Molecular Phylogenetics and Evolution 66 (2013) 17–29

Contents lists available at SciVerse ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Insights into phylogeny, sex function and age of *Fragaria* based on whole chloroplast genome sequencing

Wambui Njuguna [a,*], Aaron Liston [b], Richard Cronn [c], Tia-Lynn Ashman [d], Nahla Bassil [e]

[a] Department of Horticulture, Oregon State University, ALS 4017, Corvallis, OR 97331, USA
[b] Department of Botany and Plant Pathology, Oregon State University, Cordley Hall 2082, Corvallis, OR 97331, USA
[c] Pacific Northwest Research Station, United States Forest Service, 3200 SW Jefferson Way, Corvallis, OR 97331, USA
[d] Department of Biological Sciences, University of Pittsburgh, Pittsburgh, PA 15260, USA
[e] United States Department of Agriculture, Agricultural Research Service, National Clonal Germplasm Repository, Corvallis, OR 97333, USA

ARTICLE INFO

Article history:
Received 13 February 2012
Revised 20 August 2012
Accepted 30 August 2012
Available online 14 September 2012

Keywords:
Plastome
Polyploidy
Self-compatibility

ABSTRACT

The cultivated strawberry is one of the youngest domesticated plants, developed in France in the 1700s from chance hybridization between two western hemisphere octoploid species. However, little is known about the evolution of the species that gave rise to this important fruit crop. Phylogenetic analysis of chloroplast genome sequences of 21 *Fragaria* species and subspecies resolves the western North American diploid *F. vesca* subsp. *bracteata* as sister to the clade of octoploid/decaploid species. No extant tetraploids or hexaploids are directly involved in the maternal ancestry of the octoploids.

There is strong geographic segregation of chloroplast haplotypes in subsp. *bracteata*, and the gynodioecious Pacific Coast populations are implicated as both the maternal lineage and the source of male-sterility in the octoploid strawberries. Analysis of sexual system evolution in *Fragaria* provides evidence that the loss of male and female function can follow polyploidization, but does not seem to be associated with loss of self-incompatibility following genome doubling. Character-state mapping provided insight into sexual system evolution and its association with loss of self-incompatibility and genome doubling/merger. *Fragaria* attained its circumboreal and amphitropical distribution within the past one to four million years and the rise of the octoploid clade is dated at 0.372–2.05 million years ago.

1. Introduction

The domestication of strawberry (*Fragaria* ×*ananassa* subsp. *ananassa*) is well documented historically. Female plants of *F. chiloensis* from Chile were brought to France in 1716 by a French army officer, Amédée Frézier. He gave one of the plants to Antoine de Jussieu, the director of the King's garden in Paris, where clones of the eastern North American dioecious *F. virginiana* imported from eastern N. America in the early to mid 17th century were growing. Cultivated individuals derived from the interspecific hybrids of these two species (*F. chiloensis*, *F. virginiana*) in the early 18th century created the modern strawberry grown today (Hancock, 1999). Beyond these recent historical events, however, there is far less known about the evolution of the two species that contributed to this important fruit crop.

As is common in domesticated plants, strawberry is a polyploid (octoploid; $2n = 8x = 56$), and the genus *Fragaria* contains extensive natural ploidy variation with species known at five ploidy levels (diploid through decaploid). In addition to differences in ploidy, *Fragaria* contains a wide range of diversity in sexual systems (Staudt, 1989). Two-thirds of the 27 recognized taxa (species and subspecies) are hermaphroditic and either self-compatible (SC) (10 species) or self-incompatible (SI) (seven species). Ten species show some degree of sexual polymorphism, including: gynodioecy (females and hermaphrodites) in *F. vesca* subsp. *bracteata*; subdioecy (females, hermaphrodites and males) in *F. chiloensis*, *F. virginiana*, and their naturally occurring hybrid *F.* ×*ananassa* subsp. *cuneifolia*; and dioecy (females and males) in *F. corymbosa*, *F. gracilis*, *F. moschata*, *F. moupinensis*, *F. orientalis and F. tibetica* (Staudt, 2009). The diversity in chromosome number and mating systems makes *Fragaria* an exceptional system for understanding sexual system evolution, especially in light of recent genetic mapping that uncovered proto-sex chromosomes and sex chromosome turnover in the genus (Goldberg et al., 2010; Spigler et al., 2008).

The diploid ancestry of the octoploid parents of the cultivated strawberry, *F.* ×*ananassa* subsp. *ananassa,* has been studied through the use of interspecific hybridization and cytogenetic methods (Bringhurst, 1990; Federova, 1946; Senanayake and Bringhurst, 1967). These studies primarily implicated the widespread species *F. vesca* and its close relatives as potential diploid ancestors to these

* Corresponding author. Present address: Eurofins Lancaster Laboratories, 2425 New Holland Pike, Lancaster, PA 17605, USA. Fax: +1 541 738 4205.
E-mail addresses: WYoung@lancasterlabs.com (W. Njuguna), listona@science.oregonstate.edu (A. Liston), rcronn@fs.fed.us (R. Cronn), tia1@pitt.edu (T.-L. Ashman), nahla.bassil@ars.usda.gov (N. Bassil).

1055-7903/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ympev.2012.08.026

lineages. Molecular phylogenetic studies based on chloroplast DNA and nuclear ribosomal ITS sequences (Harrison et al., 1997; Potter et al., 2000) identified a clade of *F. vesca* and related species as the likely maternal ancestral lineage contributing to the octoploid cytoplasm. The first published study to use low-copy nuclear loci (Rousseau-Gueutin et al., 2009) also supported the *F. vesca* clade as a diploid ancestor, and further identified a Japanese species, *F. iinumae*, as contributing to the genome composition of the octoploids. Despite these advances, several important questions remain unresolved: (1) was there a single origin of the octoploids?; (2) can a single extant member of the *F. vesca* clade be identified as a direct ancestor of the octoploids?; (3) what is the likely geographic location of the octoploid origin? and (4) how long ago did this divergence take place?

Furthermore, given the diversity of sexual systems in strawberry, phylogenetic analysis can test conflicting hypotheses for the relationship between sexual system and whole-genome duplication (Baker, 1984; Jennings, 1976; Brunet and Liston, 2001; Osborn et al., 2003; Miller and Venable, 2000). For instance, we can determine whether loss of self-incompatibility with increase in ploidy facilitates the evolution of separate sexes (dioecy) (Miller and Venable, 2000) or if other mechanisms like the acquisition of male sterility or female sterility mutations (Brunet and Liston, 2001) are involved.

Phylogenetic analysis of closely related taxa using nearly-complete chloroplast genomes (Parks et al., 2009; Straub et al., 2012) can provide unprecedented insights into phylogenetic relationships and biogeographic history, as well as allow the first explicit tests of hypotheses for sexual system evolution in the genus *Fragaria*. We therefore sequenced nearly complete chloroplast genomes from 21 *Fragaria* species to assess the phylogeny, biogeography and sexual system evolution.

2. Materials and methods

2.1. Plant material

Twenty-five accessions representing 21wild *Fragaria* species, subspecies and hybrids, and one accession of *Potentilla*, a close relative of *Fragaria* in Rosaceae (Eriksson et al., 2003; Lundberg et al., 2009; Dobeš and Paule, 2010), were included in the study (Table 1).

2.2. DNA extraction and PCR

DNA was extracted from actively-growing leaves using a protocol based on the PUREGENE® kit (Gentra Systems Inc., Minneapolis, MN). Preparations of chloroplast DNA for sequencing were obtained from genomic and chloroplast PCR fragment pools (Table 1). To generate PCR fragment pools, 203 chloroplast primers (108 forward, 95 reverse) were screened in various combinations in four species (*F. iinumae*, *F. nipponica*, *F. orientalis*, and *F. virginiana*) to identify pairs that amplified fragments that were at least 2.5 kb in size, and to provide maximum coverage of the chloroplast genome. Where possible, primer pairs that amplified single bands in most or all of the species were chosen. Of these 203 primers, 141 were previously reported to amplify the *Cucumis sativus* L. chloroplast genome (Chung et al., 2007); 25 were designed from the genome sequence of *Morus indica* M. alba 'K2' (Ravi et al., 2006) and 36 primers were designed in this study from *F. vesca* 'Hawaii 4' (Shulaev et al., 2010). Sixty-three primer pairs (Njuguna, 2010) were finally chosen to amplify the entire chloroplast genomes of 17 accessions (Table 1). We used Phusion™ High-Fidelity DNA polymerase (New England Biolabs, Ipswich, MA) for long-range PCR. Amplifications were performed in 10 μl reactions containing 1× Phusion GC buffer, 2.5 mM of each dNTP, 10 μM of each primer, 5U of Phusion DNA polymerase, 0.05 μl of 3% DMSO and 5 ng of DNA template. PCR product quantification was carried out using the Quant-iT™ PicoGreen® dsDNA quantification protocol (Molecular Probes, Inc., Eugene, OR) following the manufacturer's specifications. Equimolar amounts of PCR products were pooled for each species to generate 1–5 μg of chloroplast DNA for Illumina sample preparations.

2.3. Illumina library preparation

For 11 accessions, we directly assembled the plastome from genomic sequencing (reviewed in Straub et al., 2011). For both genomic DNA and that obtained from chloroplast PCR fragment pools, sequencing libraries (Table 1) were prepared for sequencing using the sample preparation kit from Illumina (Illumina Inc., San Diego, CA) and as described by Cronn et al. (2008). Briefly, this approach utilizes Illumina sequencing technology to sequence multiple barcoded PCR amplified chloroplast genomes in one lane of a flow cell. Multiplexing of small organellar genomes in a single lane utilizes the high sequencing capacity of this platform (>40 million clusters per flowcell during this study). For details on the Illumina sequencing runs performed, refer to Table 1.

2.4. Sanger sequencing of SNPs

DNA sequences encompassing loci containing three parsimony-informative octoploid and decaploid specific SNPs were PCR amplified and sequenced using Sanger sequencing methods. Primers that can amplify three genes (*ndh*F, *ccs*A, and *rpo*C2) containing parsimony-informative SNPs between *F. vesca* subsp. *bracteata* and the octoploid clade were designed with Primer 3 (Rozen and Skaletsky, 2000). They included ndhF-715F/ndhF-715R (5′- GTAAAAGGTTTATGGACGGAGTT-3′, 5′-GCATTGTTGTTTTTAGGATCTGG-3′); ccsA-731F/ccsA-731R (5′-CCTTTGGTGAGATTCAATACGTG-3′, 5′-GACAAGGCCGAAGCTATTCTATC-3′), and rpoC2F/rpoC2R (5′-GGAATTCGAAATTCTCCCGTTT-3′, 5′- AGGGATAATCTAGAGCTTCGAGTTG-3′), respectively. PCR was carried out as described above, followed by Exonuclease-Shrimp Alkaline Phosphatase (ExoSAP) cleanup. This procedure involved mixing 4 μl of PCR product with an 8.1 μl mixture of 2 μl shrimp alkaline phosphatase (SAP; 1 unit/μl), 0.1 μl exonuclease I (Exo I; 20 units/μl) and 6 μl of water. The mixture was incubated at 37 °C for 60 min, followed by 72 °C for 15 min. The samples were submitted for sequencing at the Center for Genome Research and Biocomputing (CGRB) at Oregon State University in Corvallis, Oregon.

