. Comparative mapping of $F. \times ananassa$ and multiple subspecies of *F. chiloensis* and *F. virginiana* revealed the macro-syntenic structure of the cultivated reference genotype (Camarosa) and its progenitor species were nearly identical.

The collinear and diploidized genomes of $F. \times ananassa$ and its progenitors support octoploid *Fragaria* as an evolutionary clade which achieved a relatively high degree of genome stability prior to the speciation and sub-speciation of *F. chiloensis* and *F. virginiana*. The interspecific origin of $F. \times ananassa$ followed by successive hybridization throughout domestication is an unusual improvement pathway that frequently contributes to reproductive incompatibility or sterility in wide species hybrids (Ladizinsky, 1985; Hughes et al., 2007; Miller and Gross, 2011; Meyer and Purugganan, 2013). The preservation of genome structure among diverse octoploid *Fragaria* species and subspecies was likely essential to the unique history of the $F. \times ananassa$ lineage, which has undergone repeated cycles of homoploid hybridization without the formation of reproductive barriers or loss of fertility.

RESULTS AND DISCUSSION

Subgenomic Diversity of Octoploid Fragaria

We performed the first deep exploration of the homoeologous sequence diversity of octoploid *Fragaria* using the Camarosa v1.0 octoploid reference genome (Edger et al., 2019b) and a diversity panel of 93 strawberry individuals, including 47 $F. \times ananassa$, 24 F. chiloensis, and 22 F. virginiana individuals (**Table S1**). By incorporating subgenome specificity at the assembly level, previous barriers to copy-specific sequence alignment caused by the octoploid ancestral homology of strawberry posed a less significant obstacle than repetitive DNA elements in diploid genomes such as maize (Hamilton and Buell, 2012; Lee and Schatz, 2012; Schatz et al., 2012; Treangen and Salzberg, 2012). The fraction of uniquely aligning (MapQ > 0) PE150 sequences averaged 83.2% (**Figure S1**), and 90.9% of Camarosa PE250 sequences aligned uniquely, allowing comprehensive coverage and analysis of subgenomic diversity. Using genotype calling software FreeBayes and a series of hard-filters targeting unique sequence alignments, we identified 41.8M subgenomic SNP and INDEL mutations in $F. \times ananassa$ and its wild progenitors.

$F. \times ananassa$ has been described as "genetically narrow" due to the small number of founders in the pedigrees of modern cultivars (Sjulin and Dale, 1987; Dale and Sjulin, 1990; Stegmeir et al., 2010). Despite a small effective population size, our analyses show that massive genetic diversity has been preserved in $F. \times ananassa$, with negligible difference between

wild species and domesticated germplasm. The subgenome nucleotide diversity (π) of $F. \times ananassa$ ($\pi = 5.857 \times 10^{-3}$) was equivalent to wild progenitors *F. chiloensis* ($\pi = 5.854 \times 10^{-3}$) and *F. virginiana* ($\pi = 5.954 \times 10^{-3}$), and comparable to the sequence diversity of *Zea mays* landraces ($\pi = 4.9 \times 10^{-3}$) and wild *Zea mays* spp. *parviglumis* progenitors ($\pi = 5.9 \times 10^{-3}$) (Hufford et al., 2012). Correlations of $F. \times ananassa$, *F. chiloensis*, and F. *virginiana* diversity across the 28 octoploid chromosomes ranged from 0.93–0.97, showing that the magnitude and distribution of genomic diversity are broadly conserved among octoploid taxa. This suggested that $F. \times ananassa$ was not strongly bottlenecked by domestication, or that its domestication bottleneck was mitigated by continued introgression of allelic diversity from wild subspecies (Darrow, 1966). We found that variance in the distribution of octoploid nucleotide diversity was influenced more significantly by subgenome ancestry than domestication and breeding (**Table S2**). The diploid *F. vesca* subgenome, dominant with respect to gene abundance and expression (Edger et al., 2019b), contained the least diverse homoeolog of every ancestral chromosome, while subgenomes derived from the ancestors of the extant Asian species (*F. iinumae* and *F. nipponica*) contained greater genetic diversity (**Table S2**). Polyploid genome dominance contributes to differences in expression, gene loss, and purifying selection across subgenomes on the path to diploidy (Schnable et al., 2011; Grover et al., 2012; Parkin et al., 2014; Bird et al., 2018). Reduced nucleotide diversity on the dominant *F. vesca* subgenome supported distinct levels of purifying selection for the ancestral chromosomes of octoploid strawberry, with stronger selection in *F. vesca* ancestral sequences.

Because $F. \times ananassa$ was domesticated as an interspecific hybrid, individual performance was hypothesized to have been improved by allelic diversity between *F. chiloensis* and *F. virginiana*. To support comparisons of strawberry heterozygosity with previously studied polyploid species, we estimated individual heterozygosity based on the genomic frequency of heterozygous nucleotides (nts), a metric previously used in potato (Hardigan et al., 2017), and the frequency of heterozygosity at GBS-derived polymorphic sites, metrics previously used in blueberry and cotton (Page et al., 2013; de Bem Oliveira et al., 2019). Strawberry genomic heterozygosity ranged from 0.02-0.80% and averaged 0.46% genomic nts. This translated to an average of 11.1% heterozygosity at polymorphic marker sites. The most heterozygous octoploid individuals were early interspecific hybrids: White Carolina (PI551681; 0.80% nts), Peruvian Ambato (PI551736; 0.72% nts), Jucunda (PI551623; 0.70% nts), and Ettersberg 121 (PI551904; 0.66% nts). The average subgenomic heterozygosity of octoploid strawberry (0.46% nts) was below diploid potato (1.05% nts) and tetraploid potato (2.73% nts) (Hardigan et al., 2017). The average heterozygosity of octoploid strawberry at GBS-derived polymorphic sites was below the allo-tetraploid cotton A-genome (13% marker sites), above the cotton D-genome (< 1% marker sites) (Page et al., 2013), and below autotetraploid blueberry (32.4% marker sites) (de Bem Oliveira et al., 2019). Due to the presence of four ancestral homoeologs, traditional models of "fixed heterozygosity" applied to allopolyploid

species (Comai, 2005; Obbard et al., 2006) assume an octoploid functional heterozygosity four-fold greater than subgenomic estimates. Under this model, recent *F. virginiana* x *F. chiloensis* hybrids such as White Carolina and Jucunda could be regarded as similarly heterozygous to autopolyploid species such as potato. However, assembly of the allo-octoploid strawberry genome uncovered rampant gene silencing, gene loss, and homoeologous exchanges relative to diploid ancestors (Edger et al., 2019b), eroding the conservation of ancestral allele function. The frequency of unique sequence alignment (**Figure S1**) and unbroken distribution of subgenomic variant detection (**Figure S2**) in our analyses underscore the extensive divergence of the four subgenomes. Thus, traditional polyploid allele dosage models assuming genome-wide fixed heterozygosity may be of limited usefulness for strawberry.

Recombination Breakpoint Mapping of Octoploid Strawberry

We used WGS sequence analysis and recombination breakpoint mapping of an octoploid strawberry population to explore the breadth of disomic variation as an indicator of bivalent pairing during meiosis. Several cytogenetic and DNA marker studies have proposed the occurrence of polysomy in strawberry (Fedorova, 1946; Senanayake and Bringhurst, 1967; Lerceteau-Köhler et al., 2003), while others suggest that octoploids are mainly disomic (Byrne and Jelenkovic, 1976; Arulsekar and Bringhurst, 1981; Bringhurst, 1990; Kunihisa et al., 2005). We performed low-coverage (4-8x) sequencing and subgenomic DNA variant calling in a population ($n = 182$) derived from a cross between the *F. × ananassa* cultivar 'Camarosa' and the beach strawberry (*F. chiloensis* subsp. *lucida*) ecotype 'Del Norte'. These parents were selected to provide a dense comparison of profiles of mappable disomic polymorphisms in wild and domesticated octoploid individuals. Variant calling against the Camarosa v1.0 genome identified 3.7M subgenomic SNPs and INDELs inherited from 1.6M Camarosa heterozygous sites (AB × AA), 1.9M *F. chiloensis* subsp. *lucida* heterozygous sites (AA × AB), and 0.2M co-heterozygous sites (AB × AB). We used the high-density DNA variant data to perform haplotype mapping based on recombination breakpoint prediction and evaluated segregation ratios of parental alleles across the 28 octoploid chromosomes.
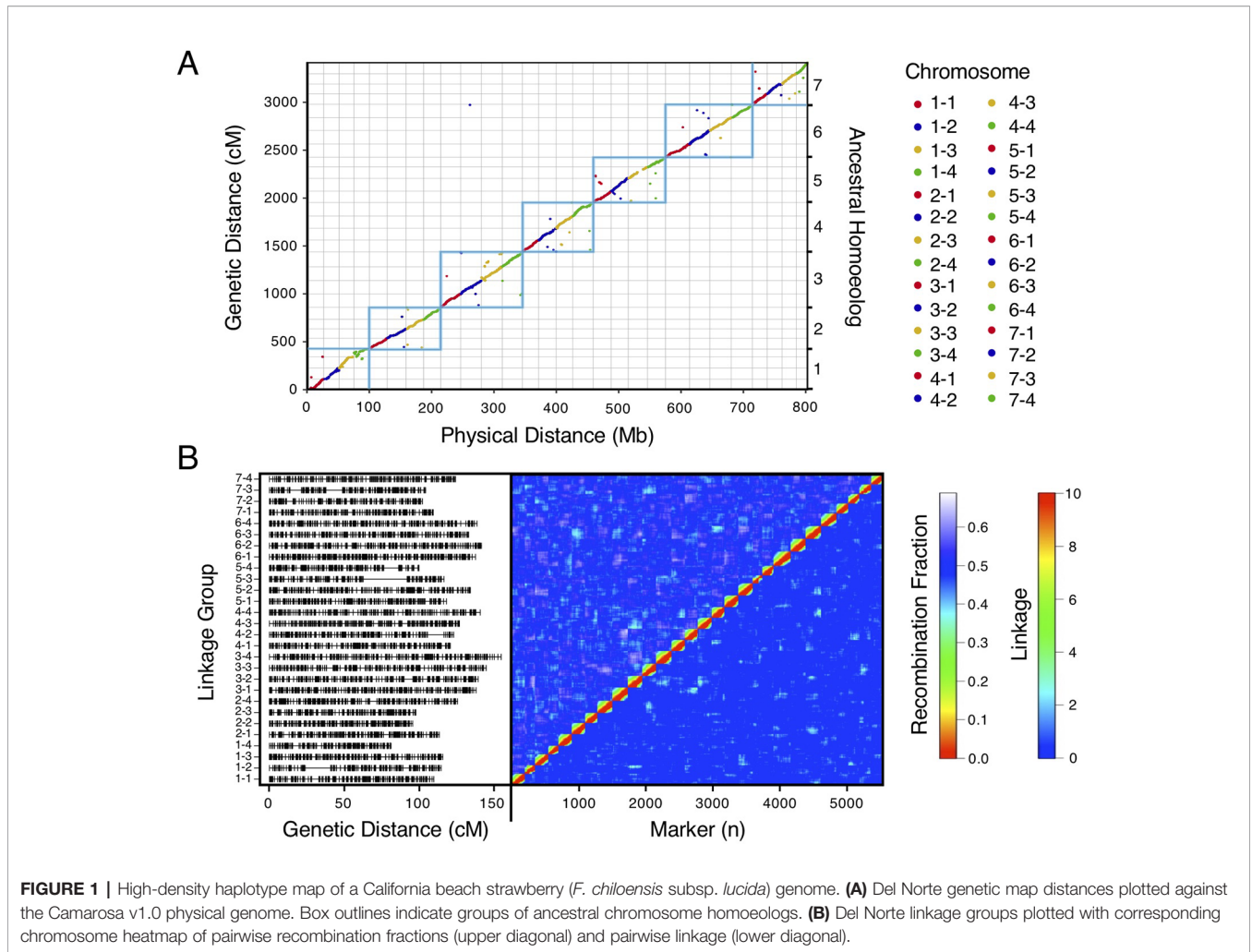
We bypassed the computational demand of analyzing pairwise linkage across millions of DNA variants with missing data and genotyping errors by implementing the haplotype calling approach proposed by Huang et al. (2009) and Marand et al. (2017). Our approach performed a sliding-window analysis to predict crossover events, then estimated the consensus of co-segregating DNA variation between recombination breakpoints to reconstruct the representative genotypes of each haploblock. The haploblocks were mapped as unique genetic markers. Using this approach, we mapped 1.9M *F. chiloensis* subsp. *lucida* DNA variants in 5,521 haploblocks spanning 3,393.86 cM, and 1.6M Camarosa DNA variants in 3,179 haploblocks spanning 2,016.95 cM (**Dataset S1**). The paternal *F. chiloensis* subsp. *lucida* map produced telomere-to-telomere coverage of the 28 octoploid chromosomes (**Figure 1**), providing the most comprehensive

