BIOLCLIMATIC, ECOLOGICAL, AND PHENOTYPIC INTERMEDIACY
AND HIGH GENETIC ADMIXTURE IN A NATURAL HYBRID
OF OCTOPOLOID STRAWBERRIES

ISABELLA SALAMONE2, RAJANIKANTH GOVINDARAJULU2, STACEY FALK3, MATTHEW PARKS4, AARON LISTON4, AND TIA-LYNN ASHMAN2,5

2Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA; 3South Park High School, Pittsburgh, Pennsylvania 15129, USA; and 4Department of Botany and Plant Pathology, Oregon State University, Corvallis, Oregon 97331, USA

Premise of the Study: Hybrid zones provide “natural laboratories” for understanding the processes of selection, reinforcement, and speciation. We sought to gain insight into the degree of introgression and the extent of ecological–phenotypic intermediacy in the natural hybrid strawberry, Fragaria × ananassa subsp. cuneifolia.

Methods: We used whole-plastome sequencing to identify parental species-specific (Fragaria chiloensis and F. virginiana) chloroplast single-nucleotide polymorphisms and combined the use of these with nuclear microsatellite markers to genetically characterize the hybrid zone. We assessed the potential role of selection in the observed geographic patterns by bioclimatically characterizing the niche of the hybrid populations and phenotypically characterizing hybrid individuals of known genomic constitution.

Key Results: Significant admixture and little overall maternal bias in chloroplast or nuclear genomes suggest a high degree of interfertility among the parental and hybrid species and point to a long history of backcrossing and genetic mixing in the hybrid zone. Even though hybrids were phenotypically intermediate to the parental species, there was a discernible fingerprint of the parental genotype within hybrid individuals. Thus, although the pattern of introgression observed suggests geographic limitations to gene flow, it may be reinforced by selection for specific parental traits in the bioclimatically intermediate habitat occupied by the hybrid.

Conclusions: This work uncovered the genetic complexity underlying the hybrid zone of the wild relatives of the cultivated strawberry. It lays the foundation for experimental dissection of the causes of genomic introgression and nuclear–cytoplasmic disassociation, and for understanding other parts of Fragaria evolutionary history.

Key words: hybrid zone; introgression; microsatellites; morphology; population genetic structure.

Hybrid zones (Barton and Hewitt, 1985), regions where hybrids and parental species coexist, provide an opportunity to study the ecological forces that influence adaptation and speciation (Buggs, 2007). Although much progress has been made, there is much still to be understood about how patterns of introgression and selection contribute to these processes. For example, introgression in admixed populations can be a direct measure of reproductive isolation (Gompert et al., 2012), and asymmetric introgression can be important in determining the direction of evolutionary change (Bacilieri et al., 1996). Yet we do not know whether asymmetric introgression under these circumstances is a sign of competitive displacement or of higher-frequency backcrossing with one parent (Buggs, 2007). Moreover, the presence of a stable hybrid zone can reflect superiority of hybrid fitness in an ecological (or bioclimatic) niche intermediate to that occupied by either parental taxon (Buggs, 2007).

The first step to using these “natural laboratories” for understanding the processes of selection, reinforcement, and speciation is to characterize the degree of introgression and the extent of ecological–phenotypic intermediacy in hybrid populations (Barton and Hewitt, 1985; Rieseberg and Blackman, 2010; Taylor et al., 2012).

Geographic patterns in genomic markers can provide a signature of past hybrid-zone dynamics. Given the differences in inheritance, markers from cytoplasmic genomes (mitochondrial and chloroplast markers) can be compared with nuclear markers to reveal patterns of admixture and whether they derive from maternal or biparental gene flow. Such studies have revealed patterns of extensive mixing among hybrids but isolation from the parents (Zeng et al., 2011) or of little admixture and hybrids associated strongly with one parental taxon (Cruzan and Arnold, 1999; Aboim et al., 2010). Hybrids of species with separate or partially separate sexes (i.e., dioecy and subdioecy) can provide additional insight into the genesis of asymmetry in hybrid zones. Specifically, such hybrids can answer how mating system (selfer vs. outcrosser) or sexual system (separate sexes vs. combined sexes) affects patterns of introgression (e.g., Engel et al., 2005; Wallace et al., 2011). The results of these analyses can be paired with geographic and bioclimatic information to create hypotheses regarding the formation, nature, and stability of the hybrid zone (Buggs, 2007; Wu and Campbell,
2007; Peñaloza-Ramírez et al., 2010; Wallace et al., 2011) and hypotheses can be tested with data on the phenotypic or functional intermediary of hybrids or the fitness of reciprocal transplants (e.g., Wu and Campbell, 2007).

*Fragaria* × *ananassa* subsp. *cuneifolia* is the natural hybrid between two sister species of octoploid *Fragaria* (Rosaceae), *F. virginiana* subsp. *platypetala* and *F. chiloensis* (Hancock and Bringhurst, 1979; Staudt, 1999). The parental species are differentiated by morphology (Staudt, 1999), ecological niche (Hancock and Bringhurst, 1979), sexual system, and, possibly, the particularities of their sex-determining chromosomes (Goldberg et al., 2010; Spigler et al., 2010; Govindarajulu et al., 2012). For instance, *F. virginiana* subsp. *platypetala* inhabits forest margins and woodlands, whereas *F. chiloensis* (sensus latu; Catling and Porebski, 1998a; Hancock et al., 2004) prefers sandy beaches and dunes (Staudt, 1999). Hancock and Bringhurst (1979) suggested that the morphological differences between these species played a major adaptive role in their ecogeographic distribution, especially in regard to water retention and use. It has been suggested that the hybrid might display morphological intermediacy (Catling and Porebski, 1998b; Staudt, 1999) and reside in an intermediate habitat, yet this has not been broadly tested, nor has there been a bioclimatic examination of the hybrid zone been undertaken. Moreover, although a previous study with limited sampling revealed that either parent can serve as maternal donor of *F. × ananassa* subsp. *cuneifolia* (Govindarajulu et al., 2012), the geographic or climatic pattern of introgression throughout the hybrid zone in terms of the nuclear and chloroplast genetic composition is unknown, so questions of the symmetry or parental dominance in the hybrid remain unanswered.