2.5. Data analysis

After the sequencing run, raw image data for each sequencing cycle was processed into base calls and alignment files through the Illumina Pipeline (version 0.2.2.6). Binning was carried out using the three nucleotide barcodes. After sorting microreads (36, 40 or 60 bp) into sample-specific bins, the barcodes and adapter tags were removed, and resulting microreads (32, 36 or 56 bp) were used for subsequent analysis. The 155,691 kb *F. vesca* 'Hawaii 4' annotated chloroplast genome (Genbank accession JF345175; Shulaev et al., 2010) was used for reference-guided microread assembly and was also included in the phylogenetic analysis. Microreads were assembled into contigs using YASRA (Ratan, 2009), and Mulan (Ovcharenko et al., 2005) was used to assemble and align contigs. Bioedit (Hall, 1999) was used to manually check and correct mis-alignments, remove primer sequences and score indels. Assembly errors resulting in high sequence divergence and insertions were recognized by low sequencing depth (<5*x*) data from the YASRA output, and were masked in the final assemblies. Calculation of variable and parsimony informative sites in

Table 1
List of 25 taxa (24 *Fragaria* L., 1 *Potentilla* L.) sequenced. Also included are the mating system, ploidy, library source (PCR amplified chloroplast DNA and/or genomic DNA), 3 bp tag, the number of contigs, estimated chloroplast genome coverage (Cov.), the number of reads, and Genbank accession number. CFRA indicates NCGR local accession number.

PI[a]	Species	Mating system	Ploidy	Illumina library[d]	Tag	Contigs	Cov. (%)	Reads	Genbank accession numbers[e]
551805	*Fragaria ×ananassa* subsp. *cuneifolia*	Subdioecious	8	PCR	aac	120	82	864943	JX117907, JX117923, JX117938, JX117952, JX117968, JX117983, JX117997, JX118012, JX118027, JX118042, JX118058, JX118077, JX118093, JX118111, JX118126, JX118144, JX118160, JX118171, JX118186, JX118205, JX118220, JX118233, JX118245
616613	*F. ×bifera*	Self compatible	2	PCR	acg	243	78	470616	JX117905, JX117936, JX117950, JX117967, JX117982, JX117996, JX118011, JX118026, JX118041, JX118056, JX118075, JX118091, JX118109, JX118124, JX118142, JX118184, JX118203, JX118218, JX118231, JX118243
551853[c]	*F. bucharica*	Self incompatible	2	Total genomic	gat	82	98	6729667	JX117909, JX117925, JX117940, JX117954, JX117970, JX117985, JX117999, JX118014, JX118029, JX118044, JX118060, JX118078, JX118095, JX118113, JX118128, JX118146, JX118162, JX118173, JX118188, JX118207, JX118222, JX118247
612318	*F. chiloensis*	Subdioecious	8	PCR and genomic	gct	621 (combined)	99	7028065	JX402801, JX402803–JX402857
616583	*F. chinensis*	Self incompatible	2	PCR	ccc	116	83	820618	JX117913, JX117929, JX117944, JX117958, JX117973, JX117988, JX118002, JX118017, JX118032, JX118047, JX118065, JX118083, JX118100, JX118116, JX118133, JX118152, JX118166, JX118177, JX118193, JX118212, JX118226, JX118252
657846	*F. corymbosa*	Dioecious	4	PCR	agc	66	84	1600377	JX117917, JX117932, JX117948, JX117962, JX117977, JX117992, JX118006, JX118021, JX118036, JX118051, JX118069, JX118087, JX118104, JX118120, JX118137, JX118156, JX118169, JX118180, JX118197, JX118228, JX118238, JX118256
641195	*F. daltoniana*	Self compatible	2	PCR	gta	120	82	621757	JX117916, JX117931, JX117947, JX117961, JX117976, JX117991, JX118005, JX118020, JX118035, JX118050, JX118068, JX118086, JX118103, JX118119, JX118136, JX118155, JX118196, JX118214, JX118237, JX118255
CFRA1973	*F. gracilis*	Dioecious	4	Total genomic	ccc	13	100	1140518	JX117918, JX117933, JX117949, JX117963, JX117978, JX117993, JX118007, JX118022, JX118037, JX118052, JX118070, JX118088, JX118105, JX118121, JX118138, JX118157, JX118170, JX118181, JX118198, JX118215, JX118229, JX118239, JX118257
637963	*F. iinumae*	Self compatible	2	PCR and genomic[d]	acg	538 (combined)	78	607838	JX117919, JX117979, JX117994, JX118008, JX118023, JX118038, JX118053, JX118072, JX118106, JX118122, JX118139, JX118158, JX118200, JX118240, JX118258
641091	*F. iturupensis*	Subdioecious	10	PCR	tgc	117	78	1921668	JX117906, JX117922, JX117937, JX117951, JX118057, JX118076, JX118092, JX118110, JX118125, JX118143, JX118185, JX118204, JX118219, JX118232, JX118244
657855	*F. mandschurica*	Self incompatible	2	PCR	agc	172	81	682461	JQ396172
551528	*F. moschata*	Dioecious	6	PCR and genomic	tgc	813 (genomic)	63	629986	JX468952–JX469013
CFRA1974	*F. moupinensis*	Dioecious	4	Total genomic	tac	115	98	731634	JX117914, JX117945, JX117959, JX117974, JX117989, JX118003, JX118018, JX118033, JX118048, JX118066, JX118084, JX118101, JX118117, JX118134, JX118153, JX118167, JX118178, JX118194, JX118213, JX118253
616672	*F. nilgerrensis*	Self compatible	2	PCR	tac	147	80	1042414	JX117911, JX117927, JX117942, JX117956, JX118063, JX118081, JX118098, JX118115, JX118131, JX118150, JX118165, JX118176, JX118191, JX118210, JX118224, JX118250
637975	*F. nipponica*	Self incompatible	2	PCR	gat	171	79	579426	JX117912, JX117928, JX117943, JX117957, JX118064, JX118082, JX118099, JX118132, JX118151, JX118192, JX118211, JX118225, JX118235, JX118251
637933	*F. orientalis*	Dioecious	4	Total genomic	atg	151	96	1011201	JX117971, JX117986, JX118000, JX118015, JX118030, JX118045, JX118061, JX118079, JX118096, JX118129, JX118148, JX118163, JX118174, JX118189, JX118208, JX118248
651568	*F. pentaphylla*	Self incompatible	2	Total genomic	gta	359	61	628309	JX469014–JX469080
651567	*F. tibetica*	Dioecious	4	PCR	ccc	87	82	1193513	JX117915, JX117930, JX117946, JX117960, JX117975, JX117990, JX118004, JX118019, JX118034, JX118049, JX118067, JX118085, JX118102, JX118118, JX118135, JX118154, JX118168, JX118179, JX118195, JX118227, JX118236, JX118254
552286	*F. vesca* subsp. *americana*	Self compatible	2	Total genomic	aac	213	91	478356	JX117965, JX117980, JX118009, JX118024, JX118039, JX118054, JX118089, JX118182, JX118241
551646	*F. vesca* subsp.	Gynodioecious	2	Total genomic	cgt	288	85	436230	JQ396171

(*continued on next page*)

Table 1 (*continued*)

PI[a]	Species	Mating system	Ploidy	Illumina library[d]	Tag	Contigs	Cov. (%)	Reads	Genbank accession numbers[e]
	bracteata								
551507	*F. vesca* subsp. *vesca*	Self compatible	2	PCR	ctg	219	80	647651	JX117904, JX117921, JX117935, JX117964, JX117966, JX117981, JX117995, JX118010, JX118025, JX118040, JX118055, JX118074, JX118090, JX118108, JX118123, JX118141, JX118183, JX118202, JX118217, JX118242
612492	*F. virginiana*	Subdioecious	8	PCR	cac	127	80	727536	JX117908, JX117924, JX117939, JX117953, JX117969, JX117984, JX117998, JX118013, JX118028, JX118043, JX118059, JX118094, JX118112, JX118127, JX118145, JX118161, JX118172, JX118187, JX118206, JX118221, JX118246
616857	*F. viridis*	Self incompatible	2	PCR	ctg	81	81	962316	JX117910, JX117926, JX117941, JX117955, JX117972, JX117987, JX118001, JX118016, JX118031, JX118046, JX118062, JX118080, JX118097, JX118114, JX118130, JX118149, JX118164, JX118175, JX118190, JX118209, JX118223, JX118234, JX118249
652552	*P. villosa*	–	–	PCR	tac	228	74	1133877	JX117920, JX117934, JX118073, JX118107, JX118140, JX118159, JX118201, JX118216, JX118230, JX118259

[a] PI refers to plant introduction number of the National Plant Germplasm System (NPGS).
[c] A presumed *F. nubicola* representative (PI 551853) was allowed to grow and displayed non-characteristic sympodial runnering and thus we consider it *F. bucharica*, also endemic to the Himalayan region.
[d] PCR and genomic samples: Contig assembly was done separately for the PCR reads and genomic reads in YASRA. The contigs were combined prior to submission to Mulan for contig assembly and alignment. The resulting alignment (containing the contigs assembled from the two illumina samples) was used for subsequent analysis.
[e] Sequence with multiple Genbank accession numbers were submitted in sections due to missing sequence between amplicons.

the alignment and calculation of pairwise distances was done in MEGA 4.0 (Tamura et al., 2007). The VISTA genome browser (Mayor et al., 2000) was used to visualize the coverage of the chloroplast genome among *Fragaria* species.

Preliminary phylogenetic analyses (Neighbor Joining, Maximum Parsimony) were performed in MEGA 4.0 (Tamura et al., 2007). Maximum likelihood (ML) analysis with rapid bootstrapping (RAxML; Stamatakis et al., 2008) was performed via the CIPRES Web Portal 2.0 (http://www.phylo.org/). The Reconstruct Ancestral State in Phylogenies (RASP) software (Yu et al., 2011) was used to infer biogeographic history based on the ML estimate of phylogeny. Wild *Fragaria* species, except for the amphitropical (Pacific coasts of North America and southern South America) *F. chiloensis,* are distributed in north temperate and holarctic zones (Hummer and Hancock, 2009). For the biogeographic analysis, three geographic regions were used in the reconstruction: East Asia, Europe and North/South America. The dispersal vicariance analysis (DIVA) was carried out with the default settings in RASP: number of chains = 10, frequency of sample = 100, discard samples = 100, temperature = 0.1, maximum number of areas = 3, state frequencies = fixed (Jukes–Cantor) and an across-site rate variation = equal.

The applicability of a strict molecular clock under the GTR+γ+I model (selected using the AIC criterion following Posada and Crandall, 1998) of nucleotide evolution was rejected by the likelihood ratio test (p = 0.00; HyPhy 2.0; Pond and Muse, 2005). For this reason, divergence time calculations were performed in BEAST v1.6.1 (Drummond et al., 2006), under an uncorrelated lognormal molecular clock model. To use a fossil calibration, plastome sequences from five members of Fabidae and the outgroups *Populus* and *Vitis* were added to the analysis (Table 2). The most recent common ancestor of *Castanea* and *Cucumis* has been dated to the Santonian, 84.0–87.5 MYA, based on the fossil *Bedellia* (Fagales) (Sims et al., 1999; Moore et al., 2010). The most recent common ancestor of *Malus* and *Prunus* has been dated to 49.42 ± 0.54 MYA, following Benedict et al. (2011). The fossil priors were treated as an exponential distribution with a mean of 5.0 and offsets of 84 MYA and 48.4 MYA, respectively. A total of 78 protein-coding genes and four ribosomal RNAs (75,341 bp) were extracted and aligned with homologous sequences from *Potentilla villosa* (Plant Introduction [PI] Accession 652552) and four *Fragaria* plastome sequences (*F. vesca* subspecies *vesca*, JF345174; *F. gracilis*, CFRA 1973; *F. chiloensis*, PI 612318; and *F. iturupensis*, PI 641091). We ran three MCMC runs with 100 million steps and sampling every 500 steps (200,000 trees). The output from BEAST was analyzed in the MCMC trace analysis package Tracer v1.5 (http://www.beast.bio.ed.ac.uk/) to determine when stationarity was reached. A consensus tree was visualized in FigTree v.1.3.1 (A. Rambaut, University of Edinburgh; http://www.tree.bio.ed.ac.uk/software/figtree/).