genetic map of an octoploid *Fragaria* genome to-date. The complete mapping of the extant homoeologs for all seven ancestral *Fragaria* chromosomes in the paternal genome and analysis of chromosome-wide segregation distortion (**Figure 2**) showed that disomic recombination is ubiquitous in the genome of *F. chiloensis*. By contrast, less than 50% of the Camarosa genome could be mapped on chromosomes 1-1, 1-2, 2-4, 3-3, 5-2, 6-2, 6-3, 6-4, and 7-3 (**Figure S3**). We analyzed Camarosa heterozygosity and segregation distortion to determine whether the inability to map large segments of the genome was the result of polysomic recombination in *F. × ananassa*. This uncovered a near total loss of polymorphism in the unmapped regions of Camarosa (**Figure 2**), showing that incomplete mapping of *F. × ananassa* resulted from depletion of heterozygosity in the hybrid genome, not polysomy. Sargent et al. (2012) previously reported extensive regions of homozygosity that affected mapping of *F. × ananassa*. Artificial selection pressure in commercially bred hybrids almost certainly accounts for the lower subgenomic heterozygosity of Camarosa relative to *F. chiloensis* subsp. *lucida*, which does not support a critical role for genome-wide interspecific heterozygosity in driving cultivar performance.

850K Octoploid Screening Array

We designed Affymetrix SNP genotyping arrays populated with subgenome-specific marker probes to enable genetic mapping, genome-wide association studies (GWAS), and genomic prediction in octoploid strawberry. DNA variants were selected for array design from the subgenomic diversity identified in the WGS panel (**Figure 3**). From 90M total unfiltered variant sites, we extracted 45M unfiltered variants that segregated in *F. × ananassa*. To identify candidate DNA variants for marker design, we selected only biallelic SNPs above a low-diversity threshold ($\pi \geq 0.05$), excluded rare alleles (MAF ≥ 0.05), required a VCF quality score > 20, and excluded sites with > 15% missing data in the diversity panel. These filters yielded 8M subgenomic SNPs segregating within the *F. × ananassa* subset of the diversity panel. We obtained 71-nt marker probes by extracting 35-nt sequences flanking each SNP site from the Camarosa v1.0 genome assembly. Marker probes for the 8M high-confidence SNP sites were then filtered to remove candidates that were problematic for array tiling. These included duplicate or near-duplicate probe sequences and probes that inherited ambiguous reference sequences (Ns), required double-tiling (A/T or C/G alleles), or that Affymetrix scored as having low buildability. We retained 6.6M probes that targeted high-confidence *F. × ananassa* variants and were acceptable for array tiling.

We applied three selection criteria for determining a subset of 850K marker probes for tiling: likelihood of probe binding interference by off-target variants, likelihood of off-target (non-single copy) probe binding, and physical genome distribution. The likelihood of probe binding interference was scored as the sum of non-reference allele frequencies for off-target variants in the 35-nt binding region adjacent to the target SNP. The likelihood of off-target probe binding was scored by performing BLAST alignment of the 71-nt probe sequences to the Camarosa v1.0 genome and quantifying the number of off-

FIGURE 1 | High-density haplotype map of a California beach strawberry (*F. chiloensis* subsp. *lucida*) genome. **(A)** Del Norte genetic map distances plotted against the Camarosa v1.0 physical genome. Box outlines indicate groups of ancestral chromosome homoeologs. **(B)** Del Norte linkage groups plotted with corresponding chromosome heatmap of pairwise recombination fractions (upper diagonal) and pairwise linkage (lower diagonal).

target alignments with query coverage above 90% and sequence identity above 90%. We then iteratively parsed the Camarosa v1.0 genome using 10 kilobase (kb) non-overlapping physical windows, extracting the best available marker from each window based on probe binding interference and off-target binding likelihoods, until reaching an 850K probe threshold. We reserved 16K positions for legacy markers from the iStraw SNP array (Bassil et al., 2015; Verma et al., 2016) that were polymorphic in a previous strawberry diversity study (Hardigan et al., 2018). The set of 850K probe sequences was submitted to Affymetrix for constructing a screening array.

We genotyped a genetically diverse sample of 384 octoploid strawberry accessions to validate subgenome-specific marker performance on the 850K screening array (**Dataset S2**). The sample fluorescence files were analyzed with the Axiom Suite in polyploid mode to generate marker clusters. Collectively, 446,644 of 850,000 marker probes produced QC-passing polymorphic SNP genotype clusters showing disomic (allopolyploid) segregation. Among these, 78.3% were classified as "PolyHighResolution" in the Axiom terminology, producing diploid co-dominant genotype clusters (AA, AB, and BB)

without detecting off-target allelic variation. Similarly, 18.8% of markers classified as "NoMinorHomozygote" in the Axiom terminology produced dominant genotype clusters in which heterozygotes (AB) clustered with one of the homozygous genotype classes (AA or BB). The remaining 2.9% of the markers detected non-target alleles and were classified as "OffTargetVariant" markers in the Axiom terminology. The HomRO statistic generated by the Axiom Suite estimates genotype cluster separation, and has been used as a metric to infer octoploid single-copy (i.e. subgenome or paralog specific) probe binding when values exceed 0.3 (Bassil et al., 2015). Based on this threshold, 74% of the QC-passing marker probes on the 850K screening array exhibited single-copy binding, in addition to measuring subgenome-specific DNA variation (**Figure S4**). The complete set of 446,644 validated probes is made available for public use (**Dataset S3**).

50K Octoploid Production Array
We selected 49,483 polymorphic marker probes from the 850K validated probe set to build a 50K production array (**Dataset S4**). We included 5,809 LD-pruned ($r^2 < 0.50$) marker probes from
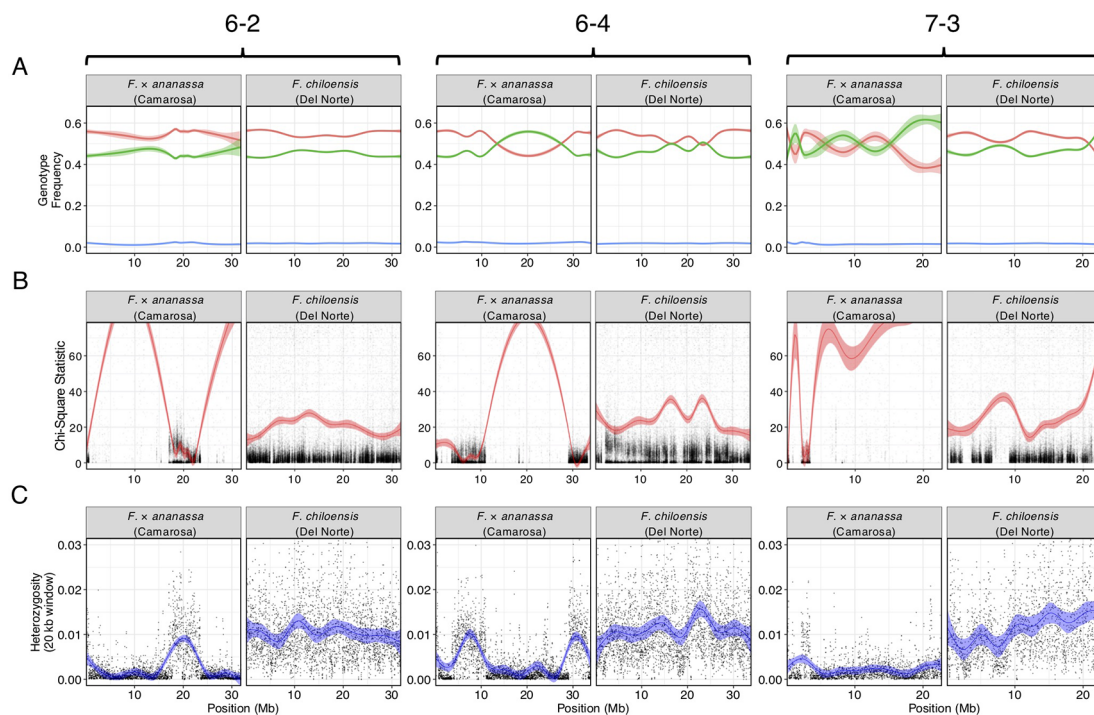
FIGURE 2 | Comparison of genomic heterozygosity and diploid segregation distortion in *F. × ananassa* (Camarosa) and *F. chiloensis* (Del Norte) on three chromosomes (6-2, 6-4, 7-3). **(A)** Frequency of AA (green), AB (red), and off-target (BB; blue) genotypes for polymorphic markers in the mapping population. **(B)** Chi-square statistic estimating segregation distortion of polymorphic markers in the mapping population. **(C)** Heterozygous nucleotide frequency of parent genotypes in 20 kb physical windows.

the iStraw35 SNP array to support cross-referencing with previous genetic and QTL mapping studies. We targeted 2,878 genes based on Camarosa v1.0 functional annotations that indicated R-gene affiliated protein domains (Edger et al., 2019b) or homology to *Fragaria vesca* genes involved in flowering and fruit development expression networks (Kang et al., 2013; Hollender et al., 2014). Candidate genes were pre-allocated up to two markers (within 1 kb) from the screening panel. We next selected a set of the most commonly segregating markers to support genetic mapping. We identified this set by selecting the marker with the highest pairwise diversity (π) in *F. × ananassa* across non-overlapping 50 kb physical genome windows. The remainder of the 50K array was populated by iteratively parsing the genome with 50 kb physical windows and selecting random QC-passing markers to provide an unbiased genome distribution. Both the 850K and 50K probe sets provide unbroken, telomere-to-telomere physical coverage of the 28 octoploid strawberry chromosomes (**Figure S5**). Within the 50K probe set, 53% of the probes were located within genes, and 79% were located within 1 kb of a gene. The 50K probe set was provided to Affymetrix for building the production array.