Accordingly, the goal of the present study was to characterize the hybrid zone of *F. × ananassa* subsp. *cuneifolia* genetically, bioclimatically, and ecologically. Specifically, we asked the following questions: (1) Is there a maternal species bias or geographic pattern to the nuclear and plastidomic genomes of *F. × ananassa* subsp. *cuneifolia* or do they show wide admixture? (2) What are the biogeographic characteristics of the *F. × ananassa* subsp. *cuneifolia* hybrid zone? (3) Are the hybrids ecologically or phenotypically intermediate? And finally, (4) is there an association between hybrid genotype and habitat or phenotype?

**MATERIALS AND METHODS**

*Species biology*—All three octoploid species of *Fragaria* (Rosaceae) (*F. × ananassa* subsp. *cuneifolia*, *F. virginiana* subsp. *platypetala*, and *F. chiloensis*) are allopolyploids (AAA′A′BBB′B′; Bringhurst, 1990) with disomic inheritance (2n = 8x = 56) (Ashley et al., 2003; Govindarajulu et al., 2012). They are all herbaceous perennial herbs that can reproduce by plantlets on stolons or by seeds on swollen receptacles (Staudt, 1999). *Fragaria chiloensis* is dioecious (males and females), whereas *F. virginiana* subsp. *platypetala* and *F. × ananassa* subsp. *cuneifolia* are subdioecious (females, males, and hermaphrodites) (Ashman, 1999; Staudt, 1999; Govindarajulu et al., 2012). Hermaphrodites are self-compatible, pollination occurs via a variety of small bees, flies, and ants (Ashman, 2000), and fruits are animal dispersed (T.-L. Ashman, personal observation).

*Fragaria virginiana* has a wide range across North America and may have been separated into two vicarious groups as the result of uplifting of the Rocky Mountains (Staudt, 1999). The primary focus of our study, the western subspecies *Fragaria virginiana* subsp. *platypetala* ranges from southern California, USA, to southern British Columbia, Canada, reaching from Colorado, USA, to the Pacific coastline; and *F. chiloensis* extends along the Pacific coast from southeastern California to the farthest west tip of Alaska, USA (Hancock and Bringhurst, 1979; Staudt, 1999). Although they share the same basic vegetative body plan, the leaf morphology of the two species reflects their habitat differences (e.g., the leaves of *F. virginiana* are thin and smooth, generally elongated and narrow, with pointed teeth, whereas those of *F. chiloensis* are short, thick, and coriaceous in texture, with crenate teeth; Staudt, 1999). Moreover, although *F. cascadenensis* is a recently described decaploid species that is morphologically similar to *F. virginiana* subsp. *platypetala* (Hummer, 2012), which makes the two species difficult to distinguish in the absence of ploidy information, none of our samples originated from within the range of *F. cascadenensis* (above 1000 m in the Oregon Cascades, USA; Hummer, 2012).

**Plastome sequencing and SNP validation**—To identify single-nucleotide polymorphisms (SNPs) specific to the parental species’ plastomes, we sequenced and assembled chloroplast genomes from two accessions: a female *F. chiloensis* (accession GP33 [HM11], Goldberg et al., 2010) and a male *F. virginiana* subsp. *virginiana* (accession e4772; Spigler et al., 2010). Total genomic DNA was extracted in triplicate from fresh frozen leaves using the standard FastDNA extraction protocol (MP Biomedicals, Solon, Ohio, USA), purified with a Qiagen PCR purification kit (Qiagen, Valencia, California, USA), and quantified using the Qubit dsDNA HS assay (Life Technologies/Invitrogen, Carlsbad, California, USA). DNA (–60 ng) from each accession was “tagmented” and amplified by polymerase chain reaction (PCR) in duplicate, with 9 and 12 cycles, according to the Nextera protocols (Epigenome Technologies, Madison, Wisconsin, USA). Enriched samples were electrophoresed on a 2% agarose gel and extracted in six narrow (~30 base pairs [bp]) bands between about 200 and 340 bp using Qiagen Gel Extraction kits. Extraction yields were quantified with the Qubit dsDNA HS assay, and sizes were verified using an Agilent Bioanalyzer (Agilent Technologies, Santa Clara, California, USA). For each accession, three cleaned bands with the highest yields were combined for sequencing. Resulting insert sizes for the libraries were approximately 230–290 bp, and these were submitted to the Oregon State University Center for Genomic Research and Biocomputing for 120-bp paired-end (PE) sequencing on the IlluminaGAIIx, and loaded at 10 PM. The resulting read sequence pools were processed using Illumina’s Real Time Analysis version 1.8 and CASAVA version 1.7. Paired-end read pools for each accession were combined and searched for chloroplast homology using a perl script (Knaus, 2012) incorporating BLAT (Kent, 2002). From the subsets of reads with chloroplast homology, random draws of 100 000 and 200 000 reads were taken in triplicate using a perl script (B. Knaus and L. Wilhelm, U.S. Department of Agriculture Forest Service, unpublished) implementing the Fisher-Yates Shuffle algorithm (Durstenfeld, 1964). The short-read assembler YASSA (Ratan, 2009), as implemented in Align-reads (Straub et al., 2011), was used for plastome assembly with the *F. vesca* chloroplast genome as a reference (GenBank NC_015206.1, with a single copy of the IR region). Read pools of 100 000 read size were chosen to target plastome assemblies at 50–100× coverage depth (Straub et al., 2011); read pools of 200 000 read size were selected to minimize areas of the assemblies affected by stochastic dips in coverage. A cutoff of <5× coverage depth was used for masking positions conserved in relation to the reference. SNPs were masked unless coverage depth ≥25×. The resulting assemblies were aligned using default settings in MAFFT version 6.520 (Katoh et al., 2005) and adjusted by hand in BioEdit version 7.0.5.3 (Hall, 1999). A final consensus plastome sequence was created for each accession from its six aligned assemblies, using majority rule where discrepancies between assemblies existed. Positions were masked where no majority assembly was present.