Table 2
Genbank numbers of additional plastome sequences used for BEAST analysis.

Family	Species	Genbank accession or other source
Rosaceae	*Malus domestica*	Velasco et al. (2010), http://www.rosaceae.org/
Rosaceae	*Prunus persica*	HQ336405.1
Moraceae	*Morus indica*	DQ226511.1
Fagaceae	*Castanea mollissima*	HQ336406.1
Cucurbitaceae	*Cucumis sativus*	AJ970307.1
Vitaceae	*Vitis vinifera* cultivar Maxxa	DQ424856.1
Salicaceae	*Populus trichocarpa*	EF489041.1

2.6. Ancestral character state reconstruction in Fragaria

We reconstructed the evolution of two reproductive traits (mating system and sexual system) and ploidy level onto the phylogeny of the genus *Fragaria*. Information on character states for these traits was obtained from the literature (Ahmadi and Bringhurst, 1989; Goldberg et al., 2010; Spigler et al., 2008; Staudt, 1967, 1968, 1989). Given that terminology used to describe polymorphic sexual systems and functional transitions between them is often indistinct, we chose to map male and female function separately rather than mapping the sexual system per se. The male-sterility mutation is reported to be dominant to the male-fertility allele in several species (*F. vesca* subsp. *bracteata, F. orientalis, F. moschata, F. chiloensis,* and *F. virginiana*), whereas both dominant (*F. orientalis* and *F. moschata*) and recessive (*F. chiloensis* and *F. virginiana*) female sterility has been found in strawberry (Ahmadi and Bringhurst, 1989; Spigler et al., 2008; Staudt, 1967, 1968). In four

species the dominance relations are unknown (*F.* ×*ananassa* subsp. *cuneifolia, F. moupinensis, F. tibetica,* and *F. corymbosa*). The character states for each trait were coded as categorical data. For the mating systems, self compatibility was coded as "0" and gametophytic self incompatibility, "1". For the sex expression, female or male function was coded as "0" for sterile and "1" for fertile. The ploidy level was coded as follows: 2x "0", 3x "1", 4x "2", 6x "3", 8x "4", and 10x "5". The datasets for these traits were generated and converted into a NEXUS format and ancestral state reconstruction was conducted in Mesquite v.2.73 (Maddison and Maddison, 2009). The maximum likelihood (ML) tree and 500 ML bootstrap analyses were performed using RAxML (Stamatakis et al., 2008) and all branches with less than 75% bootstrap support were collapsed to construct a final tree for ancestral state reconstruction. Evolution of ancestral states for all traits was mapped onto the ultrametricized phylogeny using a trace character history function in Mesquite by following the maximum likelihood ancestral state reconstruction method.

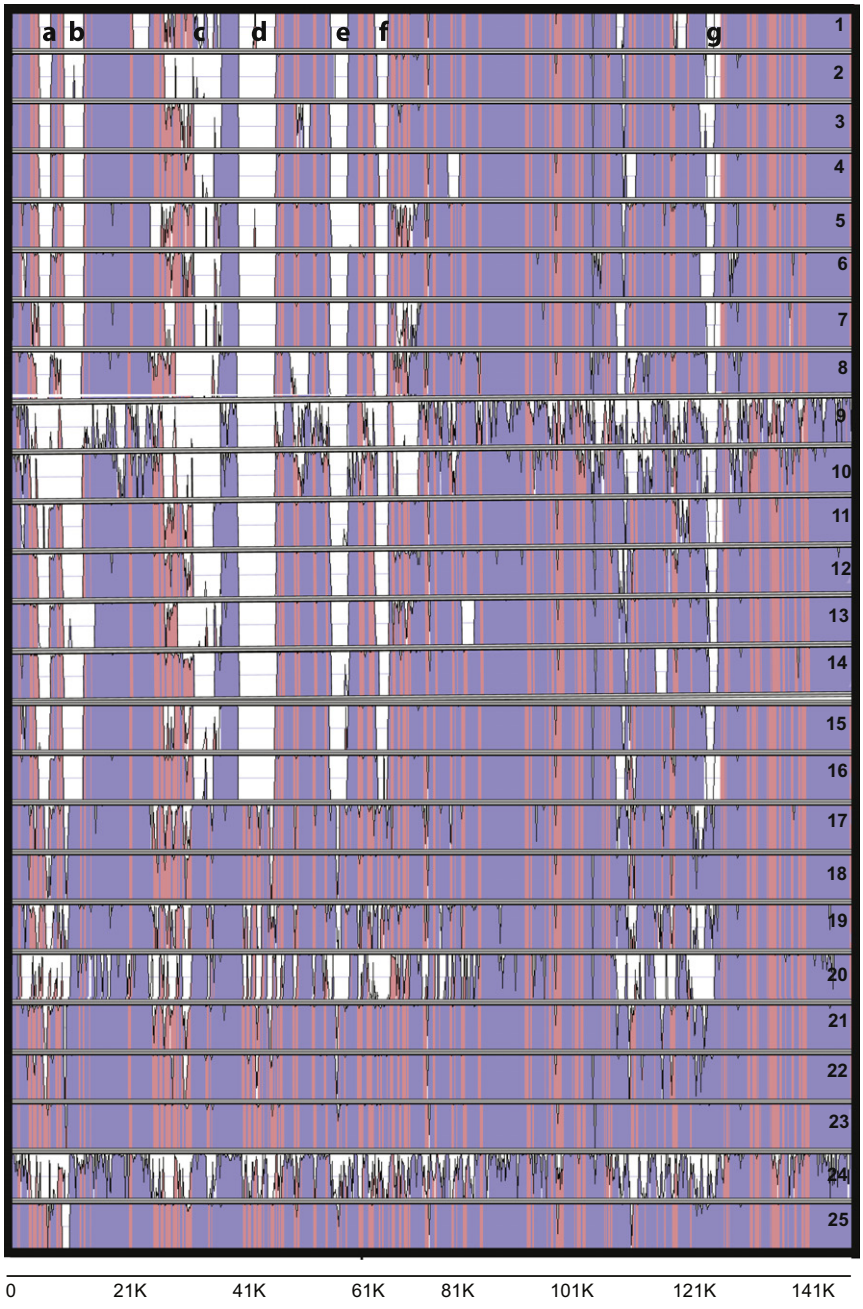

Fig. 1. A screen shot of the VISTA genome browser (http://www.pipeline.lbl.gov) output showing the chloroplast genome sequenced from each sample (PCR pools – samples 1–16, Genomic pools – samples 17–25). Peaks and valleys in the chart represent the percent conservation of sequence between each sample and the *Fragaria* reference sequence. Pink represents noncoding and dark blue illustrates exons; white regions represent missing sequence information. Regions that were highly variable and that could not be aligned were deleted from the alignment. These regions corresponded to regions of low conservation in the VISTA snap shot. The missing regions from sequenced PCR pools were found in regions between (a) *rps*16 and *trn*Q-UGG, (b) *trn*G-GCC and *atp*H, (c) *psb*D and *rps*14 (d) *psa*A and *trn*L-UAA, (e) *acc*D and *cem*A, (f) *pet*L and *rpl20* and, (g) a 1.5 kb region of the *ycf*1. (Key: 1 – *F. nilgerrensis*, 2 – *F. viridis*, 3 – *F. virginiana*, 4 – *F. tibetica*, 5 – *F. niponica*, 6 – *F. mandschurica*, 7 – *F. iinumae*, 8 – *F.* ×*bifera*, 9 – *F. bucharica* (PI: 551851), 10 – *F. chiloensis*, 11 – *F. chinensis*, 12 – *F. vesca*, 13 – *F. iturupensis*, 14 – *F. daltoniana*, 15 – *F. cuneifolia*, 16 – *F. corymbosa*, 17 – *F. vesca* subspecies *americana*, 18 – *F. bucharica* (PI: 551853), 19 – *F. vesca* subspecies *bracteata*, 20 – *F. pentaphylla*, 21 – *F. orientalis*, 22 – *F. moupinensis*, 23 – *F. gracilis*, 24 – *F. moschata*, 25 – *F. chiloensis*.

3. Results and discussion

3.1. Chloroplast genome recovery

The overall chloroplast genome assemblies obtained from sequencing genomic DNA and PCR pools of 25 accessions representing 21 *Fragaria* and one *Potentilla* ranged from 49% to 99% complete, with genomic libraries yielding significantly larger assemblies than PCR libraries (p value = 0.008) (Table 1, Fig. 1). Common missing regions from sequenced PCR products were found localized in the large single copy (LSC) region except for *ycf*1 found in the small single copy (SSC) region of the chloroplast genome (Fig. 1).

Removing regions of putative assembly errors, primer sequences and non-target amplicon sequences, in addition to exclusion of one copy of the inverted repeat (IR) region, generated an alignment of 130,493 base pairs. The number of variable sites was 4188 and that of parsimony-informative sites was 409.

3.2. Phylogenetic resolution in Fragaria

Phylogenetic analysis using ML (Fig. 2) identified and provided >98% support for two of the three diploid clades that have been observed in previous phylogenetic studies in *Fragaria* (Potter et al., 2000; Rousseau-Gueutin et al., 2009). Clade C, which contained *F. nipponica*, *F. pentaphylla*, *F. daltoniana*, and *F. chinensis,* was supported by a 100% bootstrap. Similarly high support (98%) was provided for clade A, which contained *F. vesca*, *F. mandschurica*, and *F. bucharica*. *Fragaria iinumae*, the only diploid in clade B separated from the other two clades (A and C) with a weak boostrap support (<50%). The placement of the two diploids, *F. viridis* and *F. nilgerrensis,* is unclear. There is however a strong sister relationship of *F. viridis* (80%) to clade A. *Fragaria viridis* and *F. nilgerrensis*, have either been unresolved or placed as sister to clades A and C in previous phylogenetic analyses (Harrison et al., 1997; Potter et al., 2000; Rousseau-Gueutin et al., 2009). In our analysis of complete genomes, *F. viridis* has moderate support as sister to clade A while the resolution of *F. nilgerrensis* remains uncertain (Fig. 2). Inclusion of more complete plastome sequences (i.e. obtained from direct genome sequencing) might help to resolve the relationships of these two species.