We screened 1,421 octoploid samples from multiple bi-parental populations and a large diversity panel on the 50K production array. Collectively, 42,081 markers (85%) successfully replicated QC-passing polymorphic genotype clusters when screened in the larger sample group. Of the 7,402 non-replicated markers, only

1% were excluded due to becoming monomorphic ("MonoHighResolution" Axiom class) or from increased missing data ("CallRateBelowThreshold" Axiom class). Sub-clustering and increased dispersion within the AA or BB genotype clusters ("AAvarianceX", "AAvarianceY", "BBvarianceX", "BBvarianceY" Axiom classes) accounted for 20% on non-replicated markers. The Axiom software provided no specific cause for failure for the remaining 79% of non-replicated markers. These results suggest that increasing the size and diversity of a genotyping population may affect the reproducibility of a fraction (15%) of markers on the 50K array, while a majority (85%) are highly reproducible. The fraction of polymorphic co-dominant ("PolyHighResolution") markers increased from 78% on the 850K screening array to 86% among reproducible markers 50K production array. The average correlation between genotypes predicted for the Camarosa × Del Norte bi-parental population based on low-coverage sequencing and 50K array probes at polymorphic segregating sites was 0.93.

Genetic Mapping of Wild Octoploid Ecotypes

We demonstrated that the 50K SNP array allows dense genetic mapping of heterozygous regions on all 28 chromosomes of *F. × ananassa*, *F. chiloensis*, and *F. virginiana*. Genetic mapping of *F. chiloensis* and *F. virginiana* provided telomere-to-telomere physical representation of the 28 octoploid chromosomes and near-complete representation within individual wild maps (**Dataset S1**). We
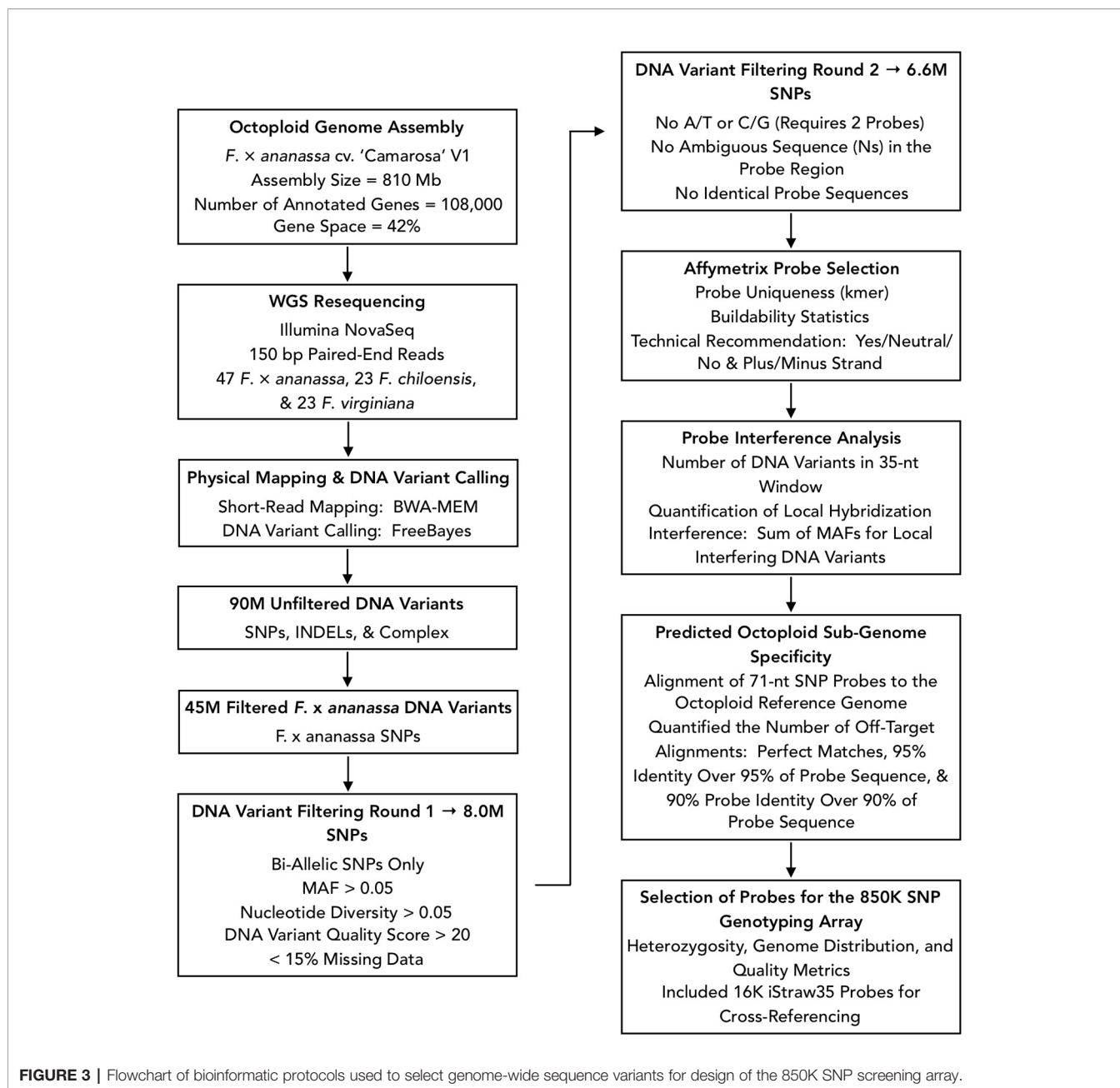
FIGURE 3 | Flowchart of bioinformatic protocols used to select genome-wide sequence variants for design of the 850K SNP screening array.

selected four octoploid parents from two outcrossing bi-parental populations genotyped on the 50K array for mapping. The first was the Camarosa × Del Norte population used for WGS recombination breakpoint mapping ($n = 182$). The second population was derived from a cross between *F. virginiana* subsp. *virginiana* accession PI552277 (female parent) and *F. virginiana* subsp. *virginiana* accession PI612493 (male parent) ($n = 96$). The number of SNPs segregating in the octoploid parent genotypes varied considerably (**Table S3**). Camarosa contained the most segregating markers (9,062), followed by PI552277 (5,575), PI612493 (5,464), and Del Norte (2,368). The larger number of markers segregating in Camarosa relative to the wild parent genotypes reflected the array

design strategy, which targeted *F. × ananassa* diversity. The unbalanced representation of *F. virginiana* and *F. chiloensis* diversity was not expected, because genome-wide variant calls showed similar profiles of heterozygosity and nucleotide diversity in the wild progenitors, and showed that *F. chiloensis* subsp. *lucida* was more heterozygous than Camarosa. The higher level of ascertainment bias against *F. chiloensis* diversity that resulted from probing *F. × ananassa* alleles supports previous findings that *F. virginiana* diversity is more prevalent in cultivated hybrids (Hardigan et al., 2018). We mapped heterozygous variant sites of the four parent octoploids using software ONEMAP (Margarido et al., 2007) to generate initial linkage groups and markers orders

and BatchMap (Schiffthaler et al., 2017) for marker re-ordering and genetic distance estimation. Despite the ascertainment bias for domesticated allelic diversity on the 50K array, the relatively unbiased distribution of genomic heterozygosity in wild genotypes (**Table S3**) provided a more complete representation of the wild octoploid genomes than F. × ananassa (Camarosa) (**Figure 4**). Large homozygous regions that produced breaks in the Camarosa WGS haplotype map and 50K array map (chromosomes 1-1, 1-2, 2-4, 3-3, 5-2, 6-2, 6-3, 6-4, and 7-3) were clearly featured in the wild genetic maps (**Figure 4**). Camarosa contained an average of 11.7 (± 6.8) SNPs/megabase (Mb) across 28 chromosomes, with as many as 25.2 SNPs/Mb (1-4) and as few as 0.8 SNPs/Mb (1-3), underscoring the scattered distribution of mappable subgenomic diversity in the cultivated strawberry hybrid.

The wild octoploid maps revealed large (Mb+) chromosomal rearrangements relative to the Camarosa v1.0 physical genome on chromosomes 1-2, 1-4, 2-1, 2-3, 6-2, and 6-4. These rearrangements were conserved across the wild species genomes, and supported by corresponding regions represented in the Camarosa genetic map (1-2, 1-4, 2-1) (**Figure S3**), indicating intra-chromosomal scaffolding errors in the physical reference genome. The fraction of SNPs that genetically mapped to non-reference chromosomes ranged from 1.5%–1.9% in the four parents, with the highest fraction observed in Camarosa. This indicated minimal inter-chromosomal errors in the physical genome and minimal inter-chromosomal marker discordance between octoploid progenitor species. Thus, the octoploid genetic maps provided no evidence of chromosome rearrangement among the wild and cultivated octoploid species.

Genome Structure of Ancestral Species

Previous studies have reported octoploid chromosome rearrangements relative to diploid Fragaria, potentially contributing to sex determination (Spigler et al., 2008; Spigler et al., 2010; Tennessen et al., 2014; Govindarajulu et al., 2015). Moreover, there is phylogenetic evidence of chromosome exchanges among the four ancestral subgenomes (Liston et al., 2014; Edger et al., 2019b). However, there is no evidence for chromosome-scale structural variation between octoploid taxa. It remains unclear to what extent the structural variation of octoploid Fragaria reflects initial polyploid 'genome shock' occurring in the common ancestor, as opposed to ongoing mutations contributing to octoploid species diversification. Through comparative mapping, we show that the genomes of diverse octoploid ecotypes contributing to the homoploid hybrid lineage of F. × ananassa are nearly completely collinear. We constructed genetic maps for two additional wild genotypes, F. chiloensis subsp. pacifica (SAL3) and F. virginiana subsp. platypetala (KB3) using publicly available DNA capture libraries to obtain a more diverse set of octoploid subspecific taxa. We aligned capture sequences from an F. chiloensis subsp. chiloensis × F. chiloensis subsp. pacifica population (GP33 × SAL3, n = 46; Tennessen et al., 2014) and an F. virginiana subsp. platypetala × F. virginiana subsp. platypetala population (KB3 × KB11, n = 46; Tennessen et al., 2018) to the Camarosa v1.0 genome assembly and predicted subgenome DNA variant

genotypes using FreeBayes. Genetic mapping of the DNA capture markers followed the protocol used for the 50K SNP datasets. Using the 50K array linkage groups and DNA capture linkage groups (**Dataset S1**), we performed comparative mapping of four octoploid subspecies—F. chiloensis subsp. lucida (Del Norte), F. chiloensis subsp. pacifica (SAL3), F. virginiana subsp. platypetala (KB3), and F. virginiana subsp. virginiana (PI552277)—in ALLMAPS using genetically mapped DNA variant sites anchored to 50-kb physical genome windows in Camarosa v1.0. The chromosomes of the octoploid progenitor subspecies were completely syntenic (**Figure 5**, **Figure S6**). Based on these results, large-scale chromosome rearrangements in octoploid Fragaria relative to the diploid ancestral genomes would have occurred before the speciation of F. chiloensis and F. virginiana.