To validate the species-specificity of the chloroplast SNPs identified by whole-plastome assembly, we sampled from 10 *F. virginiana* and 5 *F. chiloensis* from locations throughout their ranges (Appendix I[A] and Appendix S1) and cytotyped them as described below.

**Population sampling**—Wide sampling—To widely assess the distribution of maternal donors of *F. × ananassa* subsp. *cuneifolia*, we requested a loan of all herbarium specimens (or from populations identified by) the National Clonal Germplasm Repository (NCGR: http://www.ars.usda.gov/main/site_main.htm?modecode=53-58-15-00), subject to chloroplast SNP analysis (see below).
Deep sampling—Live plants were collected at ≥2 m intervals along transects in 13 populations across the ecogeographic range of *F. virginiana* subs. platypetala, *F. chiloensis*, and *F. × ananassa* subs. cuneifolia (between 2008 and 2011 [Appendix S3]). In addition to populations identified by Staude (1999) as *F. × ananassa* subs. cuneifolia (Fort Stevens [FS]; Wren [WRE]; Mary’s Peak [MP]), we collected plants from populations that were identified as *F. virginiana* by the Germplasm Resources Information Network (GRIN: http://www.ars-grin.gov) but phenotypically resembled *F. × ananassa* subs. cuneifolia (Fisherman’s Bend [FB]; Silver Falls [SF]) as well as novel sites in habitats identified as likely locations of *F. × ananassa* subs. cuneifolia (Dungeness Campground [DC], Hammond Trail [HAM]) or *F. virginiana* subs. platypetala (Emigrant Springs [ES]). Other populations of *F. virginiana* subs. platypetala (George Hudson Biological Reserve [GH], Kamiut Butte [KB]) were identified by L. Hufford (personal communication). We located sites of *F. chiloensis* using the NCGR (Honeyman State Park [HM]; http://www.ars.fas.usda.gov/main/site_main.htm; modecode=53-58-15-00), previous publications (Salishan Road [SAL]; Goldberg et al., 2010), and Calflora (Eureka [EUR]; http://www.calflora.org).

Ten to 13 plants (genotypes) per population (Appendix 1[C] and Appendix S3) were grown in a 2.1 m of Fafard no. 4 (Conrad Fafard, Agawam, Massachusetts, USA) and sand in 280 mL pots in the greenhouse at the University of Pittsburgh. In August 2011, two clones from each accession from HAM, where clones were not available), were planted in 280-mL pots of the same mix. All plants were arranged randomly on benches in January 2012 and exposed to 21°C–16°C day–night conditions with supplemental lighting for three months. The plants received fertilizer and water and were treated for pests as needed.

Genotyping, phenotyping, and environmental characterization—DNA extraction—DNA was extracted from 70–100 mg of fresh tissue from 186 plants from the 13 populations (Appendix S3) following Doyle and Doyle (1990), modified to accommodate silica-based spin columns (Epoch Life Sciences, Sugar Land, Texas, USA). One leaflet (15–60 mg dry weight) was collected per plant for 17 herbarium specimens, and DNA extraction was performed following Jobes et al. (1995).

Cytotyping—We validated two of the SNPs (petD and ndhF) identified via whole-plastome sequencing (see Results) that differentiated the chloroplast genotypes of *F. virginiana* and *F. chiloensis* in a range-wide panel of 15 samples (Appendix S1) using Sanger sequencing or dCAPS (Neff et al., 1998). We then used these SNPs to determine the maternal donor for the wide and deep sampling of *F. × ananassa* subs. cuneifolia and a set of deeply sampled populations of the parental species (Appendix S3) using either direct sequencing or dCAPS. For all plants in the deep sampling population, either ndhF or both ndhF and petD were amplified, whereas for the wide sampling both ndhF and petD were amplified by PCR following Govindaraju et al. (2012), with a modified reaction volume of 25 μL consisting of 10X Standard Taq Reaction Buffer (New England Biolabs, Ipswich, Massachusetts, USA), 10 mM dNTP, 0.5 μM each forward and reverse primer, 1 μL 100X purified BSA (New England Biolabs), 1.5 units Taq DNA Polymerase (New England Biolabs), and 1.5 μL DNA. For nine herbarium samples, we used PuReTaq Ready-To-Go PCR beads (GE Healthcare, Pittsburgh, Pennsylvania, USA).