Within Clade A, several nodes are resolved with ML bootstrap support above 95%, including a subclade of all octoploid/decaploid taxa. Ours is the first phylogenetic estimate to observe monophyly in these taxa, and this result is consistent with a single origin to this lineage of polyploid species. As expected, *F.* ×*bifera* (*F. vesca* × *F. viridis*) (Staudt et al., 2003) was sister to its maternal parent *F. vesca* subsp. *vesca* (99% bootstrap support). *Fragaria vesca* subsp. *vesca* and *F. vesca* subsp. *americana* also resolved as sister taxa (99% bootstrap support), while *F. vesca* subsp. *bracteata* is sister taxon to the octoploids and the sole decaploid, indicating they share a common cytoplasmic ancestor. Previous studies of the *F. vesca* group included only three of the four subspecies (*F. vesca* subsp. *californica*, *F. vesca* subsp. *americana* and *F. vesca* subsp. *vesca*), and a close

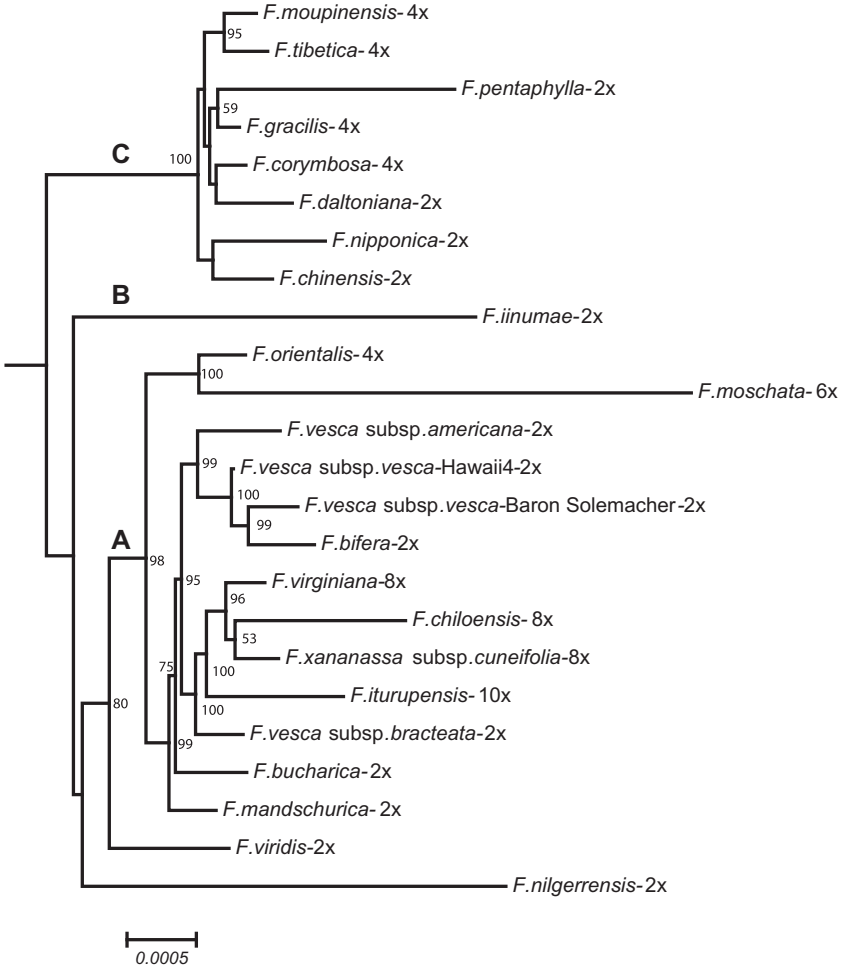

Fig. 2. Phylogenetic relationships of *Fragaria* based on ML analyses of plastomes. Only bootstraps above 50% are shown. Clades A–C refer to classifications from previous phylogenetic studies (Potter et al., 2000; Harrison et al., 1997; Rousseau-Gueutin et al., 2009).

Table 3
List of accessions used for Sanger sequencing of three parsimony-informative SNPs shared by the sole representative of each of *F. vesca* subsp. *bracteata* and the octoploid and decaploid species from Illumina sequencing (see Table 1). CFRA indicates NCGR local accession number.

PI	Taxon	Origin	Ploidy
551851	*F. bucharica*	Pakistan	2
551853	*F. bucharica*	Pakistan	2
616583	*F. chinensis*	China	2
CFRA1911	*F. corymbosa*	China	4
641195	*F. daltoniana*	China	2
1140518	*F. gracilis*	China	4
637963	*F. iinumae*	Japan	2
CFRA1947	*F. mandshurica*	Mongolia	2
–	*F. vesca* subsp. *bracteata*	Mexico	2
551528	*F. moschata*	France	6
CFRA1974	*F. moupinensis*	–	4
616672	*F. nilgerenssis*	China	2
637975	*F. nipponica*	Japan	2
616857	*F. viridis*	Sweden	2
616613	*F. ×bifera*	France	2
637933	*F. orientalis*	Russian Fed	4
651568	*F. pentaphylla*	China	2
CFRA1908	*F. tibetica*	China	4
551507	*F. vesca* subsp. *vesca*	Germany	2
551881	*F. vesca* subsp. *americana*	S. Dakota	2
651550	*F. vesca* subsp. *bracteata*	Idaho	2
551646	*F. vesca* subsp. *bracteata*	Idaho	2
657860	*F. vesca* subsp. *bracteata*	Colorado	2
652552	*P. villosa*	US	
551807	*F. vesca* subsp. *bracteata*	California	2
551835	*F. vesca* subsp. *bracteata*	Oregon	2
616651	*F. vesca* subsp. *bracteata*	British Columbia	2
551749	*F. vesca* subsp. *californica*	California	2
612492	*F. virginiana*	Canada	8
551805	*F. ×ananassa*	California	8
612318	*F. chiloensis*	Ecuador	8
641091	*F. iturupensis*	Russian Fed	10

relationship was observed among them (Potter et al., 2000; Rousseau-Gueutin et al., 2009). Using the nuclear *GBSSI*-2 (Rousseau-Gueutin et al., 2009) and nr*ITS* (Potter et al., 2000), European *F. vesca* subsp. *vesca* was differentiated from American subsp. *californica* and subsp. *americana*, but they were unresolved relative to the octoploid and decaploid species.

Monophyly of clade C was strongly supported but the resolution among species within the clade was low, in agreement with previous phylogenetic analyses using both nuclear (Rousseau-Gueutin et al., 2009) and chloroplast (Potter et al., 2000) sequences. Diploid-tetraploid relationships of species within this clade [*F. pentaphylla-F. tibetica* (Staudt and Dickoré, 2001), *F. chinensis–F. gracilis*, or *F. chinensis–F. corymbosa* (Staudt, 2009)] previously proposed based on overlapping geographical distribution and similarities in a limited number of morphological traits were not supported in this study. Close relationships among clade C species may be explained by their common distribution in the Himalayan region of China and Tibet (Staudt, 2006; Staudt and Dickoré, 2001; Darrow, 1966). *Fragaria nipponica*, distributed in Japan, Sakhalin and the Kurils in Russia (Staudt and Olbricht, 2008), is the only species in this clade that is found outside the Himalayan region (Figs. 2 and 4).

Sanger sequencing of plastome regions [NADH dehydrogenase F (*ndh*F), c-type cytochrome synthesis (*ccs*A), and RNA polymerase C2 (*rpo*C2)] encompassing three parsimony-informative SNPs that were shared between *F. vesca* subsp. *bracteata* and the octoploid and decaploid species were validated in additional *F. vesca* subspecies accessions (Table 3). Two haplotypes were observed in *F. vesca* subsp. *bracteata* accessions. The octoploid haplotype occurred in accessions collected in California, Oregon and British Columbia and the non-octoploid haplotype was observed in accessions collected from Idaho and Colorado (Fig. 3). *Fragaria vesca* subsp. *californica* also had the octoploid haplotype. *Fragaria vesca* subsp. *bracteata* is distributed along the coastal and Cascade ranges from British Columbia to California, where the octoploids *F. virginiana* subsp. *virginiana*, subsp. *glauca* and subsp. *platypetala* and *F. chiloensis* subsp. *pacifica* and subsp. *lucida* are also found (Hummer et al., 2009). *Fragaria vesca* subsp. *californica* is found near the Pacific Ocean from southern Oregon to California. This geographic segregation of chloroplast haplotypes in the Pacific Coast diploid populations implies that they are in the maternal lineage of the octoploid strawberries.

3.3. Biogeographic analysis

Based on the biogeographic analysis (Fig. 4) the chloroplast donor of the clades A–C is hypothesized to have originated from East Asia. The clade containing the octoploids, the decaploid *F. iturupensis* and diploid *F. vesca* subsp. *bracteata* (Fig. 4) is hypothesized to have originated from North/South America. The decaploid is

Taxon- Ploidy	PI	Origin	SNP1	SNP2	SNP3
F. vesca subsp. *bracteata*-2x	616651	BC/Canada	TGTCG**G**GTTGC…	TTTTT**T**GTGCT…	TTAGG**G**TTTAT
F. vesca subsp. *bracteata*-2x	551807	California/USA	TGTCG**G**GTTGC…	TTTTT**T**GTGCT…	TTAGG**G**TTTAT
F. vesca subsp. *bracteata*-2x	551835	Oregon/ USA	TGTCG**G**GTTGC…	TTTTT**T**GTGCT…	TTAGG**G**TTTAT
F. vesca subsp. *californica*-2x	551749	California/USA	TGTCG**G**GTTGC…	TTTTT**T**GTGCT…	TTAGG**G**TTTAT
F. ×ananassa subsp. *cuneifolia*-8x	551805	California/ USA	TGTCG**G**GTTGC…	TTTTT**T**GTGCT…	TTAGG**G**TTTAT
F. chiloensis-8x	612318	Ecuador/S. America	TGTCG**G**GTTGC…	TTTTT**T**GTGCT…	TTAGG**G**TTTAT
F. virginiana-8x	612492	Ontario/Canada	TGTCG**G**GTTGC…	TTTTT**T**GTGCT…	TTAGG**G**TTTAT
F. iturupensis-10x	641091	Kurile Isl/Rus. Fed.	TGTCG**G**GTTGC…	TTTTT**T**GTGCT…	TTAGG**G**TTTAT
F. vesca subsp. *americana*-2x	551881	Dakota/USA	TGTCG**A**GTTGC…	TTTTT**A**GTGCT…	TTAGG**A**TTTAT
F. vesca subsp. *bracteata*-2x	551646	Idaho/ USA	TGTCG**A**GTTGC…	TTTTT**A**GTGCT…	TTAGG**A**TTTAT
F. vesca subsp. *bracteata*-2x	657860	Colorado/ USA	TGTCG**A**GTTGC…	TTTTT**A**GTGCT…	TTAGG**A**TTTAT
F. vesca subsp. *bracteata*-2x	651550	Idaho/USA	TGTCG**A**GTTGC…	TTTTT**A**GTGCT…	TTAGG**A**TTTAT
F. vesca subsp. *bracteata*-2x	-	Hidalgo/Mexico	TGTCG**A**GTTGC…	TTTTT**A**GTGCT…	TTAGG**A**TTTAT
F. vesca subsp. *vesca*-2x	551507	Germany/Europe	TGTCG**A**GTTGC…	TTTTT**A**GTGCT…	TTAGG**A**TTTAT
F. ×bifera-2x	616613	France/Europe	TGTCG**A**GTTGC…	TTTTT**A**GTGCT…	TTAGG**A**TTTAT
F. viridis	616857	Sweden/Europe	TGTCG**A**GTTGC…	TTTTT**A**GTGCT…	TTAGG**A**TTTAT

Fig. 3. Sequence alignment of three chloroplast genome regions [SNP1- RNA polymerase C2 (*rpo*C2), SNP2-NADH dehydrogenase F (*ndh*F), and SNP3-c-type cytochrome synthesis (*ccs*A)] encompassing three parsimony-informative SNPs shared between *F. vesca* subsp. *bracteata* and the octoploid and decaploid species (red). Other species not shown here collected from E. Asia (*F. bucharica*, *F. chinensis*, *F. corymbosa*, *F. daltoniana*, *F. gracilis*, *F. iinumae*, *F. moupinensis*, *F. nilgerenssis*, *F. nipponica*, *F. orientalis*, *F. pentaphylla*, *F. tibetica*, the sole hexaploid from Europe, *F. moschata*, and the outgroup *Potentilla villosa* did not contain the octoploid haplotype. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

restricted to a single island of the Kuril Archipelago, but it is speculated that it occurs on other islands of the Kurils (Staudt, 1989). Even though this species has a limited distribution it shares leaf characteristics, color and texture with *F. virginiana* (Staudt, 1989). The close relationship of *F. vesca* subsp. *bracteata* and subsp. *californica* with the decaploid and octoploid species distributed across the Pacific Ocean suggests a northwestern North American or Beringian location for the origin of the octoploids.