Comparative genomic analysis of the four octoploid strawberry subgenomes, in addition to diploid strawberry species F. vesca and F. iinumae and black raspberry (Rubus occidentalis), revealed a remarkable degree of karyotypic stability not only within octoploid strawberry, but within related species that have evolved independently for millions of years (**Figure 6**). While several rearrangements and inversions were detected between ancestral diploid chromosomes, they primarily consisted of large, syntenic blocks with high degrees of collinearity (**Figure 6**). These results demonstrate that the structural conservation retained between octoploid species extends to broader Fragaria taxa, including strawberry species that are estimated to have diverged ~7.5 million years ago and beyond that to a distantly related Rosaceous species estimated to have diverged 33 M years ago (Njuguna et al., 2013; Qiao et al., 2016). This is unique compared to extensive karyotype evolution documented in polyploid species involving numerous chromosomal fusion and rearrangement events, e.g. Camelina sativa (Mandáková et al., 2019).

Unification of Octoploid Chromosome Nomenclatures

Previous octoploid genetic mapping studies relied on a variety of DNA marker technologies including early PCR-based assays (e.g., simple sequence repeats and amplified fragment length polymorphisms) and technical ploidy reduction of sequence variants called against the diploid F. vesca genome (Rousseau-Gueutin et al., 2008; Spigler et al., 2008; Spigler et al., 2010; Tennessen et al., 2014; Bassil et al., 2015; Vining et al., 2017; Tennessen et al., 2018). This diversity of DNA marker genotyping strategies without information linking to the F. × ananassa physical genome has caused a proliferation of disconnected strawberry chromosome nomenclatures that may not accurately reflect the phylogenetic origins of its respective subgenomes. The Camarosa v1.0 reference genome provides an anchoring point for unifying the existing octoploid nomenclatures. We aligned all historic Fragaria microsatellite markers in the Rosacea Genomic Database (GDR) to the Camarosa v1.0 genome and anchored the Spigler et al. (2010) nomenclature to the Camarosa physical genome, which provided the corresponding linkage groups for anchoring the Tennessen
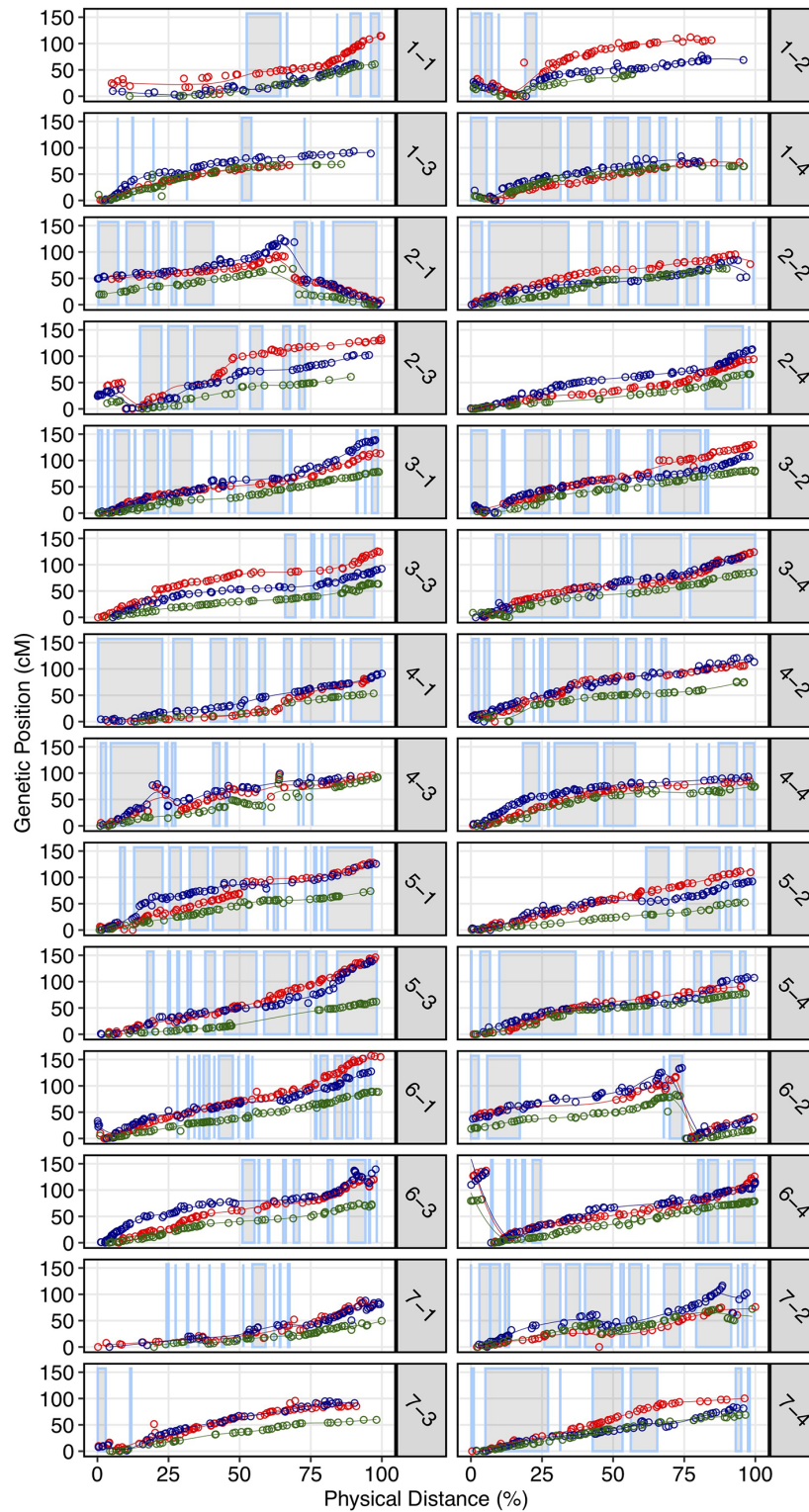
FIGURE 4 | Genetic maps of three wild octoploid strawberry genotypes (PI552277–red; PI612493–blue; Del Norte–green) based on 50K SNP array genotypes plotted against the Camarosa v1.0 physical genome. Grey highlighted chromosome segments indicate contiguous (up to 500 kb) regions of the physical genome represented by the Camarosa 50K SNP array map.
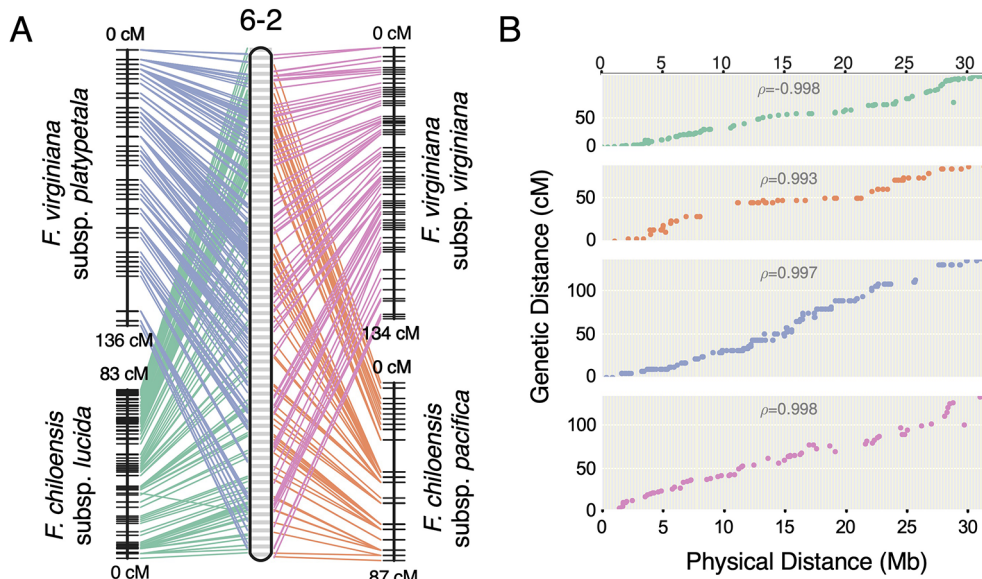
FIGURE 5 | Comparative mapping of four wild octoploid *Fragaria* subspecies (*F. chiloensis* subsp. *lucida*, *F. chiloensis* subsp. *pacifica*, *F. virginiana* subsp. *platypetala*, *F. virginiana* subsp. *virginiana*) on chromosome 6-2. **(A)** Marker collinearity between individual species maps and the consensus ordering of 50-kb physical windows from the Camarosa v1.0 physical genome. **(B)** Marker genetic distances plotted against the consensus ordering of 50-kb physical windows.
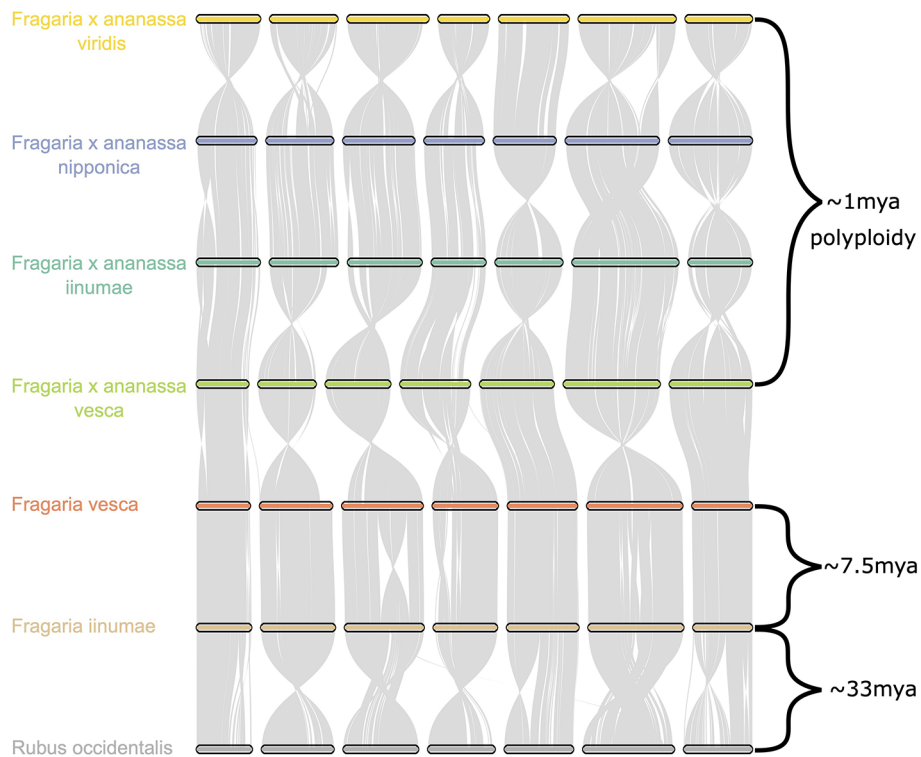

FIGURE 6 | Chromosome scale collinearity of the *F. × ananassa* subgenomes, *F. vesca*, *F. iinumae*, and *Rubus occidentalis*. Large inversions reflect the orientations of genome assemblies, not whole-chromosome inversions. Divergence estimates for *F. vesca* and *F. iinumae*, and *Fragaria and Rubus* are from Qiao et al. (2016) and estimate for date of *F. × ananassa* polyploid are from Njuguna et al. (2013).

et al. (2014) nomenclature. We then utilized legacy iStraw probes retained on the 50K array to link the Sargent and van Dijk chromosome nomenclatures (van Dijk et al., 2014; Sargent et al., 2016) to the Camarosa v1.0 genome, which was scaffolded using the map published by Davik et al. (2015). In total, five of the most widely cited octoploid strawberry chromosome nomenclatures were unified in relation to the physical genome (**Table 1**). The existing octoploid nomenclatures each contained subgenome assignments that were incongruent with ancestral chromosomal origins determined by phylogenetic analysis of the physical genome (Edger et al., 2019b; Edger et al., 2019a), particularly with respect to the non-*F. vesca* and non-*F. iinumae* subgenomes. The unmasking of octoploid homoeologous chromosome lineages (Edger et al., 2019b; Edger et al., 2019a) and construction of genetic maps showing complete collinearity among ancestral species provide the foundation for a unified octoploid nomenclature reflecting the phylogenetic origins of its subgenomes.