To identify SNPs in the wide samples, the amplified PCR product was Sanger sequenced and aligned using Sequecher version 4.8 (Gene Codes, Ann Arbor, Michigan, USA). For the deep samples, the amplified PCR product was purified using a Qagen PCR purification kit, and we performed dCAPS following Govindaraju et al. (2012) to elucidate the SNP genotypes. For this assay 6 μL of purified ndhF product from each sample was digested with 5 units of the restriction enzyme MspI for 2 h at 37°C, and 6 μL of purified petD product with 5 units of TaqI at 65°C for 2 h. The recognition site of MspI includes the bolded variable site (CATTTAGAAT/CTATTGAACT) within ndhF, and that of TaqI includes the bolded variable site (TCGACTCAAA) in petD. Products were assayed on agarose gels and species-specific cytotypes identified as follows: (1) ndhF loci, a 500-bp (uncut) product, identified as *F. virginiana*, whereas two fragments (336 and 166 bp product) corresponded to the *F. chiloensis* cytotype: (2) two petD fragments (56 and 44 bp) differentiated the *F. chiloensis* cytotype from the *F. virginiana* (uncut) cytotype. In sum, both genes were scored in 134 of the 212 total samples, verifying parental assignments.

Nuclear marker genotyping—A total of five nuclear microsatellite primer-pairs (mSSRs) (ARSFL04, ARSFL014, ARSFL022, ARSFL007 [Lewers et al., 2005], and CFVCT017 [Sargent et al., 2009]) were genotyped for 12–15 individuals per population (Appendix S3). The PCR amplification with each primer pair was performed separately, according to the protocol as previously described (Schuelke, 2000; Spigler et al., 2008; Govindaraju et al., 2012) on a PTC-225 thermal cycler (MJ Research, GMI, Ramsey, Minnesota, USA) in 15 μL volume. Four positive control samples were included on each 96-well plate. We multiplexed PCR products from two or three primers by mixing 1.5-μL aliquots from each reaction with 0.2 μL LIZ500 standard and 10.5 μL Hi-Di formamide (Applied Biosystems, Life Technologies, Carlsbad, California, USA). Fragment analysis and genotyping were conducted using ABI 3730XL DNA analyzer and GeneMapper (Applied Biosystems). Scoring products of SSR primer pairs in polyplodies can be challenging because of high levels of heterozygosity and difficulty in assessing which products occur in more than one copy (Arroyo et al., 2010). Thus, we adopted a commonly used approach of treating products as dominant markers and scoring presence or absence of a given product (“allel” size) (Arroyo et al., 2010). In total, 81 products (ARSFL04 [10] ARSFL014 [19], ARSFL022 [15], ARSFL007 [18], and CFVCT017 [19]) were scored on 185 individuals from the 13 populations.

Morphological characteristics—We scored six morphological characters that have been identified as differentiating the two sister species (coriaceousness, vein reticulation, leaflet length, leaflet width, orientation of trichomes on the petiole, and thickness of the petiole and leaf) and mapped the coordinates from this data set to visualize the distribution pattern of each species. We obtained records from the Global Biodiversity Information Facility (GBIF; http://www.gbif.org) and augmented GBIF records with those of our sampled populations and herbarium specimens. In instances where GPS coordinates were not recorded on herbarium sheets, coordinates were extrapolated from Google Earth (http://google.com/earth/index.html) when city and state (or province) were indicated. We assumed that records in GBIF were properly identified, but we acknowledge that this may not be true in some cases because of difficulties in distinguishing between the species. We obtained 571 records (300 *F. virginiana* subs. platypetala; 209 *F. chiloensis; and 62 *F. × ananassa* subs. cuneifolia) and obtained bioclimatic variables from WorldClim data (2.5 × 2.5 min resolution; http://www.worldclim.org/bioclim) using DIVA-GIS (DIVA-GIS version 7.5; http://www.diva-gis.org). We selected altitude, annual mean temperature, and annual precipitation for analysis. We mapped the coordinates from this data set to visualize the distribution pattern of the hybrid and parental species.

Statistical analyses—Morphometric analysis—We used principal component analysis (PCA) to reduce leaf trait information to few axes, selected axes, and used these to test for intermediacy in leaf morphology of the hybrid species, as well as to determine whether maternal cytotype affects the phenotypic intermediacy or affinity with the parental cytoplasmic donor. The PCA analysis was performed with both qualitative and quantitative data (as in Schmickl and Koch, 2011) using MVSP version 3.131 (Kovach Computing Services, UK).
We tested for differences among species in PC1 and PC2 with one-way analysis of variance (PROC GLM, SAS version 9.1; SAS Institute, Cary, North Carolina, USA) and for a cytotype effect within F. × ananassa subsp. cuneifolia with a two-sample t-test with unequal variances (Welch’s t-test with Satterthwaite’s approximate degrees of freedom) using Analyst in SAS. In addition, we used canonical discriminant function analysis (PROC CANDISC and PROC DISCRIM in SAS) to phenotypically differentiate hybrids (populations and individuals) from the parental species. Raw data used for morphometric analysis were deposited at Dryad (http://dx.doi.org/10.5061/dryad.84t05).

Cytotypic and nuclear marker analysis—To assess maternal-species bias and geographic patterns in chloroplast donor of F. × ananassa subsp. cuneifolia, we performed chi-square analyses on all widely sampled populations plus the majority type from deeply sampled ones. To infer population admixture based on nSSR markers, we performed a Bayesian model-based clustering analysis in Structure version 2.3.3 (Pritchard et al., 2000) with the 178 individuals that had marker data for three to five SSR loci. This approach assigns individuals to a population or to two or more populations (assuming K populations in a model) if their genotypes indicate that they are admixed. We signified individuals to a population or to two or more populations (assuming K populations in a model) if their genotypes indicate that they are admixed. We used the admixture ancestry model with independent allele frequencies and predefined population information for a range of K values starting from 1 to 13. We used 10,000 burn-ins and MCMC replicates for each run, with 10 replicate runs conducted for each K value. Finally, we used the online tool Structure Harvester that implements the method from Evanno et al. (2005) to extract a ΔK value that indicates the optimal number of genetic clusters from data. Data used for structure analysis were deposited at Dryad (http://dx.doi.org/10.5061/dryad.84t05).