Geographical distributions coupled with morphological characters have in the past been used to draw phylogenetic relationships in *Fragaria*. For example, Staudt (2003) proposed *F. mandshurica* as the diploid ancestor to the supposed autotetraploid (AAAA) *F. orientalis*. An allotetraploid (A′A′A″A″) hypothesis for *F. orientalis* was deemed possible using low copy nuclear genes with A′ and A″ (Y′ and Y″, Rousseau-Gueutin et al., 2009) representing *F. vesca* and *F. mandschurica* (Rousseau-Gueutin et al., 2009). Strong support for a close relationship of this tetraploid to the sole hexaploid, *F. moschata* in this study (Fig. 2) substantiates the conclusions of Harrison et al. (1997) based on chloroplast RFLP fragments, that the two represent a polyploid series. While *F. mandschurica* might not be the maternal donor to *F. orientalis*, auto- or allo-tetraploid origins are possible. However a diploid progenitor for these two species, *F. orientalis* and *F. moschata*, could not be identified here.

The overlapping distributions in Europe of *F. viridis*, *F. vesca* and *F. moschata* suggested that the hexaploid is an allopolyploid of the two diploids (Staudt et al., 2003). In previous studies, *F. viridis* was implicated as the maternal donor. By comparison of two indels (or indel mapping) in the chloroplast spacers, *psb*J–*psb*F and *rps*18–*rpl*20, *F. viridis* was favored over *F. vesca* and *F. bucharica* as the maternal donor to *F. moschata* (Lin and Davis, 2000). *Fragaria viridis* and *F. moschata* accessions also contained similar chloroplast SSR haplotypes (Njuguna, 2010). However, this relationship between *F. viridis* and *F. moschata* was either not supported (based on plastid *trn*L/F and *trn*S/G sequences, Lundberg et al., 2009), or unresolved (using chloroplast *trn*L intron and *trn*L–*trn*F spacer, Potter et al., 2000) in other studies. Alignment of the *rps*18–*rpl*20 intergenic spacer obtained by chloroplast sequencing in the large number of *Fragaria* species (Njuguna, 2010) used in this study (as

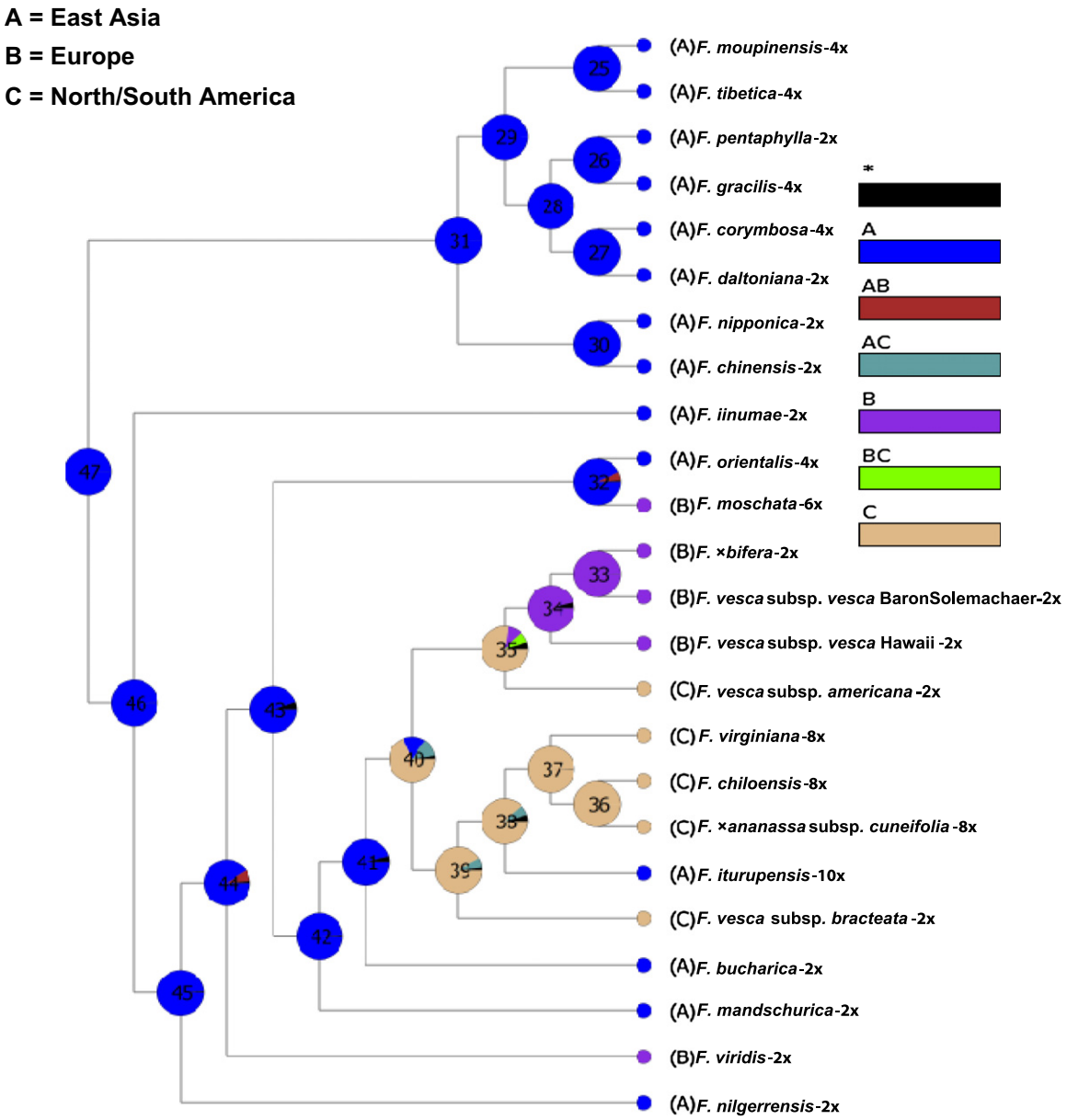

Fig. 4. Biogeographic reconstruction of *Fragaria*. Species were assigned one of three geographic regions: A = East Asia, B = Europe, C = North/South America.

opposed to its prior analysis that included four species; Lin and Davis, 2000) revealed that the 10 bp insertion observed in *F. moschata* and *F. viridis* was also found in diploids *F. iinumae*, *F. mandschurica* and *F. daltoniana,* tetraploids *F. gracilis*, *F. moupinensis*, *F. orientalis* and *F. tibetica*, octoploid *F. chiloensis* and decaploid *F. iturupensis*. Indel mapping can be useful in identifying species groups. However, without comprehensive taxon sampling, phylogenetic conclusions can be misleading; in addition, chloroplast rearrangements have been demonstrated to occur recurrently in independent lineages (Kelchner and Wendel, 1996). *Fragaria viridis* could be a possible progenitor of *F. moschata* but they did not form a clade (Fig. 2), suggesting that the actual maternal donor of the hexaploid and possibly the tetraploid *F. orientalis* remains unidentified.

3.4. Evolution of sex function

Character-state mapping provided insight into sexual system evolution and its association with loss of self-incompatibility and genome doubling/merger. Based on the conservative collapsed tree structure, several conclusions can be drawn. First, with respect to mating system (Fig. 5a), lack of resolution at the Clade C node makes it impossible to determine whether there were one, two, or three independent gains of gametophytic self-incompatibility (GSI) (*F. nipponica*, *F. chinensis*, and *F. pentaphylla*) from a self-compatible ancestor (proportional likelihood = 0.93). Thus the figure is consistent with as few as one and as many as three independent gains.

In Clade A, there are two equally parsimonious scenarios: (1) an ancient gain of GSI followed by two subsequent losses, one of which is associated with an increase in ploidy (the lineage to *F. moschata* and *F. orientalis*); and (2) three recent and independent gains of GSI in *F. viridis*, *F. mandshurica* and *F. bucharica*. While some argue that loss of SI is common but SI is seldom regained (Igic et al., 2008), transitions in both directions have been found in other groups (e.g., Ferrer and Good-Avila, 2007).

Second, with respect to male sterility, which also exists in both major clades, there are five gains (and one loss) or six gains of male sterility (Fig. 5b). In Clade C, lack of resolution prevents us from concluding one, two or three independent gains of male sterility (proportional likelihood = 0.95). In contrast, Clade A provides strong support for one gain of male sterility in the *F. moschata*/*F. orientalis* lineage (proportional likelihood for a sterile ancestor to both = 0.88), plus two nearly equally likely scenarios of character evolution in the octoploid clade. One scenario involves male sterility arising in a diploid ancestor early in the evolutionary history of the clade (proportional likelihood = 0.56) and then being lost in the decaploid *F. iturupensis*. Alternatively, male sterility arose independently in diploid *F. vesca* subsp. *bracteata* and in the ancestor that gave rise to the octoploids (proportional likelihood = 0.44). However, when additional evidence such as our current understanding of the dominance of male sterility (Ahmadi and Bringhurst, 1989; Goldberg et al., 2010; Spigler et al., 2010) is considered, the scales may be tipped in the direction of the former, as dominant male sterility is very rare in angiosperms (Kaul, 1988). Furthermore, there are no clear cases of gains of male sterility associated with increases in ploidy that implicated self-incompatible diploids. In the Asian clade, the current resolution of the tree does not provide support for a specific GSI ancestor of any of the tetraploids with male sterility (*F. corymbosa*, *F. moupinensis, F. gracilis* and *F. tibetica*), and the weight of the evidence is in favor of a SC ancestor of the octoploid clade (proportional likelihood = 0.95, Fig. 5a). In contrast, the gain of female sterility was accompanied by an increase in ploidy in all five cases (Fig. 5c) illustrating the evolution of the: octoploid Clade; *F. moschata* (6×)/*F. orientalis* (4×); *F. corymbosa* (4×), *F. gracilis* (4×), and *F. moupinensis* (4×)/*F. tibetica* (4×).