CONCLUSION

Using the first octoploid genome-guided approach to subgenomic (diploid) DNA variant discovery, we have demonstrated that the genomes of the octoploid progenitors of *F. × ananassa* are highly collinear and diploidized (**Figure 5**, **Figure S6**). Octoploid *Fragaria* taxa do not follow the common polyploid rule book for chromosome rearrangement (Ramsey

and Schemske, 2002; Gaeta et al., 2007; Leitch and Leitch, 2008; Cifuentes et al., 2010; Gaeta and Pires, 2010; Chester et al., 2012; Renny-Byfield and Wendel, 2014; Wendel et al., 2016; Alix et al., 2017), but instead exhibit incredible karyotypic stability across biogeographically diverse subspecies. Strikingly, we did not observe any large-scale (Mb+) structural rearrangements (e.g., translocations or inversions) in the genomes of *F. chiloensis*, *F. virginiana*, or *F. × ananassa* (**Figure 4**). The broad conservation of chromosome structure across diverse progenitor taxa partly explains the absence of reproductive barriers and ease of gene flow between wild octoploid species and the domesticated hybrid lineage. In this regard, octoploid *Fragaria* species appear part of a minority among polyploid plants, though similar examples of karyotypic stability have been described in monocots and dicots (Sun et al., 2017; VanBuren et al., 2019). Because of the ubiquity of polyploidy in angiosperms and the diversity of chromosome-restructuring outcomes along the pathway to diploidization, universal rules do not necessarily apply (Cifuentes et al., 2010; Le Comber et al., 2010; Renny-Byfield and Wendel, 2014; Wendel et al., 2016). The remarkable karyotypic stability and absence of chromosome rearrangements among octoploid *Fragaria* taxa are indicative of regular diploid meiotic behavior and suggest that homoeologous recombination has failed to disrupt the ancestral octoploid karyotype, which has been preserved over the past 0.4-2.1 million years (Tennessen et al., 2014). The unique history of strawberry as a crop lineage, including its origin as an interspecific hybrid and frequent use of interspecific hybridization throughout domestication, was

TABLE 1 | Published octoploid strawberry linkage group nomenclatures anchored to corresponding physical chromosomes in the Camarosa v1.0 reference genome.

Octoploid Chromosome	Proposed Origin (Edger et al., 2019b)	Spigler et al. (2010)	Tennessen et al. (2014)	van Dijk et al. (2014)	Davik et al. (2015)	Sargent et al. (2016)
1-2	*F. iinumae*	I.D	I-Bi	1C	1-B	1b
2-4	*F. iinumae*	II.B	II-Bi	2C	2-D	2b
3-2	*F. iinumae*	III.A	III-Bi	3D	3-B	3b
4-4	*F. iinumae*	IV.C	IV-Bi	4A	4-D	4b
5-3	*F. iinumae*	V.D	V-Bi	5B	5-C	5b
6-3	*F. iinumae*	VI.D	VI-Bi	6B	6-C	6X2
7-3	*F. iinumae*	VII.A	VII-Bi	7D	7-C	7b
1-3	*F. nipponica*	I.C	I-B2	1B	1-C	1X2
2-1	*F. nipponica*	II.D	II-B1	2D	2-A	2X2
3-3	*F. nipponica*	III.C	III-B1	3B	3-C	3X2
4-2	*F. nipponica*	IV.D	IV-B2	4D	4-B	4X2
5-4	*F. nipponica*	V.B	V-B2	5D	5-D	5X1
6-2	*F. nipponica*	VI.B	VI-B1	6D	6-B	6b
7-1	*F. nipponica*	VII.D	VII-B2	7C	7-A	7X2
1-4	*F. vesca*	I.A	I-Av	1A	1-D	1A
2-2	*F. vesca*	II.A	II-Av	2A	2-B	2A
3-4	*F. vesca*	III.D	III-Av	3A	3-D	3A
4-3	*F. vesca*	IV.A	IV-Av	4B	4-C	4A
5-1	*F. vesca*	V.C	V-Av	5A	5-A	5A
6-1	*F. vesca*	VI.A	VI-Av	6A	6-A	6A
7-2	*F. vesca*	VII.B	VII-Av	7A	7-B	7A
1-1	*F. viridis*	I.B	I-B1	1D	1-A	1X1
2-3	*F. viridis*	II.C	II-B2	2B	2-C	2X1
3-1	*F. viridis*	III.B	III-B2	3C	3-A	3X1
4-1	*F. viridis*	IV.B	IV-B1	4C	4-A	4X1
5-2	*F. viridis*	V.A	V-B1	5C	5-B	5X2
6-4	*F. viridis*	VI.C	VI-B2	6C	6-D	6X1
7-4	*F. viridis*	VII.C	VII-B1	7B	7-D	7X1

almost certainly supported by the uncommon stability of its progenitor genomes, allowing unrestricted gene flow across octoploid genetic backgrounds. Comparative genomic analysis of octoploid strawberry with related diploid species and *Rubus occidentalis* revealed this stability extends back over 30 million years. Thus, the high degree of structural conservation among genomes of diverse ancestral *Fragaria* species may factor into the absence of any chromosomal fusions and/or rearrangements arising as a result of polyploid or homoploid hybridization.

The global importance and rapid commercial success of *F. × ananassa* over the last 250 years has been attributed to the interspecific homoploid hybrid component of heterosis (Spangelo et al., 1971; Shaw, 1995; Stegmeir et al., 2010; Rho et al., 2012). Quantitative evidence for heterosis in *F. × ananassa* is limited (Spangelo et al., 1971; Shaw, 1995; Stegmeir et al., 2010; Rho et al., 2012). Heterosis is an often-cited advantage of polyploidy, where genomic heterozygosity preserved by subgenomic recombination maintains "fixed heterosis" in allopolyploids (Comai, 2005). The *F. × ananassa* genome is riddled with ancient homoeologous exchanges (Edger et al., 2019b), a hallmark of inter-subgenomic recombination in early generations. Nevertheless, the genomes of present-day octoploid taxa appear to be highly diploidized. We observed disomic inheritance of DNA variants across the genomes of the octoploids in the present study, and similar ranges of subgenomic heterozygosity for wild individuals and commercial hybrids. The success of *F. × ananassa* should not be solely attributed to "fixed heterosis" because neither octoploid progenitor species, which share the effects of fixed heterosis and show similar subgenomic heterozygosity, was commercially successful before the hybrid (Darrow, 1966; Finn et al., 2013). We hypothesize that interspecific complementation, a broader pool of potentially adaptive alleles, and masking of deleterious mutations could be more important than fixed heterosis in *F. × ananassa* (Comai, 2005; Alix et al., 2017).

We have shown that the inherent complexity and previous intractability of octoploid strawberry genomics were largely associated with the technical challenge of distinguishing subgenome level variation from the broader pool of ancestral sequence homology. The use of an allo-octoploid reference genome addressed this problem by allowing variant calling based on unique sequence alignments to the respective subgenomes. While local subgenome homology could remain an issue, we identified a nearly continuous distribution of subgenome-specific variation spanning the octoploid genome by traditional short-read sequencing. With the design of the 850K and 50K arrays, facile high-density genotyping of octoploid-anchored disomic SNP markers has been further enabled, reducing the bioinformatic requirements for octoploid strawberry research. In addition to expanding and validating the current molecular toolset, we have demonstrated that allopolyploid reference genomes facilitate the use of straightforward diploid approaches for genomic analysis and quantitative genetics of octoploid strawberry. In doing so, the results of this study help pave the way for molecular breeding of a historically difficult plant genome.

MATERIALS AND METHODS

WGS Sequence Datasets

We generated Illumina sequencing libraries for a diversity panel of 84 wild and domesticated octoploid genotypes (PE150), and the Camarosa reference genotype (PE250). Eight sequenced octoploid libraries (PE100) were obtained from the NCBI sequence read archive (SRA) (SRR1513906, SRR1513893, SRR1513905, SRR1513903, SRR1513892, SRR1513904, SRR1513867, SRR1513873), providing a total of 93 sample libraries in the diversity panel. We generated additional Illumina libraries (PE150) for 182 progeny of an *F. × ananassa* x *F. chiloensis* subsp. *lucida* mapping population (Camarosa x Del Norte). Genomic DNA was extracted from immature leaf tissue using the E-Z 96 Plant DNA kit (Omega Bio-Tek, Norcross, GA, USA) with Proteinase K was added to the initial buffer, and RNase treatment delayed until lysate was removed from the cellular debris. An additional spin was added, and incubation steps were heated to 65 C during elution. Libraries were prepared using the KAPA Hyper Plus kit using BIOO Nextflex adapters. DNA was sheared using the Covaris E220 and size selected for an average insert size of 300-nt using magnetic beads (Mag-Bind® RxnPure Plus, Omega Bio-tek). Libraries were pooled and sequenced on the Novaseq at UCSF for average 8x genome coverage in the diversity panel, and 4-8x coverage in the mapping population. DNA capture-based Illumina sequences for the *F. chiloensis* subsp. *lucida* x *F. chiloensis* subsp. *pacifica* population (GP33 x SAL3: n = 46) described in Tennessen et al. (2014), and the *F. virginiana* subsp. *platypetala* x *F. virginiana* subsp. *platypetala* population (KB3 x KB11: n = 46) described in Tennessen et al. (2018) were downloaded from the NCBI SRA.

Subgenomic WGS Variant Calling

We predicted SNP and INDEL variants in the Camarosa v1.0 subgenomes using sequenced Illumina short-read libraries for the octoploid diversity panel, Camarosa x Del Norte bi-parental population, and DNA capture sequences downloaded from the SRA. Short-read sequences were quality-trimmed with CutAdapt v1.8 using default parameters and a minimum Phred score of 25. Trimmed short-reads were aligned to the Camarosa v1.0 genome assembly (Edger et al., 2019b) using BWA-mem v0.7.16, processed to mark optical and PCR duplicates using Picard Tools v2.18, and INDEL-realigned using GATK v3.8. Genomic variants were predicted based on uniquely mapped reads (MapQ > 20) using FreeBayes v1.2 and filtered with vcflib. For analysis of subgenomic variation, a set of hard-filters was applied to remove variants with low site quality (vcflib: QUAL > 40), low contribution of allele observations to site quality (vcflib: QUAL/AO > 10), low read coverage (vcflib: DP > 500), strand bias (vcflib: SAF > 0 & SAR > 0), read-placement bias (RPR > 1 & RPL > 1), unbalanced mapping quality of reference and alternate alleles (vcflib: 0.4 ≤ [MQM/MQMR] ≤ 2.5), unbalanced allele frequencies at heterozygous sites (vcflib: 0.2 ≤ AB ≤ 0.8), low end-placement probability score (EPP ≥ 3), and low strand-bias probability score (vcflib: SRP ≥ 3 & SAP ≥ 3). Sample genotypes

were required to have individual read coverage ≥ 4, and at least two reads and a minimum of 0.20 read observations supporting each allele.