Geographic and bioclimatic variables—We determined whether geographic or bioclimatic variables varied among the species using one-way analyses of variance (ANOVA) and Tukey’s tests to detect significant differences between species (PROC GLM in SAS). Altitude and annual mean temperature were natural log transformed, and mean annual precipitation was square-root transformed to fit model assumptions of normality. We then tested for a significant cytotype effect within F. × ananassa subsp. cuneifolia populations using a two-sample t-test (Analyst in SAS). Annual mean temperature was not normally distributed, so we tested for a cytotype effect using a Mann-Whitney-Wilcoxon two-sample test (PROC NPAR1WAY in SAS). Across all five environmental characteristics measured in situ, we tested whether F. × ananassa subsp. cuneifolia sites were intermediate between the parental species using a binomial probability test (i.e., intermediate or not).

RESULTS

Plastome sequencing and SNP validation—Plastome assembly replicates were consistent within and between species. In brief, across both species, mean (± SE) coverage depth for 100,000 and 200,000 read assemblies was 81.42 ± 0.07 and 163.09 ± 0.13, and the number of contigs in assemblies was slightly higher for 100,000-bp assemblies than for 200,000-bp assemblies (13.83 ± 1.76 vs. 10.83 ± 1.01). Plastome length for F. chiloensis (GenBank JN884816) and F. virginiana (GenBank JN884817) assemblies averaged 129,492 ± 66 bp and 129,324 ± 259 bp, respectively, with 7089 ± 43 and 1735 ± 1246 positions masked. Final consensus sequences were 129,659 bp and 129,676 bp for F. chiloensis and F. virginiana, respectively, with 134 and 136 positions masked. With inclusion of both inverted repeat copies, plastome lengths were 155,604 bp and 155,621 bp for F. chiloensis and F. virginiana, respectively, again with 134 and 136 positions masked.

A total of 60 SNPs were identified within the aligned final consensus sequences of F. chiloensis and F. virginiana plastomes (Fig. 1; Appendix S4). The majority (42/60) of polymorphisms were located in intergenic regions, whereas smaller proportions were found in protein-coding (11/60) and intronic (7/60) regions. Twenty-five polymorphisms were classified as substitutions or indels directly associated with mono- or dinucleotide repeats, 26 were substitutions that did not affect repeats, and 10 were associated with complex repetitive units or palindromic repeats.

Regardless of geographic location or subspecies designation, all F. virginiana shared the same SNPs in petD (G) and ndhF (A) (Appendix S1), and these differed from all F. chiloensis, which shared the same SNPs (petD: A; ndhF: G).

Nuclear and cytoplasmic genome structure—In total, 212 samples were cytotyped from the three species (Fig. 2A, B). All 31 samples of F. chiloensis were found to have the F. chiloensis cytotype, and all 32 samples of F. virginiana subsp. platypetala were found to have the F. virginiana cytotype. We found no significant bias in maternal species cytotype across all populations of F. × ananassa subsp. cuneifolia, 42% had F. virginiana major cytotype, and 58% had F. chiloensis major cytotype (χ² = 0.61, df = 1, P > 0.40, n = 26; Fig. 2A). To characterize the longitudinal distribution of cytotypes, we divided the hybrid zone into eastern (toward inland North America) and western (toward the Pacific Ocean) regions using the mean longitude (123.449°W) of the populations. We found that half as many of the populations in the eastern region had the F. chiloensis cytotype (3/14) as in the western region (8/12).

Within the deeply sampled populations of F. × ananassa subsp. cuneifolia, we found variation in cytotype composition, with the percentage of the F. chiloensis cytotype ranging from 6% to 100% (Fig. 2B; Appendix S5). In the eastern region, one F. × ananassa subsp. cuneifolia population had a majority of the F. chiloensis cytotype (MP, 94%) and three had a majority of F. virginiana cytotype (SF, 92%; FB, 92%; WREN, 94%). By contrast, in the western region all three sites had a majority of the F. chiloensis cytotype (DC, 76%; FTs, 100%; HAM, 100%), again suggesting a geographic bias to maternity.

Our Bayesian analyses of population structure (Fig. 2C) revealed a maximum ΔK value for two clusters, indicating two distinct nuclear genetic groups. Populations of F. chiloensis (EUR, HM, SAL) showed a high proportion of membership to cluster 1 (74 ± 14%), whereas those of F. virginiana subsp. platypetala (ES, GH, KB) to cluster 2 (69 ± 10%), although neither of the parental species displayed exclusive membership to either cluster. On average, populations of F. × ananassa subsp. cuneifolia reflected nearly equivalent membership to each cluster (e.g., cluster 1: 53 ± 12%), but membership varied widely (between 8% and 90%) among populations. For instance, MP (92%), WREN (75%), and FB (72%) showed a high proportion of membership to cluster 2, whereas FTs (90%), HAM (67%), SF (81%), and DC (71%) showed higher values of membership to cluster 1. There was no consistent association between the chloroplast and nuclear genomic affiliations. For instance, MP had a majority of F. chiloensis cytotype but showed a high proportion of membership to cluster 2, and SF had a majority of F. virginiana cytotype but displayed a higher proportion of individuals with membership to cluster 1. Within populations, individuals with the minority cytotype tended to be representative of their population’s nuclear constitution rather than that predicted by their cytotype; for example, the individual with F. chiloensis cytotype from FB showed 63% membership to cluster 2. These results suggest a high degree of introgression among both the hybrid and parental populations.
Phenotypic intermediacy—The PCA on morphological traits identified two main principal components. The first (PC1) accounted for 77% of the observed variation. Leaf coriaceousness and venation were the main contributors to PC1, and both were positively correlated with it (r = 0.61 and 0.67, respectively). The second (PC2) accounted for 13% of the variation. Petiole pubescence was positively associated with PC2, but leaf specific mass was negatively associated with it (r = 0.68 and −0.66, respectively). One-way ANOVA showed a significant species effect for PC1 and PC2 (F2, 127 = 132.88; F2, 127 = 122.98, respectively; P < 0.0001 for both; Fig. 3A, B). Post hoc comparisons showed that all three species differed significantly for PC1: F. chiloensis had highest values, followed by F. ×ananassa subsp. cuneifolia and then F. virginiana subsp. platypetala (Fig. 3A). For PC2, only F. ×ananassa subsp. cuneifolia differed significantly from the parental species, which did not differ from each other (Fragaria × ananassa subsp. cuneifolia vs. F. virginiana subsp. platypetala or F. chiloensis) (Fig. 3B).