Fragaria, like several other genera (e.g., *Fuchsia* [Berry et al., 2004; Sytsma and Smith, 1991], *Lycium* [Miller, 2002], *Schiedia* [Weller et al., 1995], *Bryonia* [Volz and Renner, 2008], *Galium* [Soza and Olmstead, 2010] and *Thalictrum* [Soza et al., 2012]), shows that dimorphic sexual systems can evolve multiple times within a genus. But unlike *Lycium* [Yeung et al., 2005] and like *Bryonia* [Volz and Renner, 2008], our analysis of sexual system evolution in *Fragaria* does not support the hypothesis that loss of GSI with genome doubling is the underlying mechanism. In contrast, our results

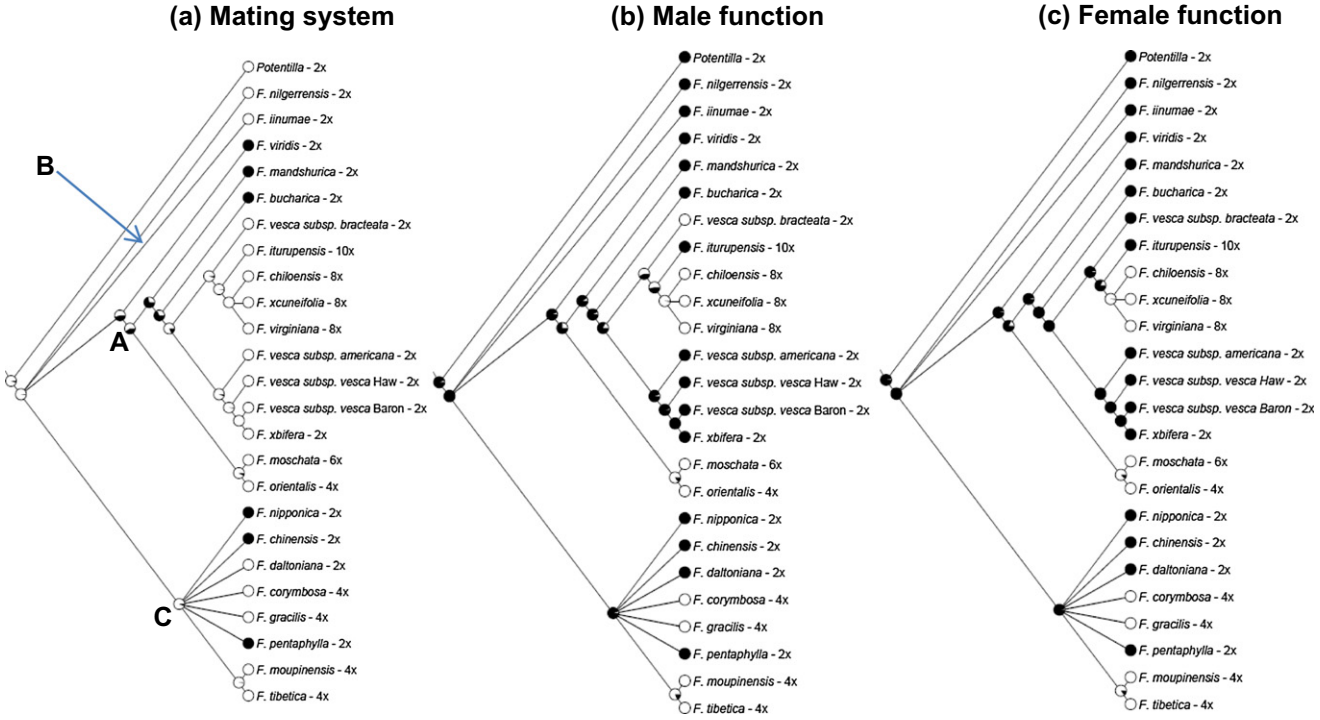

Fig. 5. Mating system and sexual system mapping across *Fragaria*. Mapping of (a) mating system (white: self- compatible, black: self incompatible); (b) male function (black: fertile, white: sterile); (c) female function (black: fertile, white: sterile). Maximum likelihood probabilities denoted at branch nodes.

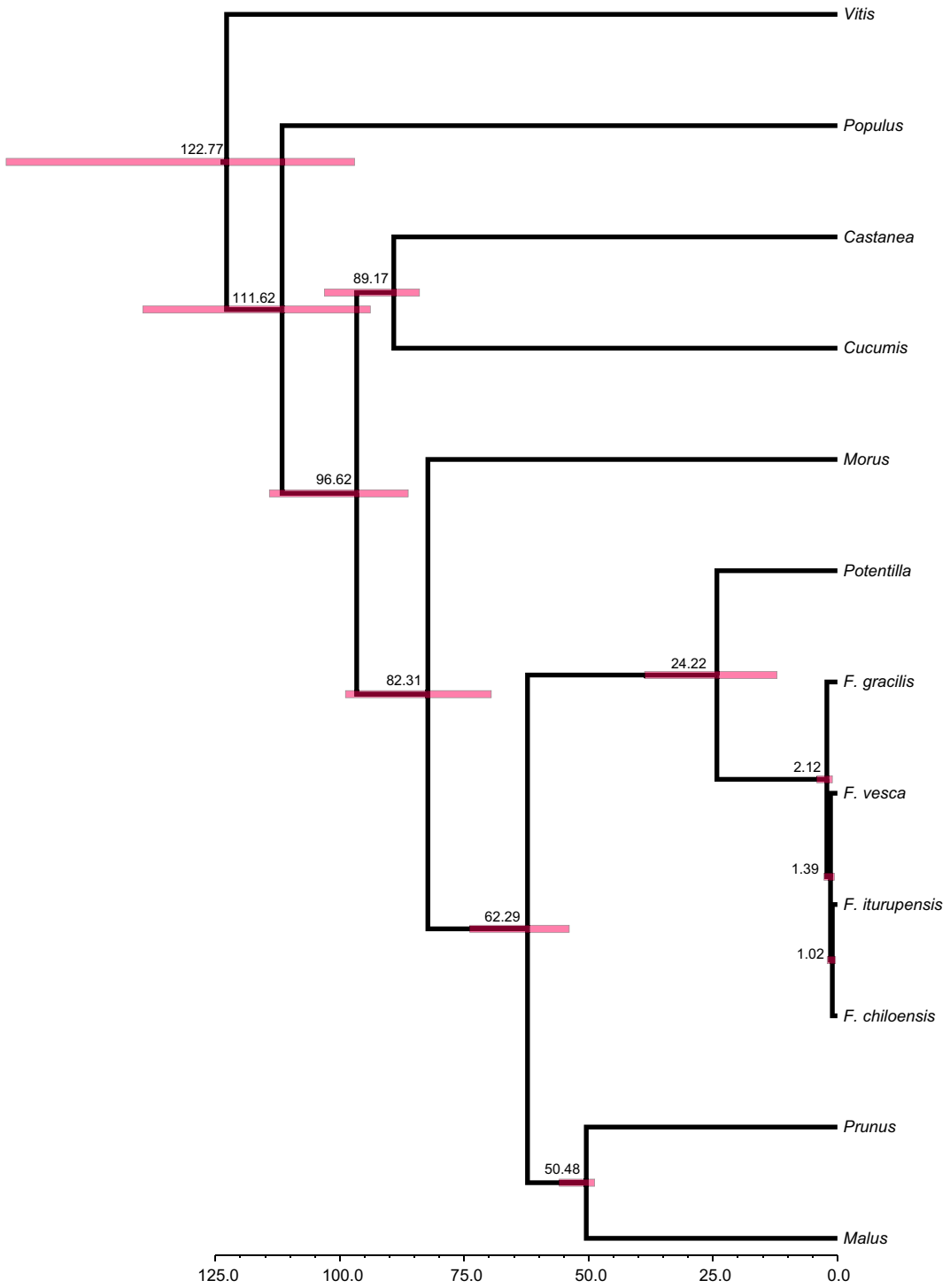

Fig. 6. Chronogram for *Fragaria*, based on 82 plastid loci, and calibrated with an 84.0–87.5 MYA divergence for *Castanea* and *Cucumis* and a 49.42 ± 0.54 MYA divergence for *Malus* and *Prunus*. Bars represent the 95% highest posterior density, and median ages (MY) are shown for each node. All posterior probabilities are 1.0.

support the idea that other changes associated with polyploidization are important. In fact, by mapping sex function (rather than sexual system), we provide evidence that loss of male and female function can follow polyploidization, perhaps as a result of gene silencing, methylation changes or gene loss which are common following polyploidization (reviewed in Pikaard (2001)). We must note, however, that since this reconstruction involves only maternal contributions, we cannot rule out the possibility that mutations causing loss of sex function were the result of hybridization in species thought to be allopolyploids, e.g., *F. virginiana*, *F. chiloensis*, *F. orientalis* and *F. moschata* (Bringhurst, 1990; Senanayake and Bringhurst, 1967).

One case of a reversal to hermaphroditism, involving *F. iturupensis*, is evident in our reconstruction of the *Fragaria* phylogeny and this may be associated with long-distance dispersal to the Iturup Islands (Hummer et al., 2009). Such a possibility is in agreement with Baker's law (Baker, 1967) which states that only single propagules that are self-fertile can initiate new populations. Such reversals have been found in other species (reviewed in Schaefer and Renner, 2010), but the association here with in-

Table 4
Fossil-calibrated relaxed molecular clock results for the genus *Fragaria* and the octoploid clade. TMRCA = time to most recent common ancestor. MYA = million years ago. HPD = highest posterior density.

Clade	Mean TMRCA (MYA)	Median TMRCA (MYA)	95% HPD (lower)	95% HPD (upper)	Effective sample size
Fragaria	2.326	2.119	1.017	4.139	762
Octoploids	1.116	1.021	0.372	2.05	1304

creased ploidy (10x) also highlights the fact that polyploidy can be associated with loss, as well as gain, of dioecy [e.g., *Bryonia* (Volz and Renner, 2008), *Mercurialis* (Obbard et al., 2006)]. Such losses may be associated with breakdown of nuclear sex determination (Westergaard, 1958) or complementarities associated with allopolyploidy (Gorelick, 2003), or whole chromosome loss in neopolyploids (see Zhiyong et al., 2011).

Fragaria vesca subsp. *bracteata* was implicated as the diploid contributor to the octoploids, and this is corroborated by character-state mapping of male sterility and functional analyses of gynodioecy (Li et al., 2012). This reconstruction predicts that male sterility in *F. virginiana* and *F. chiloensis* is homologous to that in *F. vesca* subsp. *bracteata* and that the female sterility, the second mutation required to create subdioecy or dioecy and a sex chromosome (reviewed in Charlesworth (2008)), arose only after polyploidization. Our result suggest that the second mutation, may have arisen as a result of direct genetic effects of genome doubling/merger (cited above) or those combined with selection imposed by invasion of novel habitats and ecological conditions. In fact, it is postulated that these two octoploid species diverged after migration across the Bering Strait to NW North America from Asia, where *F. chiloensis* spread southward and *F. virginiana* spread eastward and southward, adapting to dry (sand dunes, rocky cliff faces) and moister (meadows, disturbed areas) environments, respectively (Potter et al., 2000; Staudt, 1999). These results predict that female sterility should be homologous in *F. virginiana* and *F. chiloensis* and that evidence of different genome locations in the two species is the result of rearrangement (Goldberg et al., 2010). Tests of this prediction await characterization and comparison of DNA sequence of the sex determining regions in the two species.

3.5. Young age of Fragaria

The BEAST MCMC runs required 10%, 25% and 45% of the cycles to reach stationarity and convergence. After removal of pre-stationarity cycles, a total of 22 million cycles and 44,000 samples were summarized. Effective sample size (ESS) was >200 for all parameters except for the mean rate (ESS = 163.3) under the uncorrelated log-normal relaxed molecular clock. Based on these analyses, the most recent common ancestor of *Fragaria*¸ and the origin of the octoploid clade, are estimated to be of Pliocene–Pleistocene age (Fig. 6, Table 4). These results are comparable to an alternative estimate that used 1.35×10^{-9} substitutions/site/year (Wolfe et al., 1987) as an average rate of chloroplast sequence evolution, dating the origin of *Fragaria* to 1.52–4.44 MYA and the octoploids to 0.19–0.86 MYA (Njuguna, 2010). Preliminary analyses with all *Fragaria* sequences and with Fabaceae plastomes failed to reach convergence and gave much older ages for the origin of *Fragaria* (results not shown). These results were likely due to the bimodal distribution of speciation rates (*Fragaria* species vs. representative genera and families) and the elevated rate of sequence evolution observed in Fabaceae plastomes (Moore et al., 2010).

The accuracy of clade age estimates based on relaxed molecular clocks is highly dependent on the reliability of the calibration points (Clarke et al., 2011). We used two fossils that clearly possess synapomorphies of their respective clades as crown group calibrations. Although Rosaceae has a fossil record beginning in the early Eocene (Devore and Pigg, 2007), it is dominated by compressed leaves, which are not as reliable as flowers and fruit for attribution to genus. The recent documentation of flowers and fruits from two distinct lineages of Prunoideae (Benedict et al., 2011), combined with radiometric dating, provides a robust calibration for this subfamily, represented here by *Malus* and *Prunus* (Fig. 6).