Octoploid Genomic Diversity

We estimated octoploid genetic diversity metrics from a VCF file containing genotype calls for 45M filtered subgenomic SNPs and INDELs. We calculated the population-level diversity (π) and per-sample heterozygosity of sequence variants using a custom perl script. Chromosomal and genome-wide population nucleotide diversity estimates were calculated as the sum of pairwise diversity for all variant sites divided by total non-gap (N) genomic nucleotides. Sample heterozygosity was calculated as the sum of heterozygous variant sites divided by total non-gap (N) genomic nucleotides, and the fraction of total variant sites that were heterozygous.

Array Design and Genotyping

Unfiltered genomic variants were filtered to retain sites segregating in *F.* × *ananassa* cultivars. Cultivar variants were filtered to retain biallelic SNP sites with minor allele frequency ≥ 0.05, marker diversity ≥ 0.05, variant QUAL score > 20, and missing data < 15%. Variants requiring 2-probe assays (A/T or C/G) were excluded. 71-nt marker probe sequences were obtained by retrieving 35-nt SNP flanking sequences from the Camarosa v1.0 assembly. Markers containing ambiguous sequences (Ns), or identical probes were excluded. A set of 6.6M probes was submitted to Affymetrix for scoring and recommendation based on strand, kmer uniqueness, and buildability. Probes were scored for likelihood of binding interference or non-specific binding based on off-target variant counts in the binding region, the sum of minor allele frequencies of interfering variants, and counting off-target BLAST alignments (> 90% id, > 90% query length) in the genome. A final screening panel of 850,000 markers, including 16,000 iStraw probes (Bassil et al., 2015; Verma et al., 2016), was submitted to Affymetrix for constructing the 850K screening array. A panel of 384 octoploid strawberry genotypes was screened on the 850K array. Marker genotype clusters were scored using the Axiom Analysis Suite. Clustering was performed in "polyploid" mode with a marker call-rate threshold of 0.89. Samples were filtered with a dQC threshold of 0.82 and QC CR threshold of 93. A subset of 49,483 probes was selected from polymorphic, QC-passing markers ("PolyHighResolution", "NoMinorHomozygote", "OffTargetVariant") on the 850K screening array to populate the 50K production array. 5,809 LD-pruned (r^2 < 0.50) probes were pre-selected from the iStraw design, in addition to 47 probes associated with QTL for *Fusarium oxysporum* resistance and the Wasatch day neutral flowering locus (unpublished data). We assigned two markers per gene to a set of 2,878 genes located in expression networks related to flowering and fruit development (Kang et al., 2013; Hollender et al., 2014), or associated with R-gene domains. Non-overlapping 50 kb physical windows were parsed to select single markers containing the highest pairwise diversity in *F.* × *ananassa*

genotypes. The remainder of the 50K array was populated by iteratively parsing 50 kb physical windows to select random QC-passing markers for uniform genomic distribution. 1,421 octoploid samples, including the Camarosa x Del Norte mapping population (n = 182), and PI552277 x PI612493 mapping population (n = 96), were genotyped on the 50K array and processed using the Axiom Analysis Suite using the same settings as the 850K dataset.

WGS Haplotype Linkage Mapping

We used 1.6M female parent informative variant calls (AB x AA), and 1.9M male parent informative variant calls (AA x AB) to generate haploblock markers for mapping genome-wide variant calls in the Camarosa x Del Norte bi-parental population. Camarosa-informative and Del Norte-informative variant calls were divided into parent-specific marker sets, then split by chromosome. For each chromosome, we performed pairwise linkage disequilibrium (LD)-clustering of markers (LD ≥ 0.96) in an initial seed region containing the first 250 chromosome variants, to identify marker groups called in the same phase relative to the unphased Camarosa v1.0 genome. The genotype phase of the LD cluster containing the largest number of markers were selected as the "seed phase". A 50 kb sliding window was initiated in the seed region and moved across the chromosome, identifying downstream markers in negative LD with the seed phase and reversing the repulsion genotype calls, in order to phase the chromosome into an artificial backcross. If a phasing window skipped a physical region larger than 100 kb without markers, reached a window with fewer than 25 markers, or the average downstream marker LD fell below 0.90, the chromosome was then fragmented at the breakpoint, seed-phase clustering repeated, and the sliding window reset for the subsequent downstream region. We used the software PhaseLD (Marand et al., 2017) with a 50-marker window and 1-marker step size to predict crossover events in the backcross-phased chromosome blocks. Window-specific variant calls lying between the predicted recombination breakpoints were used to generate consensus genotypes representing the haploblock. We mapped the haploblock markers using software ONEMAP (Margarido et al., 2007) and BatchMap (Schiffthaler et al., 2017) in outcross mode. ONEMAP was used to bin co-segregating markers, calculate pairwise recombination fractions, determine optimal LOD thresholds, then cluster markers into linkage groups based on a LOD threshold of 8, and maximum recombination fraction of 0.22. Marker orders and genetic distances were re-estimated in parallel with BatchMap using a window of 50 markers, window overlap of 30, and ripple window of 6 markers.

Array and DNA Capture Linkage Mapping

We performed single-marker linkage mapping of populations genotyped using the 50K array or DNA capture sequences because each contained fewer than 10,000 segregating markers per parent. Individual parent genotypes were mapped separately

using their respective informative marker subsets. We filtered markers based on a chi-square test for segregation distortion (p-value < 0.10), and excluded markers with >5% missing data. ONEMAP was used to bin co-segregating markers, calculate pairwise recombination fractions, determine optimal LOD thresholds, and cluster markers into linkage groups based on a LOD threshold of 8, and maximum recombination fraction of 0.22. Marker orders and genetic distances were re-estimated in parallel with BatchMap using a window of 20 markers, window overlap of 15, and ripple window of 6 markers.

GDR Microsatellite Alignment

We obtained the complete set of microsatellite primers developed in *Fragaria* species from the Rosaceae Genomics Database (GDR). Primers pairs were aligned to the Camarosa v1.0 genome in an orientation-aware manner using IPCRESS with a maximum amplicon fragment size of 500 bp and allowing 1 mismatch per primer.

Comparative Genomic Analysis

The subgenomes of *F. × ananassa* cultivar Camarosa and the genomes of *F. vesca* (Edger et al., 2017), *F. iinumae* (Edger et al., 2019a), and *Rubus occidentalis* (VanBuren et al., 2018) were compared with the MCScan Toolkit v1.1 (Wang et al., 2012) Syntenic gene pairs were identified using all-vs-all BLAST search, filtered to remove pairs with cscores below 0.7, and clustered into syntenic blocks in MCScan. Macrosynteny plots were constructed using the python version of MCScan (https://github.com/tanghaibao/jcvi/wiki/MCscan) (Python-version).

DATA AVAILABILITY STATEMENT

Sequence datasets used in this study were deposited in the NCBI Sequence Read Archive (https://www.ncbi.nlm.nih.gov/sra), BioProject ID PRJNA578384. Supplemental data files, genetic maps, and Affymetrix SNP genotypes were deposited in the DRYAD database (DOI https://doi.org/10.25338/B8R31Q).

AUTHOR CONTRIBUTIONS

MH, PE, and SK contributed conception and design of the study. MH, AL, KB, and MF performed the statistical analysis. RF, CA, and GC provided genetic material and collected, submitted DNA samples. MH and SK wrote the first draft of the manuscript. All authors contributed to manuscript revision, read and approved the submitted version.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2019.01789/full#supplementary-material

REFERENCES

Abbott, R. J., Hegarty, M. J., Hiscock, S. J., and Brennan, A. C. (2010). Homoploid hybrid speciation in action. *Taxon* 59, 1375–1386. doi: 10.1002/tax.595005

Alix, K., Gérard, P. R., Schwarzacher, T., and Heslop-Harrison, J. S. (2017). Polyploidy and interspecific hybridization: partners for adaptation, speciation and evolution in plants. *Ann. Bot.* 120, 183–194. doi: 10.1093/aob/mcx079

Arulsekar, S., and Bringhurst, R. S. (1981). Genetic model for the enzyme marker PGI in diploid California Fragaria vesca L: its variability and use in elucidating the mating system. *J. Hered.* 72, 117–120. doi: 10.1093/oxfordjournals.jhered.a109438

Barb, J. G., Bowers, J. E., Renaut, S., Rey, J. I., Knapp, S. J., Rieseberg, L. H., et al. (2014). Chromosomal evolution and patterns of introgression in Helianthus. *Genetics* 197, 969–979. doi: 10.1534/genetics.114.165548

Bassil, N. V., Davis, T. M., Zhang, H., Ficklin, S., Mittmann, M., Webster, T., et al. (2015). Development and preliminary evaluation of a 90 K Axiom® SNP array for the allo-octoploid cultivated strawberry Fragaria× ananassa. *BMC Genom.* 16, 155. doi: 10.1186/s12864-015-1310-1

Bird, K. A., VanBuren, R., Puzey, J. R., and Edger, P. P. (2018). The causes and consequences of subgenome dominance in hybrids and recent polyploids. *New Phytol.* 220, 87–93. doi: 10.1111/nph.15256

Bringhurst, R. S. (1990). Cytogenetics and evolution in American Fragaria. *HortScience* 25, 879–881. doi: 10.21273/HORTSCI.25.8.879

Burke, J. M., Lai, Z., Salmaso, M., Nakazato, T., Tang, S., Heesacker, A., et al. (2004). Comparative mapping and rapid karyotypic evolution in the genus Helianthus. *Genetics* 167, 449–457. doi: 10.1534/genetics.167.1.449

Byrne, D., and Jelenkovic, G. (1976). Cytological diploidization in the cultivated octoploid strawberry Fragaria× ananassa. *Can. J. Genet. Cytol.* 18, 653–659. doi: 10.1139/g76-076

Chester, M., Gallagher, J. P., Symonds, V. V., da Silva, A. V. C., Mavrodiev, E. V., Leitch, A. R., et al. (2012). Extensive chromosomal variation in a recently formed natural allopolyploid species, Tragopogon miscellus (Asteraceae). *Proc. Natl. Acad. Sci.* 109, 1176–1181. doi: 10.1073/pnas.1112041109

Cifuentes, M., Grandont, L., Moore, G., Chèvre, A. M., and Jenczewski, E. (2010). Genetic regulation of meiosis in polyploid species: new insights into an old question. *New Phytol.* 186, 29–36. doi: 10.1111/j.1469-8137.2009.03084.x

Comai, L. (2005). The advantages and disadvantages of being polyploid. *Nat. Rev. Genet.* 6, 836. doi: 10.1038/nrg1711

Dale, A., and Sjulin, T. M. (1990). Few cytoplasm contribute to North American strawberry cultivars. *HortScience* 25, 1341–1342. doi: 10.21273/HORTSCI.25.11.1341

Darrow, G. M. (1966). *The strawberry. History, breeding and physiology* (New York: Holt, Rinehart & Winston).