Within F. ×ananassa subsp. cuneifolia, we found a significant cytotype effect for PC1 (t1, 72.7 = 5.2, P = 0.0001) and PC2 (t1, 70.9 = −2.7, P = 0.01). Chloroplast genotype mirrored parental phenotype values for PC1 but not PC2. Moreover, among individuals, there were significant correlations between PCs and genetic cluster membership. Specifically, PC1 and genetic cluster 1 were positively correlated (r = 0.42, P < 0.0005, n = 64)—that is, plants with greater genetic affiliation to F. chiloensis had more coriaceous leaves and pronounced venation; whereas PC2 was negatively correlated with genetic cluster 1 (r = −0.35, P < 0.004, n = 64), which means that individuals with lower genetic affiliation to F. chiloensis had greater petiole pubescence but less leaf-specific mass.

Canonical discriminant function analyses produced two canonical variables: the first (CAN1) accounted for 81% of the variation and had high loadings of leaf coriaceousness and petiole pubescence (canonical coefficients of 1.6 and 3.8, respectively) whereas the second (CAN2) accounted for 19% and loaded heavily on leaf-specific mass and petiole pubescence (canonical coefficients of 1.66 and −1.7 respectively). The discriminant function analysis confirmed our classification for the three species overall 93% of the time, with 91% for F. chiloensis had greater petiole pubescence but less leaf-specific mass.

Canonical discriminant function analyses produced two canonical variables: the first (CAN1) accounted for 81% of the variation and had high loadings of leaf coriaceousness and petiole pubescence (canonical coefficients of 1.6 and 3.8, respectively) whereas the second (CAN2) accounted for 19% and loaded heavily on leaf-specific mass and petiole pubescence (canonical coefficients of 1.66 and −1.7 respectively). The discriminant function analysis confirmed our classification for the three species overall 93% of the time, with 91% for F. chiloensis, 89% for F. ×ananassa subsp. cuneifolia, and 100% for F. virginiana subsp. platypetala (Fig. 4). Of the 11% of F. ×ananassa subsp. cuneifolia that were classified as F. chiloensis in this analysis, seven individuals were from FtS and one (a minority cytotype) from WREN.

Environmental characteristics—Site characteristics of F. ×ananassa subsp. cuneifolia populations were in-between the parental species for all five measured environmental variables...
cated that hybrids between widely and deeply sampled populations revealed complex quencing and assembly identified 60 species-specific SNPs, lion years ago (Njuguna et al., 2013) whole-plastome se-
cytoplasmic genomes —

phism within populations is consistent with ongoing introgression. First, although there was significant admixture and little overall maternal bias in chloroplast or nuclear genomes within the hybrid, there was wide heterogeneity in the signature of introgression. This plus geographic structure suggests differences in the prevailing direction of gene flow or selection. Second, the hybrid expressed both ecological (i.e., bioclimatic) and pheno-
typic intermediacy, but there was still a discernible fingerprint in the prevailing direction of gene flow or selection. This plus geographic structure suggests differences that mirrored the habitat of the cytotype donor species, there was no evidence of cyto-
type differences for any of the bioclimatic variables (Fig. 5A–C; altitude: $t_{1,22.4} = -0.58$; temperature $U_{11,15} = 151.1$; precipita-
tion: $t_{1,16.3} = 0.29$; all $P > 0.45$).

DISCUSSION

Our genetic and ecological characterization of the hybrid zone of F. × ananassa subsp. cuneifolia revealed several patterns. First, although there was significant admixture and little overall maternal bias in chloroplast or nuclear genomes within the hybrid, there was wide heterogeneity in the signature of introgression. This plus geographic structure suggests differences in the prevailing direction of gene flow or selection. Second, the hybrid expressed both ecological (i.e., bioclimatic) and pheno-
typic intermediacy, but there was still a discernible fingerprint of the parental species within the hybrid. That is, there was an association between phenotype and genotype across hybrid indi-
viduals, and this was evident from markers in both nuclear and chloroplast genomes. Nonetheless, significant polymor-
phism within populations is consistent with ongoing introgression in this hybrid.