The estimated ages for deep nodes (Fig. 6) have large intervals, and these are likely inflated by the presence of heterotachy (Wertheim et al., 2011). Our estimate for the origin of Rosaceae, 53.9–73.9 MYA, is somewhat younger than a previous estimate of 76 MYA based on an angiosperm-wide analysis (Wikström et al., 2001).The most recent common ancestor of *Potentilla* and *Fragaria* is estimated at 12.1–38.8 MYA in our study. This does not overlap with the 45.1–53.4 MYA interval estimated by Dobeš and Paule (2010). However, their calibration was based on a compression fossil assigned to *Rosa* in an unpublished master's thesis. The early Eocence origin of the genus *Rosa* has been discounted in a recent review of the fossil history of Rosaceae (Devore and Pigg, 2007); these authors consider the late Eocene or early Oligocene as the first reliable records of this genus. Dobeš and Paule (2010) conducted a second calibration using the Wikström et al. (2001) age for the origin of Rosaceae; this interval of 26.9–42.5 MYA is almost entirely within the range of our estimate.

Based on its circumboreal distribution, Staudt (1999) suggested a Cretaceous origin of the genus *Fragaria,* and he placed the divergence of *F. vesca* subspecies in the Eocene (Staudt, 1989). Our fossil calibrated relaxed molecular clock analysis clearly contradicts such an ancient origin of the genus (Fig. 6). In fact, the estimated age of the genus (1.0–4.1 MYA) places its origin in the Pliocene to Pleistocene (Table 4). If our predictions are accurate, this analysis identifies the octoploid species of *Fragaria* as one of the youngest lineages known to possess sex chromosomes in angiosperms. Assuming the sex chromosomes originated at the same time as the octoploid clade (0.37–2.05 MYA), they are similar in age to the homomorphic sex chromosomes of *Carica papaya* (~0.5–2.2 MYA) and muchyounger than the strongly heteromorphic sex chromosomes of *Silene latifolia* that are estimated to have originated 8–24 MYA (Ming et al., 2011). The error associated with these kinds of estimates is difficult to estimate, particularly given the incomplete nature of the fossil record. Irrespective of the actual age of *Fragaria* linages, the limited differentiation of the octoploid subspecies observed in genetic diversity studies of wild populations of *F. chiloensis* and *F. virginiana* (Hokanson et al., 2006), and the conservation of synteny observed in comparative linkage mapping (Rousseau-Gueutin et al., 2008; Sargent et al., 2009), are consistent with the estimated young age of this clade.

Acknowledgments

The authors would like to thank Theodore Bunch, Michael Dossett, Rajanikanth Govindarajulu, Brian Knaus, Stephen Meyers, Matthew Parks, Sushma Naithani, April Nyberg, and Sarah Sundholm for their invaluable technical support during this project. We would also like to acknowledge the OSU Center for Genome Research and Biotechnology for their invaluable help in sequencing (Mark Dasenko) and bioinformatic assistance and data transfer (Chris Sullivan). We also thank Manuel González Ledesma and Da-

vid Gernandt for providing specimen for F. vesca subsp. bracteata from Hidalgo, Mexico. Thanks also to Dr. Jim Hancock and Dr. Daniel Potter for thorough review of this manuscript. T.L.A. acknowledges NSF DEB 1020523 and A.L. acknowledges DEB 1020271 for logistical support.

References

Ahmadi, H., Bringhurst, R.S., 1989. Genetics of sex expression in *Fragaria* species. Am. J. Bot. 78, 504–514.

Baker, H.G., 1967. Support for Baker's Law-As a rule. Evolution 21, 853–856.

Baker, H.G., 1984. Some functions of dioecy in seed plants. Am. Nat. 124, 149–158.

Benedict, J.C., DeVore, M.L., Pigg, K.B., 2011. *Prunus* and *Oemleria* (Rosaceae) flowers from the Late Early Eocene Republic Flora of Northeastern Washington State, USA. Int. J. Plant Sci. 172, 948–958.

Berry, P.E., Hahn, W.J., Sytsma, K.J., Hall, J.C., Mast, A., 2004. Phylogenetic relationships and biogeography of *Fuchsia* (Onagraceae) based on noncoding nuclear and chloroplast DNA data. Am. J. Bot. 91, 601.

Bringhurst, R.S., 1990. Cytogenetics and evolution in American *Fragaria*. HortScience 25, 879–881.

Brunet, J., Liston, A., 2001. Polyploidy and gender dimorphism. Science 291, 1441a.

Charlesworth, D., 2008. Sex chromosome origins and evolution. In: Pagel, M., Pamiankowski, A. (Eds.), Evolutionary Genomics and Proteomics. Sinauer Associates Inc., Sunderland, MA, pp. 207–239.

Chung, S.M., Gordon, V.S., Staub, J.E., 2007. Sequencing cucumber (*Cucumis sativus* L.) chloroplast genomes identifies differences between chilling-tolerant and -susceptible cucumber lines. Genome 50, 215–225.

Clarke, J.T., Warnock, R.C.M., Donoghue, P.C.J., 2011. Establishing a time-scale for plant evolution. New Phytol. 192, 266–301.

Cronn, R., Liston, A., Parks, Gernandt, G.S., Shen, R., Mockler, T., 2008. Multiplex sequencing of plant chloroplast genomes using Solexa sequencing-by-synthesis technology. Nucleic Acids Res. 36, e122.

Darrow, G.M., 1966. The Strawberry: History Breeding and Physiology, first ed. New York.

DeVore, M., Pigg, K., 2007. A brief review of the fossil history of the family Rosaceae with a focus on the Eocene Okanogan Highlands of eastern Washington State, USA, and British Columbia, Canada. Plant Syst. Evol. 266, 45–57.

Dobeš, C., Paule, J., 2010. A comprehensive chloroplast DNA-based phylogeny of the genus *Potentilla* (Rosaceae): implications for its geographic origin, phylogeography and generic circumscription. Mol. Phylogenet. Evol. 56, 156–175.

Drummond, A.J., Ho, S.Y.W., Phillips, M.J., Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4, e88.

Eriksson, T., Hibbs, M.S., Yoder, A.D., Delwiche, C.F., Donoghue, M.J., 2003. The phylogeny of Rosoideae (Rosaceae) based on sequences of the internal transcribed spacers (*ITS*) of nuclear ribosomal DNA and the *trn*L/F region of chloroplast DNA. Int. J. Plant Sci. 164, 197–211.

Federova, M.J., 1946. Crossability and phylogenetic relations in the main European species of *Fragaria*. Compt. Rend. Acad. Sci. 52, 545–547.

Ferrer, M.M., Good-Avila, S.V., 2007. Macrophylogenetic analyses of the gain and loss of self-incompatibility in the Asteraceae. New Phytol. 173, 401–414.

Goldberg, M., Spigler, R., Ashman, T.-L., 2010. Comparative genetic mapping points to different sex chromosomes in sibling species of wild strawberry (*Fragaria*). Genetics 186, 1425–1433.

Gorelick, R., 2003. Evolution of dioecy and sex chromosomes via methylation driving Muller's ratchet. Biol. J. Linn. Soc. 80, 353–368.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.

Hancock, J.F., 1999. Strawberries. CAB International, Wallingford, UK.

Harrison, R.E., Luby, J.J., Furnier, G.R., 1997. Chloroplast DNA restriction fragment variation among strawberry (*Fragaria* spp.) taxa. J. Am. Soc. Hortic. Sci. 122, 63–68.

Hokanson, K.E., Smith, M.J., Connor, A.M., Luby, J.J., Hancock, J.F., 2006. Relationships among subspecies of New World octoploid strawberry species *Fragaria virginiana* and *Fragaria* chiloensis based on simple sequence repeat marker analysis. Can. J. Bot. 84, 1829–1841.

Hummer, K., Hancock, J., 2009. Strawberry genomics: botanical history, cultivation, traditional breeding, and new technologies. In: Folta, K.M., Gardiner, S.E. (Eds.), Plant Genetics and Genomics. Crops and Models. Springer, New York, pp. 413–436.

Hummer, K., Nathewet, P., Yanagi, T., 2009. Decaploidy in *Fragaria iturupensis* Staudt (Rosaceae). Am. J. Bot. 96, 713–716.

Igic, B., Lande, R., Kohn, J.R., 2008. Loss of self-incompatibility and its evolutionary consequences. Int. J. Plant. Sci. 169, 93–104.

Jennings, D.L., 1976. Raspberries and blackberries. In: Simmonds, N.W. (Ed.), Evolution of Crop Plants. Longman Group Ltd., New York, pp. 251–254.

Kaul, M.L.H., 1988. Male Sterility in Higher Plants. Springer-Verlag, Berlin.

Kelchner, S.A., Wendel, W.F., 1996. Hairpins create minute inversions in non-coding regions of chloroplast DNA. Curr Genet. 30, 259–262.

Li, J., Koski, M.H., Ashman, T.-L., 2012. Functional characterization of gynodioecy in *Fragaria vesca* ssp. *bracteata*. Ann. Bot. 109, 545–552.

Lin, J., Davis, T.M., 2000. S1 analysis of long PCR heteroduplexes: detection of chloroplast indel polymorphisms in *Fragaria*. Theor. Appl. Genet. 101, 415–420.

Lundberg, M., Töpel, M., Eriksen, B., Nylander, J.A.A., Eriksson, T., 2009. Allopolyploidy in Fragariinae (Rosaceae): comparing four DNA sequence regions with comments on classification. Mol. Phylogenet. Evol. 51, 269–280.

Maddison, W., Maddison, D., 2009. Mesquite: A Modular System for Evolutionary Analysis. <http://www.mesquiteproject.org/>.

Mayor, C., Brudno, M., Schwartz, J.R., Poliakov, A., Rubin, E.M., Frazer, K.A., Pachter, L.S., Dubchak, I., 2000. VISTA: visual global DNA sequence alignments of arbitrary length. Bioinformatics 16, 1046.

Miller, J., 2002. Phylogenetic relationships and the evolution of gender dimorphism in Lycium (Solanaceae). Syst. Bot. 27, 416–428.

Miller, J.S., Venable, D.L., 2000. Polyploidy and the evolution of gender dimorphism in plants. Science 289, 2335–2538.

Ming, R., Bendahmane, A., Renner, S.S., 2011. Sex chromosomes in land plants. Annu. Rev. Plant. Biol. 62, 485–514.

Moore, M.J., Soltis, P.S., Bell, C.D., Burleigh, J.G., Soltis, D.E., 2010. Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. Proc. Natl. Acad. Sci. USA 107, 4623–4628.

Njuguna, W., 2010. Development and Use of Molecular Tools in *Fragaria* L. PhD Dissertation. Oregon State University, Corvallis, Oregon.

Obbard, D., Harris, S., Buggs, R.J., Pannell, J.R., 2006. Hybridization, polyploidy, and the evolution of sexual systems in Mercurialis (Euphorbiaceae). Evolution 60, 1801–1815.

Osborn, T.C., Butrulle, D.V., Sharpe, A.G., Pickering, K.G., Parkin, I.A.P., Parker, J.S., Lydiate, D.J., 2003. Detection and effects of a homeologous reciprocal transposition in *Brassica napus*. Genetics 165, 1569–1577.

Ovcharenko, I., Giardine, B.M., Hou, M., Ma, J., Hardison, R.C., Stubbs, L., Miller, W., 2005. Mulan: multiple-sequence local alignment and visualization for studying function and evolution. Genome Res. 15, 184–194.

Parks, M., Cronn, R., Liston, A., 2009. Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. BMC Biol. 7, 84.

Pikaard, C.S., 2001. Genomic change and gene silencing in polyploids. Trends Genet. 17, 675–677.