Davik, J., Sargent, D. J., Brurberg, M. B., Lien, S., Kent, M., and Alsheikh, M. (2015). A ddRAD based linkage map of the cultivated strawberry, Fragaria xananassa. *PLoS One* 10, e0137746. doi: 10.1371/journal.pone.0137746

de Bem Oliveira, I., Resende, M. F. R., Ferrão, L. F. V., Amadeu, R. R., Endelman, J. B., Kirst, M., et al. (2019). Genomic prediction of autotetraploids; influence of relationship matrices, allele dosage, and continuous genotyping calls in phenotype prediction. *Genes Genom. Genet.* 9, 1189–1198. doi: 10.1534/g3.119.400059

Duchesne, A.-N. (1766). *Histoire naturelle des fraisiers contenant les vues d'économie réunies à la botanique, et suivie de remarques particulières sur plusieurs points qui ont rapport à l'histoire naturelle générale, par M. Duchesne fils.* chez Didot le jeune.

Edger, P. P., VanBuren, R., Colle, M., Poorten, T. J., Wai, C. M., Niederhuth, C. E., et al. (2017). Single-molecule sequencing and optical mapping yields an improved genome of woodland strawberry (Fragaria vesca) with chromosome-scale contiguity. *Gigascience* 7, 1–7. gix124. doi: 10.1093/gigascience/gix124

Edger, P. P., McKain, M. R., Yocca, A. E., Knapp, S. J., Qiao, Q., and Zhang, T. (2019a). Phylogenetic analyses and Fragaria iinumae genome support octopoid strawberry origin. *Nat. Genet.*

Edger, P. P., Poorten, T. J., VanBuren, R., Hardigan, M. A., Colle, M., McKain, M. R., et al. (2019b). Origin and evolution of the octoploid strawberry genome. *Nat. Genet.* 51, 541–547. doi: 10.1038/s41588-019-0356-4

Fedorova, N. J. (1946). Crossability and phylogenetic relations in the main European species of Fragaria. *Compil. Natl. Acad. Sci. USSR* 52, 545–547.

Finn, C. E., Retamales, J. B., Lobos, G. A., and Hancock, J. F. (2013). The Chilean strawberry (Fragaria chiloensis): over 1000 years of domestication. *HortScience* 48, 418–421. doi: 10.21273/HORTSCI.48.4.418

Folta, K. M., and Davis, T. M. (2006). Strawberry genes and genomics. *CRC Crit. Rev. Plant Sci.* 25, 399–415. doi: 10.1080/07352680600824831

Gaeta, R. T., and Pires, J. C. (2010). Homoeologous recombination in allopolyploids: the polyploid ratchet. *New Phytol.* 186, 18–28. doi: 10.1111/j.1469-8137.2009.03089.x

Gaeta, R. T., Pires, J. C., Iniguez-Luy, F., Leon, E., and Osborn, T. C. (2007). Genomic changes in resynthesized Brassica napus and their effect on gene expression and phenotype. *Plant Cell* 19, 3403–3417. doi: 10.1105/tpc.107.054346

Govindarajulu, R., Parks, M., Tennessen, J. A., Liston, A., and Ashman, T. (2015). Comparison of nuclear, plastid, and mitochondrial phylogenies and the origin of wild octoploid strawberry species. *Am. J. Bot.* 102, 544–554. doi: 10.3732/ajb.1500026

Grover, C. E., Gallagher, J. P., Szadkowski, E. P., Yoo, M. J., Flagel, L. E., and Wendel, J. F. (2012). Homoeolog expression bias and expression level dominance in allopolyploids. *New Phytol.* 196, 966–971. doi: 10.1111/j.1469-8137.2012.04365.x

Hamilton, J. P., and Buell, C. R. (2012). Advances in plant genome sequencing. *Plant J.* 70, 177–190. doi: 10.1111/j.1365-313X.2012.04894.x

Hancock, J. F.Jr., and Bringhurst, R. S. (1979). Ecological differentiation in perennial, octoploid species of Fragaria. *Am. J. Bot.* 66, 367–375. doi: 10.1002/j.1537-2197.1979.tb06237.x

Hardigan, M. A., Laimbeer, F. P. E., Newton, L., Crisovan, E., Hamilton, J. P., Vaillancourt, B., et al. (2017). Genome diversity of tuber-bearing Solanum uncovers complex evolutionary history and targets of domestication in the cultivated potato. *Proc. Natl. Acad. Sci.* 114, E9999–E10008. doi: 10.1073/pnas.1714380114

Hardigan, M. A., Poorten, T. J., Acharya, C. B., Cole, G. S., Hummer, K. E., Bassil, N., et al. (2018). Domestication of temperate and coastal hybrids with distinct ancestral gene selection in octoploid strawberry. *Plant Genome* 11. doi: 10.3835/plantgenome2018.07.0049

Hirakawa, H., Shirasawa, K., Kosugi, S., Tashiro, K., Nakayama, S., Yamada, M., et al. (2014). Dissection of the octoploid strawberry genome by deep sequencing of the genomes of Fragaria species. *DNA Res.* 21, 169–181. doi: 10.1093/dnares/dst049

Hirsch, C. N., and Buell, C. R. (2013). Tapping the promise of genomics in species with complex, nonmodel genomes. *Annu. Rev. Plant Biol.* 64, 89–110. doi: 10.1146/annurev-arplant-050312-120237

Hollender, C. A., Kang, C., Darwish, O., Geretz, A., Matthews, B. F., Slovin, J., et al. (2014). Floral transcriptomes in woodland strawberry uncover developing receptacle and anther gene networks. *Plant Physiol.* 165, 1062–1075. doi: 10.1104/pp.114.237529

Huang, X., Feng, Q., Qian, Q., Zhao, Q., Wang, L., Wang, A., et al. (2009). High-throughput genotyping by whole-genome resequencing. *Genome Res.* 19, 1068–1076. doi: 10.1101/gr.089516.108

Hufford, M. B., Xu, X., Van Heerwaarden, J., Pyhäjärvi, T., Chia, J.-M., Cartwright, R. A., et al. (2012). Comparative population genomics of maize domestication and improvement. *Nat. Genet.* 44, 808. doi: 10.1038/ng.2309

Hughes, C. E., Govindarajulu, R., Robertson, A., Filer, D. L., Harris, S. A., and Bailey, C. D. (2007). Serendipitous backyard hybridization and the origin of crops. *Proc. Natl. Acad. Sci.* 104, 14389–14394. doi: 10.1073/pnas.0702193104

Jiao, Y., Wickett, N. J., Ayyampalayam, S., Chanderbali, A. S., Landherr, L., Ralph, P. E., et al. (2011). Ancestral polyploidy in seed plants and angiosperms. *Nature* 473, 97. doi: 10.1038/nature09916

Kang, C., Darwish, O., Geretz, A., Shahan, R., Alkharouf, N., and Liu, Z. (2013). Genome-scale transcriptomic insights into early-stage fruit development in woodland strawberry Fragaria vesca. *Plant Cell* 25, 1960–1978. doi: 10.1105/tpc.113.111732

Koskela, E. A., Sønsteby, A., Flachowsky, H., Heide, O. M., Hanke, M., Elomaa, P., et al. (2016). TERMINAL FLOWER 1 is a breeding target for a novel everbearing trait and tailored flowering responses in cultivated strawberry (Fragaria× ananassa Duch.). *Plant Biotechnol. J.* 14, 1852–1861. doi: 10.1111/pbi.12545

Kunihisa, M., Fukino, N., and Matsumoto, S. (2005). CAPS markers improved by cluster-specific amplification for identification of octoploid strawberry (Fragaria× ananassa Duch.) cultivars, and their disomic inheritance. *Theor. Appl. Genet.* 110, 1410–1418. doi: 10.1007/s00122-005-1956-1

Ladizinsky, G. (1985). Founder effect in crop-plant evolution. *Econ. Bot.* 39, 191–199. doi: 10.1007/BF02907844

Latta, R. G., Bekele, W. A., Wight, C. P., and Tinker, N. A. (2019). Comparative linkage mapping of diploid, tetraploid, and hexaploid Avena species suggests extensive chromosome rearrangement in ancestral diploids. *Sci. Rep.* 9, 1–12. doi: 10.1038/s41598-019-48639-7

Le Comber, S. C., Ainouche, M. L., Kovarik, A., and Leitch, A. R. (2010). Making a functional diploid: from polysomic to disomic inheritance. *New Phytol.* 186, 113–122. doi: 10.1111/j.1469-8137.2009.03117.x

Lee, H., and Schatz, M. C. (2012). Genomic dark matter: the reliability of short read mapping illustrated by the genome mappability score. *Bioinformatics* 28, 2097–2105. doi: 10.1093/bioinformatics/bts330

Leitch, A. R., and Leitch, I. J. (2008). Genomic plasticity and the diversity of polyploid plants. *Science* 320, 481–483. doi: 10.1126/science.1153585

Lerceteau-Köhler, E., Guerin, G., Laigret, F., and Denoyes-Rothan, B. (2003). Characterization of mixed disomic and polysomic inheritance in the octoploid strawberry (Fragaria× ananassa) using AFLP mapping. *Theor. Appl. Genet.* 107, 619–628. doi: 10.1007/s00122-003-1300-6

Liston, A., Cronn, R., and Ashman, T. (2014). Fragaria: a genus with deep historical roots and ripe for evolutionary and ecological insights. *Am. J. Bot.* 101, 1686–1699. doi: 10.3732/ajb.1400140

Luby, J. J., Hancock, J. F., and Ballington, J. R. (1992). Collection of native strawberry germplasm in the Pacific Northwest and Northern Rocky Mountains of the United States. *HortScience* 27, 12–17. doi: 10.21273/HORTSCI.27.1.12

Mandáková, T., Pouch, M., Brock, J. R., Al-Shehbaz, I. A., and Lysak, M. A. (2019). Origin and evolution of diploid and allopolyploid camelina genomes was accompanied by chromosome shattering. *Plant Cell.* 31, 2596-2612 tpc-00366. doi: 10.1105/tpc.19.00366

Marand, A. P., Jansky, S. H., Zhao, H., Leisner, C. P., Zhu, X., Zeng, Z., et al. (2017). Meiotic crossovers are associated with open chromatin and enriched with Stowaway transposons in potato. *Genome Biol.* 18, 203. doi: 10.1186/s13059-017-1326-8

Margarido, G. R. A., Souza, A. P., and Garcia, A. A. F. (2007). OneMap: software for genetic mapping in outcrossing species. *Hereditas* 144, 78–79. doi: 10.1111/j.2007.0018-0661.02000.x

McKain, M. R., Tang, H., McNeal, J. R., Ayyampalayam, S., Davis, J. I., dePamphilis, C. W., et al. (2016). A phylogenomic assessment of ancient polyploidy and genome evolution across the Poales. *Genome Biol. Evol.* 8, 1150–1164. doi: 10.1093/gbe/evw060

Meyer, R. S., and Purugganan, M. D. (2013). Evolution of crop species: genetics of domestication and diversification. *Nat. Rev. Genet.* 14, 840. doi: 10.1038/nrg3605

Miller, A. J., and Gross, B. L. (2011). From forest to field: perennial fruit crop domestication. *Am. J. Bot.* 98, 1389–1414. doi: 10.3732/ajb.1000522

Njuguna, W., Liston, A., Cronn, R., Ashman, T.-L., and Bassil, N. (2013). Insights into phylogeny, sex function and age of Fragaria based on whole chloroplast genome sequencing. *Mol. Phylogenet. Evol.* 66, 17–29. doi: 10.1016/j.ympev.2012.08.026