Evidence of pattern of hybridization from nuclear and cytoplasmic genomes—Consideration of both diagnostic chloroplast SNPs and nSSR confirmed the value of multiple genomes to assess hybridization between two closely related taxa (Bacilieri et al., 1996; Aboim et al., 2010). Even though F. chiloensis and F. virginiana diverged only 0.19–0.86 mil-

bias. Although all populations had a clear majority cytotype, both cytotypes were evident in 5 out of 7 deeply sampled F. × ananassa subsp. cuneifolia populations, indicating con-
tinued gene (seed) flow or incomplete cytotype exclusion, be-
cause paternal leakage (i.e., nonmaternal inheritance; McCauley et al., 2007) of the chloroplast has not been detected in hybrid crosses of F. × ananassa subsp. cuneifolia (Govindarajulu et al., 2012) or other Fragaria species (Davis et al., 2010). Ad-
mixed individuals were identified in all hybrid populations (i.e., proportional membership scores ranged from 2% to 99% cluster 1), and some populations showed a distinct incongruity between nuclear and cytotype affiliation (Fig. 2B, C). Even in the case of populations that showed little admixture within indi-
viduals, the majority nuclear and cytotypic affinities revealed introgression (i.e., MP was highly associated with cluster 2 but had F. chiloensis as the majority cytotype, whereas SF was highly associated with cluster 1 but had F. virginiana majority cytotype). That is not to say, however, that all populations of F. × ananassa subsp. cuneifolia showed a cyto–nuclear incongruity; cytotypic and nuclear affiliations corresponded to the same parental species at FtS and FB.

It is also interesting to note that Structure analysis de-
tected levels of nuclear admixture in some of the parental species’ populations (e.g., SAL and ES; mean cluster 1: 47% and 51%, respectively) that were equivalent to or greater than the average level of nuclear admixture across F. × anan-
assa subsp. cuneifolia populations (mean cluster 1: 53%). Populations like these (and the “misclassified” F. chiloensis individuals at FtS; Fig. 4) could reflect ongoing backcross-
ning between the hybrid and the parental species, as suggested by Staudt (1999) on the basis of phenotypic observations and habitat disturbance level. By contrast, however, we found no evidence of admixture in chloroplast genomes in the six pa-
rental species populations that were deeply sampled (Appen-
dix S3), which suggests that the observed admixture in nuclear genome of SAL and ES could simply reflect the recent diver-
gence of the parental species (Njuguna et al., 2013) and lack of private alleles at the neutral nSSR loci (Beaumont et al., 2001; Mirol et al., 2010). Additional data from highly vari-
able nuclear loci would help resolve this issue, because the high assignment probabilities based on phenotypic data sup-
port the dissimilarity of the parental species (Fig. 4).

Although, on average, both parental species contributed equally to the genomic constitution of F. × ananassa subsp. cuneifolia, populations of this hybrid species showed geo-
graphic structure. The F. chiloensis type dominated in the western region, and the F. virginiana type was more com-
mon in the inland eastern region. This pattern was seen most strongly in the maternally inherited chloroplast, which could suggest a geographic limitation to gene flow via seeds, perhaps

Fig. 2. (A) Geographic representation of Fragaria chiloensis (red +), F. virginiana subsp. platypetala (blue +), and F × ananassa subsp. cuneifolia (circles) populations from GBIF records, wide samples, and deep sampled populations. Fragaria × ananassa subsp. cuneifolia records are broken down by cytotype (red circle = F. chiloensis; blue circle = F. virginiana; yellow circle = unknown cytotype). Note: 47 F. chiloensis sites along the Pacific coastline to Alaska are not pictured. (B) Geographic representation of cytotype of F. × ananassa subsp. cuneifolia populations in the hybrid zone (red circle = F. chiloensis; blue circle = F. virginiana). Pie-chart inserts show the breakdown of maternity within deeply sampled populations (red = F. chiloensis; blue = F. virginiana). Sites marked by (×) are locations of F. chiloensis (red) and F. virginiana subsp. platypetala (blue) populations that were deeply sampled. (C) Bar plot of individual assignment probability for each individuation of deeply sampled populations of F. chiloensis, F. × ananassa subsp. cuneifolia, and F. virginiana subsp. platypetala produced by Structure. Populations are organized in order of increasing longitude within each species. Red = cluster 1, high values of membership held by F. chiloensis; blue = cluster 2, high values of membership held by F. virginiana subsp. platypetala. The highlighted asterisks on top indicate individuals with a minority cytotype within a given deeply sampled population.
Phenotypic and bioclimatic patterns of hybridization—
Hybrids occupied sites bioclimatically and ecologically intermediate between the parental species (Table 1; Fig. 5), and those with *F. chiloensis* cytotype tended toward sites more like the *F. chiloensis* habitat (similarly for *F. virginiana*). Such a pattern could reflect selection for coordination between the nuclear and chloroplast genomes for physiological competency (especially photosynthesis and WUE) and high performance in parent-like sites (Wu and Campbell, 2007). Consistent with this hypothesis, we found that greenhouse-grown hybrids with *F. chiloensis* cytotype were more similar to *F. chiloensis* in PC1 (leaf coriaceousness and reticulation of venation) and that there was a significant correlation between PC1 and nuclear affinity to *F. chiloensis* (e.g., genetic cluster 1). In addition, Hancock et al. (1989) found a positive relation between photosynthetic rate and proportion *F. chiloensis* cultivated hybrids. In sum, these data support an ecologically based selective hypothesis for the hybrid-zone heterogeneity, but experiments with genetically characterized natural hybrids or experimentally produced ones grown and phenotyped under field conditions would be needed to confirm this (Wu and Campbell, 2007).

Two populations were fixed for one species’ chloroplast genome, but with strong nuclear affinity to the other species (MP and SF: 7% and 81% *F. chiloensis*, respectively), provide an interesting contradiction to the cyto–nuclear coordination hypothesis. These two populations are in the meadows of the Coast and Cascades mountain ranges and could reflect long-distance bird-mediated seed dispersal followed by widespread pollen swamping. Such patterns have been proposed in other species whereby

Fig. 3. (A) PC1 and (B) PC2 from a principal component analysis based on six morphological traits. Mean ± SE are shown for *Fragaria chiloensis*, *F. ×ananassa* subsp. *cuneifolia*, and *F. virginiana* subsp. *platypetala* with inserts reflecting the subset of *F. ×ananassa* subsp. *cuneifolia* with *F. chiloensis* (striped bars) or *F. virginiana* (white bars) cytotype.