Pond, S.K., Muse, S.V., 2005. Site-to-site variation of synonymous substitution rates. Mol. Biol. Evol. 22, 2375–2385.

Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 9, 817–818.

Potter, D., Luby, J.J., Harrison, R.E., 2000. Phylogenetic relationships among species of *Fragaria* (Rosaceae) inferred from non-coding nuclear and chloroplast DNA sequences. Syst. Bot. 25, 337–348.

Ratan, A., 2009. Assembly Algorithms for Next Generation Sequence Data. PhD Dissertation. Pennsylvania State University.

Ravi, V., Khurana, J., Tyagi, A., Khurana, P., 2006. The chloroplast genome of mulberry: complete nucleotide sequence gene organization and comparative analysis. Tree Genet. Genomes 3, 49–59.

Rousseau-Gueutin, M.A., Lerceteau-Kohler, E., Barrot, L., Sargent, D.J., Monfort, A., Simpson, D., Arus, P., Guerin, G., Denoyes-Rothan, B., 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.

Rousseau-Gueutin, M.A., Gaston, A., Aïnouche, A., Aïnouche, M.L., Olbricht, K., Staudt, G., Richard, L., Denoyes-Rothan, B., 2009. Tracking the evolutionary history of polyploidy in *Fragaria* L. (strawberry): new insights from phylogenetic analyses of low-copy nuclear genes. Mol. Phylogenet. Evol. 51, 515–530.

Rozen, S., Skaletsky, H.J., 2000. Primer3 on the WWW for general users and for biologist programmers. In: Krawetz, S., Misener, S. (Eds.), Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa New Jersey, pp. 365–386.

Sargent, D.J., Fernandéz-Fernandéz, F., Ruiz-Roja, J.J., Sutherland, B.G., Passey, A., Whitehouse, A.B., Simpson, D.W., 2009. A genetic linkage map of the cultivated strawberry (*Fragaria* ×*ananassa*) and its comparison to the diploid *Fragaria* reference map. Mol. Breed. 24, 293–303.

Schaefer, H., Renner, S.S., 2010. A three-genome phylogeny of *Momordica* (Cucurbitaceae) suggests seven returns from dioecy to monoecy and recent long-distance dispersal to Asia. Mol. Phylogenet. Evol. 54, 553–560.

Senanayake, Y.D.A., Bringhurst, R.S., 1967. Origin of *Fragaria* polyploids I. Cytological analysis. Am. J. Bot. 51, 221–228.

Shulaev, V., Sargent, D.J., Crowhurst, R.N., Mockler, T.C., Folkerts, O., Delcher, A.L., Jaiswal, P., Mockaitis, K., Liston, A., Mane, S.P., Burns, P., Davis, T.M., Slovin, J.P., Bassil, N., Hellens, R.P., Evans, C., Harkins, T., Kodira, C., Desany, B., Crasta, O.R., Jensen, R.V., Allan, A.C., Michael, T.P., Setubal, J.C., Celton, J.-M., Rees, D., Jasper, G., Williams, K.P., Holt, S.H., Rojas, J.J.R., Chatterjee, M., Liu, B., Silva, H., Meisel, L., Adato, A., Filichkin, S.A., Troggio, M., Viola, R., Ashman, T.-L., Wang, H., Dharmawardhana, P., Elser, J., Raja, R., Priest, H., Bryant, D.W., Fox, S.E., Givan, S.A., Wilhelm, L.J., Naithani, S., Christoffels, A., Salama, D.Y., Carter, J., Girona, E.L., Zdepski, A., Wang, W., Kerstetter, R.A., Schwab, W., Korban, S.S., Davik, J., Monfort, A., Denoyes-Rothan, B., Arus, P., Mittler, R., Flinn, B., Aharoni, A., Bennetzen, J.L., Salzberg, S.L., Dickerman, A.W., Velasco, R., Borodovsky, M., Veilleux, R.E., Folta, K.M., 2010. The genome of woodland strawberry (*Fragaria vesca*). Nat. Genet. 43, 109–116.

Sims, H.J., Herendeen, P.S., Lupia, R., Christopher, R.A., Crane, P.R., 1999. Fossil flowers with Normapolles pollen from the Upper Cretaceous of southeastern North America. Rev. Palaeobot. Palynol. 106, 131–151.

Soza, V.L., Olmstead, R.G., 2010. Evolution of breeding systems and fruits in New World Galium and relatives (Rubiaceae). Am. J. Bot. 97, 1630.

Soza, V.L., Brunet, J., Liston, A.I., Smill, P.S., Stilio, V.S., 2012. Phylogenetic insights into the correlates of dioecy in meadow-rues (*Thalictrum*, Ranunculaceae). Mol. Phylogenet. Evol.. http://dx.doi.org/10.1016/j.ympev.2012.01.009.

Spigler, R.B., Lewers, K.S., Main, D.S., Ashman, T.-L., 2008. Genetic mapping of sex determination in a wild strawberry, *Fragaria virginiana*, reveals earliest form of sex chromosome. Heredity 101, 507–517.

Spigler, R.B., Lewers, K.S., Johnson, A.L., Ashman, T.-L., 2010. Comparative mapping reveals autosomal origin of sex chromosome in octoploid *Fragaria virginiana*. J. Hered. 2010, 101.

Stamatakis, A., Hoover, P., Rougemont, J., 2008. A fast bootstrapping algorithm for the RAxML Web-Servers. Syst. Biol. 57, 758–771.

Staudt, G., 1967. The genetics and evolution of heteroecy in the genus Fragaria. Investigations on *Fragaria orientalis*. Z. Pflanzenzuchtung 58, 245–277.

Staudt, G., 1968. The genetics and evolution of heteroecy in the genus *Fragaria*. III. Investigations of hexaploid and octoploid species. Z. Pflanzenzuchtg. 59, 83–102.

Staudt, G., 1989. The species of *Fragaria*, their taxonomic and geographical distribution. Acta Hort. 265, 23–33.

Staudt, G., 1999. Systematics and geographic distribution of the American strawberry species. Univ. Calif. Publ. 81, 1–162.

Staudt, G., 2003. Notes on Asiatic *Fragaria* species: III *Fragaria orientalis* Losinsk and *Fragaria mandushurica* spec nov. Bot. Jahrb. Syst. 124, 397–419.

Staudt, G., 2006. Himalayan species of *Fragaria* (Rosaceae). Bot. Jahrb. Syst. 126, 483–508.

Staudt, G., 2009. Strawberry biogeography genetics and systematic. In: Proceedings of the 6th International Strawberry Symposium, vol. 842, pp. 71–84.

Staudt, G., Dickoré, W.B., 2001. Notes on Asiatic *Fragaria* species: *Fragaria pentaphylla* Losinsk and *Fragaria tibetica* spec nov. Bot. Jahrb. Syst. 123, 341–354.

Staudt, G., Olbricht, K., 2008. Notes on Asiatic *Fragaria* species V: *F nipponica* and *F iturupensis*. Bot. Jahrb. Syst. 127, 317–341.

Staudt, G., DiMeglio, L.M., Davis, T.M., Gerstberger, P., 2003. *Fragaria* ×*bifera* Duch: origin and taxonomy. Bot. Jahrb. Syst. 125, 53–72.

Straub, S.C.K., Fishbein, M., Livshultz, T., Foster, Z., Parks, M., Weitemier, K., Cronn, R.C., Liston, A., 2011. Building a model: developing genomic resources for common milkweed (*Asclepias syriaca*) with low coverage genome sequencing. BMC Genomics 12, 211.

Straub, S.C.K., Parks, M., Weitemier, K., Fishbein, M., Cronn, R.C., Liston, A., 2012. Navigating the tip of the genomic iceberg: next-generation sequencing for plant systematics. Am. J. Bot. 99, 349–364.

Sytsma, K., Smith, J.F., 1991. The use of chloroplast DNA to assess biogeography and evolution of morphology, breeding systems, and flavonoids in *Fuchsia* sect. *Skinnera* (Onagraceae). Sys. Bot. 16, 257–269.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4. Mol. Biol. Evol. 24, 1596–1599.

Velasco, R., Zharkikh, A., Affourtit, J., Dhingra, A., Cestaro, A., Kalyanaraman, A., Fontana, P., Bhatnagar, S.K., Troggio, M., Pruss, D., Salvi, S., Pindo, M., Baldi, P., Castelletti, S., Cavaiuolo, M., Coppola, G., Costa, F., Cova, V., Dal, R.A., Goremykin, V., Komjanc, M., Longhi, S., Magnago, P., Malacarne, G., Malnoy, M., Micheletti, D., Moretto, M., Perazzolli, M., Si-Ammour, A., Vezzulli, S., Zini, E., Eldredge, G., Fitzgerald, L., Gutin, N., Lanchbury, J., Macalma, T., Mitchell, J., Reid, J., Wardell, B., Kodira, C., Chen, Z., Desany, B., Niazi, F., Palmer, M., Koepke, T., Jiwan, D., Schaeffer, S., Krishnan, V., Wu, C., Chu, V., King, S.T., Vick, J., Tao, Q., Mraz, A., Stormo, A., Stormo, K., Bogden, R., Ederle, D., Stella, A., Vecchietti, A., Kater, M.M., Masiero, S., Lasserre, P., Lespinasse, Y., Allan, A.C., Bus, V., Chagne, D., Crowhurst, R.N., Gleave, A.P., Lavezzo, E., Fawcett, J.A., Proost, S., Rouze, P., Sterck, L., Toppo, S., Lazzari, B., Hellens, R., Durel, C.-E., Gutin, A., Bumgarner, R.E., Gardiner, S.E., Skolnick, M., Egholm, M., Van de Peer, Y., Salamini, F., Viola, R., 2010. The genome of the domesticated apple (*Malus* ×*domestica* Borkh.). Nat. Genet. 42, 833–839.

Volz, S.M., Renner, S.S., 2008. Hybridization, polyploidy, and evolutionary transitions between monoecy and dioecy in *Bryonia* (Cucurbitaceae). Am. J. Bot. 95, 1297–1306.

Weller, S.G., Wagner, W.L., Sakai, A.K., 1995. A phyologenetic analysis of *Schediea* and *Alsinidendron* (Caryophyllaceae: Alsinoideae): implications for the evolution of breeding systems. Syst. Bot. 20, 315–337.

Wertheim, J.O., Fourment, M., Pond, S.L.K., 2011. Inconsistencies in estimating the age of HIV-1 subtypes due to heterotachy. Mol. Biol. Evol.. http://dx.doi.org/10.1093/molbev/msr266.

Westergaard, M., 1958. The mechanism of sex determination in dioecious flowering plants. Adv. Genet. 9, 217–281.

Wikström, N., Savolainen, V., Chase, M.W., 2001. Evolution of the angiosperms: calibrating the family tree. Proc. R. Soc. Lond. B. 268, 2211–2220.

Wolfe, K.H., Li, W.H., Sharp, P.M., 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc. Nat. Acad. Sci. U.S.A. 84, 9054–9058.

Yeung, K., Miller, J.S., Savage, A.E., Husband, B.C., Igic, C., Koh, J.R., 2005. Association of ploidy and sexual sstem in *Lycium Californicum* (Solanaceae). Evolution 59, 2048–2055.

Yu, Y., Harris, A.J., He, X.J., 2011. RASP (Reconstruct Ancestral State in Phylogenies). Version 2.0.1.0. <http://www.mnh.scu.edu.cn/soft/blog/RASP>.

Zhiyong, X.R.T., Gaeta, Pires J.C., 2011. Homoeologous shuffling and chromosome compensation maintain genome balance in resynthesized allopolyploid *Brassica napus*. Proc. Natl. Acad. Sci. USA. http://dx.doi.org/10.1073/pnas.1014138108.