Obbard, D. J., Harris, S. A., and Pannell, J. R. (2006). Simple allelic-phenotype diversity and differentiation statistics for allopolyploids. *Hered. (Edinb)* 97, 296. doi: 10.1038/sj.hdy.6800862

Page, J. T., Huynh, M. D., Liechty, Z. S., Grupp, K., Stelly, D., Hulse, A. M., et al. (2013). Insights into the evolution of cotton diploids and polyploids from whole-genome re-sequencing. *Genes Genom. Genet.* 3, 1809–1818. doi: 10.1534/g3.113.007229

Parkin, I. A. P., Koh, C., Tang, H., Robinson, S. J., Kagale, S., Clarke, W. E., et al. (2014). Transcriptome and methylome profiling reveals relics of genome dominance in the mesopolyploid Brassica oleracea. *Genome Biol.* 15, R77. doi: 10.1186/gb-2014-15-6-r77

Pincot, D. D. A., Poorten, T. J., Hardigan, M. A., Harshman, J. M., Acharya, C. B., Cole, G. S., et al. (2018). Genome-wide association mapping uncovers Fw1, a dominant gene conferring resistance to Fusarium wilt in strawberry. *Genes Genom. Genet.* 8, 1817–1828. doi: 10.1534/g3.118.200129

Qiao, Q., Xue, L., Wang, Q., Sun, H., Zhong, Y., Huang, J., et al. (2016). Comparative transcriptomics of strawberries (Fragaria spp.) provides insights into evolutionary patterns. *Front. Plant Sci.* 7, 1839. doi: 10.3389/fpls.2016.01839

Ramsey, J., and Schemske, D. W. (2002). Neopolyploidy in flowering plants. *Annu. Rev. Ecol. Syst.* 33, 589–639. doi: 10.1146/annurev.ecolsys.33.010802.150437

Renny-Byfield, S., and Wendel, J. F. (2014). Doubling down on genomes: polyploidy and crop plants. *Am. J. Bot.* 101, 1711–1725. doi: 10.3732/ajb.1400119

Rho, I. R., Woo, J. G., Jeong, H. J., Jeon, H. Y., and Lee, C. (2012). Characteristics of F1 hybrids and inbred lines in octoploid strawberry (Fragaria× ananassa Duchesne). *Plant Breed.* 131, 550–554. doi: 10.1111/j.1439-0523.2012.01958.x

Rieseberg, L. H. (1997). Hybrid origins of plant species. *Annu. Rev. Ecol. Syst.* 28, 359–389. doi: 10.1146/annurev.ecolsys.28.1.359

Rousseau-Gueutin, M., Lerceteau-Köhler, E., Barrot, L., Sargent, D. J., Monfort, A., Simpson, D., et al. (2008). Comparative genetic mapping between octoploid and diploid Fragaria species reveals a high level of colinearity between their genomes and the essentially disomic behavior of the cultivated octoploid strawberry. *Genetics* 179, 2045–2060. doi: 10.1534/genetics.107.083840

Salamone, I., Govindarajulu, R., Falk, S., Parks, M., Liston, A., and Ashman, T. (2013). Bioclimatic, ecological, and phenotypic intermediacy and high genetic admixture in a natural hybrid of octoploid strawberries. *Am. J. Bot.* 100, 939–950. doi: 10.3732/ajb.1200624

Sargent, D. J., Passey, T., Šurbanovski, N., Girona, E. L., Kuchta, P., Davik, J., et al. (2012). A microsatellite linkage map for the cultivated strawberry (Fragaria× ananassa) suggests extensive regions of homozygosity in the genome that may have resulted from breeding and selection. *Theor. Appl. Genet.* 124, 1229–1240. doi: 10.1007/s00122-011-1782-6

Sargent, D. J., Yang, Y., Šurbanovski, N., Bianco, L., Buti, M., Velasco, R., et al. (2016). HaploSNP affinities and linkage map positions illuminate subgenome composition in the octoploid, cultivated strawberry (Fragaria× ananassa). *Plant Sci.* 242, 140–150. doi: 10.1016/j.plantsci.2015.07.004

Schatz, M. C., Witkowski, J., and McCombie, W. R. (2012). Current challenges in *de novo* plant genome sequencing. *Genome Biol.* 13, 243. doi: 10.1186/gb-2012-13-4-243

Schiffthaler, B., Bernhardsson, C., Ingvarsson, P. K., and Street, N. R. (2017). BatchMap: a parallel implementation of the OneMap R package for fast computation of F1 linkage maps in outcrossing species. *PLoS One* 12, e0189256. doi: 10.1371/journal.pone.0189256

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

Senanayake, Y. D. A., and Bringhurst, R. S. (1967). Origin of Fragaria polyploids. I. Cytological analysis. *Am. J. Bot.* 54, 221–228. doi: 10.1002/j.1537-2197.1967.tb06912.x

Shaw, D. V. (1995). Comparison of ancestral and current-generation inbreeding in an experimental strawberry breeding population. *Theor. Appl. Genet.* 90, 237–241. doi: 10.1007/BF00222207

Shulaev, V., Sargent, D. J., Crowhurst, R. N., Mockler, T. C., Folkerts, O., Delcher, A. L., et al. (2011). The genome of woodland strawberry (Fragaria vesca). *Nat. Genet.* 43, 109. doi: 10.1038/ng.740

Sjulin, T. M., and Dale, A. (1987). Genetic diversity of North American strawberry cultivars. *J. Hortic. Sci. Biotechnol.* 112, 375–385.

Soltis, D. E., Segovia-Salcedo, M. C., Jordon-Thaden, I., Majure, L., Miles, N. M., Mavrodiev, E. V., et al. (2014a). Are polyploids really evolutionary dead-ends (again)? A critical reappraisal of Mayrose et al. (2011). *New Phytol.* 202, 1105–1117. doi: 10.1111/nph.12756

Soltis, P. S., Liu, X., Marchant, D. B., Visger, C. J., and Soltis, D. E. (2014b). Polyploidy and novelty: gottlieb's legacy. *Philos. Trans. R. Soc. B Biol. Sci.* 369, 20130351. doi: 10.1098/rstb.2013.0351

Soltis, D. E., Visger, C. J., Marchant, D. B., and Soltis, P. S. (2016). Polyploidy: pitfalls and paths to a paradigm. *Am. J. Bot.* 103, 1146–1166. doi: 10.3732/ajb.1500501

Spangelo, L. P. S., Hsu, C. S., Fejer, S. O., and Watkins, R. (1971). Inbred line× tester analysis and the potential of inbreeding in strawberry breeding. *Can. J. Genet. Cytol.* 13, 460–469. doi: 10.1139/g71-070

Spigler, R. B., Lewers, K. S., Main, D. S., and Ashman, T. L. (2008). Genetic mapping of sex determination in a wild strawberry, Fragaria virginiana, reveals earliest form of sex chromosome. *Hered. (Edinb)* 101, 507. doi: 10.1038/hdy.2008.100

Spigler, R. B., Lewers, K. S., Johnson, A. L., and Ashman, T.-L. (2010). Comparative mapping reveals autosomal origin of sex chromosome in octoploid Fragaria virginiana. *J. Hered.* 101, S107–S117. doi: 10.1093/jhered/esq001

Staudt, G. (1999). *Systematics and geographic distribution of the American strawberry species: Taxonomic studies in the genus Fragaria (Rosaceae: Potentilleae)* (Berkeley, CA, USA: Univ of California Press).

Stegmeir, T. L., Finn, C. E., Warner, R. M., and Hancock, J. F. (2010). Performance of an elite strawberry population derived from wild germplasm of Fragaria chiloensis, and F. virginiana. *HortScience* 45, 1140–1145. doi: 10.21273/HORTSCI.45.8.1140

Sun, H., Wu, S., Zhang, G., Jiao, C., Guo, S., Ren, Y., et al. (2017). Karyotype stability and unbiased fractionation in the paleo-allotetraploid Cucurbita genomes. *Mol. Plant* 10, 1293–1306. doi: 10.1016/j.molp.2017.09.003

Tennessen, J. A., Govindarajulu, R., Ashman, T.-L., and Liston, A. (2014). Evolutionary origins and dynamics of octoploid strawberry subgenomes revealed by dense targeted capture linkage maps. *Genome Biol. Evol.* 6, 3295–3313. doi: 10.1093/gbe/evu261

Tennessen, J. A., Wei, N., Straub, S. C. K., Govindarajulu, R., Liston, A., and Ashman, T.-L. (2018). Repeated translocation of a gene cassette drives sex-chromosome turnover in strawberries. *PLoS Biol.* 16, e2006062. doi: 10.1371/journal.pbio.2006062

Treangen, T. J., and Salzberg, S. L. (2012). Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nat. Rev. Genet.* 13, 36. doi: 10.1038/nrg3117

Vallejo-Marín, M., Buggs, R. J. A., Cooley, A. M., and Puzey, J. R. (2015). Speciation by genome duplication: repeated origins and genomic composition of the recently formed allopolyploid species Mimulus peregrinus. *Evol. (N. Y.)* 69, 1487–1500. doi: 10.1111/evo.12678

van Dijk, T., Pagliarani, G., Pikunova, A., Noordijk, Y., Yilmaz-Temel, H., Meulenbroek, B., et al. (2014). Genomic rearrangements and signatures of breeding in the allo-octoploid strawberry as revealed through an allele dose based SSR linkage map. *BMC Plant Biol.* 14, 55. doi: 10.1186/1471-2229-14-55

VanBuren, R., Wai, C. M., Colle, M., Wang, J., Sullivan, S., Bushakra, J. M., et al. (2018). A near complete, chromosome-scale assembly of the black raspberry (Rubus occidentalis) genome. *Gigascience* 7, giy094. doi: 10.1093/gigascience/giy094

VanBuren, R., Wai, C. M., Pardo, J., Yocca, A. E., Wang, X., Wang, H., et al. (2019). Exceptional subgenome stability and functional divergence in allotetraploid teff, the primary cereal crop in Ethiopia. *bioRxiv* 580720. doi: 10.1101/580720

Verma, S., Bassil, N. V., van de Weg, E., Harrison, R. J., Monfort, A., Hidalgo, J. M., et al. (2016). "Development and evaluation of the Axiom® IStraw35 384HT array for the allo-octoploid cultivated strawberry Fragaria× ananassa," in *VIII International Strawberry Symposium* (Belgium: ISHS Acta Horticulturae) 1156 75–82. doi: 10.17660/ActaHortic.2017.1156.10

Vining, K. J., Salinas, N., Tennessen, J. A., Zurn, J. D., Sargent, D. J., Hancock, J., et al. (2017). Genotyping-by-sequencing enables linkage mapping in three octoploid cultivated strawberry families. *PeerJ* 5, e3731. doi: 10.7717/peerj.3731

Wang, Y., Tang, H., DeBarry, J. D., Tan, X., Li, J., Wang, X., et al. (2012). MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 40, e49–e49. doi: 10.1093/nar/gkr1293

Wendel, J. F., Jackson, S. A., Meyers, B. C., and Wing, R. A. (2016). Evolution of plant genome architecture. *Genome Biol.* 17, 37. doi: 10.1186/s13059-016-0908-1

Yakimowski, S. B., and Rieseberg, L. H. (2014). The role of homoploid hybridization in evolution: a century of studies synthesizing genetics and ecology. *Am. J. Bot.* 101, 1247–1258. doi: 10.3732/ajb.1400201