Fig. 4. Scatter plot of canonical discriminant scores (CAN1 and CAN2) for each plant (genotype) in the deep population sampling. Circles are the parental species (*F. chiloensis* [filled] and *F. virginiana* subsp. *platypetala* [open]). Diamonds are *F. ×ananassa* subsp. *cuneifolia* with *F. chiloensis* cytotype (filled) or *F. virginiana* subsp. *platypetala* cytotype (open).
Conclusions—Despite the heterogeneity among populations, the genetic results as a whole suggest a high degree of interfertility between the two parental and the hybrid species and point to a long history of backcrossing and genetic mixing in the hybrid zone. Our results also lend support to conjectures made by previous authors that *F. virginiana* ssp. *platypetala* is morphologically and genetically more similar to *F. chiloensis* than to *F. virginiana* (Harrison et al., 1997), although this appears more clearly in the phenotype than in the genotype. The present work is novel in that it uncovers the genetic complexity that underlies the hybrid zone first hypothesized by Staudt (1999) and provides a geographic fingerprint of introgression that reflects distance to the parental species’ range margins. Whether this is due to gene flow and/or reinforced by selection for specific parental traits in the bioclimatically intermediate habitat occupied by the hybrid will require additional work to dissect the causes of genomic introgression and nuclear–cytoplasmic disassociation. Nonetheless, the present work provides evidence of gene flow between *F. virginiana* and *F. chiloensis* and suggests that this hybrid zone may provide an exciting ecological

<table>
<thead>
<tr>
<th>Species</th>
<th>Percent full sun</th>
<th>Percent moisture</th>
<th>N (%)</th>
<th>P (ppm)</th>
<th>K (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>F. chiloensis</em></td>
<td>71.9 ± 19.9</td>
<td>3.6 ± 1.4</td>
<td>0.04 ± 0.03</td>
<td>16.5 ± 3.5</td>
<td>39.1 ± 0</td>
</tr>
<tr>
<td><em>F. × ananassa</em> ssp. <em>cuneifolia</em></td>
<td>29.2 ± 7.2</td>
<td>16.9 ± 5.1</td>
<td>0.23 ± 0.09</td>
<td>54.1 ± 17.7</td>
<td>162.0 ± 21.6</td>
</tr>
<tr>
<td><em>F. virginiana</em> ssp. <em>platypetala</em></td>
<td>23.2 ± 6.8</td>
<td>27.7 ± 6.2</td>
<td>0.34 ± 0.11</td>
<td>66.7 ± 35.7</td>
<td>404.0 ± 175.3</td>
</tr>
</tbody>
</table>

Table 1. Habitat characteristics (mean ± SE: sun exposure, soil moisture, and soil nutrient content: nitrogen [N]; phosphorus [P]; potassium [K]) for *Fragaria chiloensis*, *F. × ananassa* ssp. *cuneifolia*, and *F. virginiana* ssp. *platypetala* populations.

long-distance dispersal leads to a cytotypic bottleneck (e.g., Le Corre et al., 1997), and unidirectional pollen swamping can lead to the extinction of one nuclear gene pool over time (Beatty et al., 2010). In addition, similar to asymmetries derived from the mating system or pollination (Wallace et al., 2011), one might predict that loss of one parental nuclear genome may occur faster when hybridizing species are dioecious (separate sexes) because half of the propagules do not produce pollen (females) and, thus, may be more likely to be pollinated by resident individuals (i.e., the other species) than by comigrants. Engel et al. (2005) proposed a similar idea when explaining the asymmetric introgression between hermaphroditic *Fucus spiralis* and dioecious *F. vesiculosus*: lower introgression of *F. spiralis* into *F. vesiculosus* was attributed to an asymmetry in sperm production. Alternatively, because the two hybridizing species may have different sex-determining regions (Goldberg et al., 2010; Govindarajulu et al., 2012), it is possible that sterility–inviability that results from hybridization (Haldane’s rule; Orr, 1997) could lead to asymmetries in reproduction or survival that are reflected in the incongruity of genomes (R. Wang, unpublished data). These ideas can be effectively explored by comparing introgression of markers linked to the sex-determining region with others elsewhere in the genome (Payseur, 2010). Integrating genetic linkage maps (Goldberg et al., 2010; Govindarajulu et al., 2012) and the genome sequence of *Fragaria vesca* (Shulaev et al., 2011) will help facilitate this endeavor.

![Fig. 5. Mean ± SE (A) altitude, (B) annual temperature, and (C) annual precipitation for *F. chiloensis*, *F. × ananassa* ssp. *cuneifolia*, and *F. virginiana* ssp. *platypetala*, with inserts reflecting the subset of *F. × ananassa* ssp. *cuneifolia* with *F. chiloensis* (striped bars) or *F. virginiana* (white bars) cytotype.](image-url)
context for elucidating other parts of the evolutionary history of octoploid Fragaria, such as the dynamic nature of sex-chromosome evolution in these species (Govindaraju et al., 2012).

LITERATURE CITED


Taxon, accession number—locality—GenBank number.

F. virginiana subsp. virginiana Duchesne, PI 612492 - Ontario, Canada

F. virginiana subsp. glauca (S.Watson) Staudt, BOX5 - Idaho, USA - 44.047N, 111.391W—JX064443, JX064458. BOX10 - Idaho, USA - 44.047N, 111.391W—JX064442, JX064457.

F. virginiana subsp. platypetala (Ryd.) Staudt, ES17 - Oregon, USA - 45.343N, 112.615W—JX064429, JX064468.


F. virginiana subsp. virginiana Duchesne, PI 612492 - Ontario, Canada