Multilocus Sex Determination Revealed in Two Populations of Gynodioecious Wild Strawberry, Fragaria vesca subsp bracteata

Multilocus Sex Determination Revealed in Two Populations of Gynodioecious Wild Strawberry, *Fragaria vesca* subsp. *bracteata*

Tia-Lynn Ashman,*,[1] Jacob A. Tennessen,[†] Rebecca M. Dalton,* Rajanikanth Govindarajulu,*
Matthew H. Koski,* and Aaron Liston[‡]

*Department of Biological Sciences, University of Pittsburgh, Pennsylvania 15260, [†]Department of Integrative Biology,
and [‡]Department of Botany and Plant Pathology, Oregon State University, Corvallis, Oregon 97331

ABSTRACT Gynodioecy, the coexistence of females and hermaphrodites, occurs in 20% of angiosperm families and often enables transitions between hermaphroditism and dioecy. Clarifying mechanisms of sex determination in gynodioecious species can thus illuminate sexual system evolution. Genetic determination of gynodioecy, however, can be complex and is not fully characterized in any wild species. We used targeted sequence capture to genetically map a novel nuclear contributor to male sterility in a self-pollinated hermaphrodite of *Fragaria vesca* subsp. *bracteata* from the southern portion of its range. To understand its interaction with another identified locus and possibly additional loci, we performed crosses within and between two populations separated by 2000 km, phenotyped the progeny and sequenced candidate markers at both sex-determining loci. The newly mapped locus contains a high density of pentatricopeptide repeat genes, a class commonly involved in restoration of fertility caused by cytoplasmic male sterility. Examination of all crosses revealed three unlinked epistatically interacting loci that determine sexual phenotype and vary in frequency between populations. *Fragaria vesca* subsp. *bracteata* represents the first wild gynodioecious species with genomic evidence of both cytoplasmic and nuclear genes in sex determination. We propose a model for the interactions between these loci and new hypotheses for the evolution of sex determining chromosomes in the subdioecious and dioecious *Fragaria*.

Diversity in sexual systems and sex determination is a hallmark of plants (Bachtrog *et al.* 2014; Renner 2014). In angiosperms in particular, this variety is indicative of the myriad ways unisexual individuals can evolve from combined sex phenotypes (hermaphrodites or cosexuals) (Diggle *et al.* 2011; Renner 2014). One canonical pathway to entirely separate sexes (dioecy) involves an intermediate sexual system known as gynodioecy (females and hermaphrodites) (Charlesworth and Charlesworth 1978), which is found in nearly 20% of families and 2% of genera of

flowering plants (Dufaÿ *et al.* 2014). Gynodioecy, however, does not always lead to dioecy and may be a stable sexual system in its own right (Lewis 1941; Charlesworth and Charlesworth 1978; Dufaÿ *et al.* 2007), or may transition back to hermaphroditism (Delph *et al.* 2007; Goldberg *et al.* unpublished results). The likelihood of these transitions is influenced by the underlying genetics, among other factors such as mating system and frequency dependent selection (Ehlers and Bataillon 2007; Crossman and Charlesworth 2014).

Characterizing the mechanisms of sex determination in gynodioecious species can be key for understanding sexual system transitions (Charlesworth and Charlesworth 1978; Schultz 1994; Maurice and Fleming 1995; Bailey and Delph 2007), as well as for elucidating aspects of male and female reproductive developmental pathways (Wang *et al.* 2013; Diggle *et al.* 2011). Known sex determination systems can involve cytoplasmic and/or nuclear genes. In cyto-nuclear gynodioecy, mitochondrial male-sterility mutations [cytoplasmic male sterility (CMS)] are counteracted by alleles at nuclear 'restorer' (*Rf*) loci (*e.g.*, Dufaÿ *et al.* 2009). Nuclear sex determination can result when CMS is fixed and only restorers segregate (Klaas and Olson 2006), or when nuclear mutations cause loss-of-function (LOF) in the pollen

development pathway irrespective of cytoplasmic genes (Bailey and Delph 2007). Finally, in some species both cytoplasmic and nuclear mutations segregate within populations and there can be more than one nuclear locus (*e.g.*, Van Damme *et al.* 2004; Garraud *et al.* 2011). Reversions to hermaphroditism can occur within populations when CMS or nuclear LOF genes are lost or when both CMS and restorers become fixed (*e.g.*, *Mimulus gutattus*, Barr and Fishman 2010). As a consequence, disparate populations of a widespread species that have been exposed to dissimilar selection regimes or differentially subjected to drift (population size or founder events) can be differentiated for sex determination mechanisms. Such variation has been revealed in species with cyto-nuclear gynodioecy (*e.g.*, *Lobelia siphiltica*, Dudle *et al.* 2001), but also can be seen in species with nuclear sex determination or sex chromosomes (*e.g.*, Dufresnes *et al.* 2014). In these cases, transitions in sex determining regions (or chromosomes) are thought to be driven by additional selective forces, such as sexually antagonistic selection, meiotic drivers and genetic load (*e.g.*, Blaser *et al.* 2014; Ubeda *et al.* 2015).

Model organisms and agriculturally important taxa have been instrumental in characterizing sex determination pathways and have revealed that transitions between sexual systems can involve new genes, new alleles or entirely novel means of sex determination (*e.g.*, Vicoso and Bachtrog 2013; Akagi *et al.* 2014). For instance while restorers of CMS have yet to be genetically characterized in wild gynodioecious species, those identified from crop hybrids are commonly pentatricopeptide repeat (PPR) genes that act in dominant fashion so only one copy is required to restore pollen function (sterility at *Rf* loci is typically recessive) (reviewed by Chen and Liu 2014). Nevertheless, it has been noted that the same mechanism can be regulated by different genes, and thus developmental networks can evolve rapidly (Cui *et al.* 2012; Wang *et al.* 2013). Understanding sex determination in wild plants is particularly valuable because it is in these species that data can be linked to ecological processes that are prevailing drivers of sexual system variation (Frank 1989; Jacobs and Wade 2003). Gynodioecious species are often closely related to hermaphroditic or dioecious ones (Dufaÿ *et al.* 2014), so characterization of their sex determination can also provide insight into genetic, developmental and evolutionary dynamics of sexual systems (Spigler and Ashman 2012; Russell and Pannell 2015).

We explored variation in sex determination in the widespread North American diploid strawberry *Fragaria vesca* subsp. *bracteata* for two main reasons. First, it is gynodioecious and the maternal donor to the sexually dimorphic octoploid *Fragaria* (dioecious *F. chiloensis* and subdioecious *F. virginiana*) (Njuguna *et al.* 2013; Tennessen *et al.* 2014; Govindarajulu *et al.* 2015), which were hybridized to produce the cultivated hermaphroditic *Fragaria×ananassa*. As such it represents both a tractable and evolutionarily appropriate genetic model for teasing apart sex differentiation in this crop and its wild relatives (Liston *et al.* 2014). Second, several lines of evidence allude to variation in sex determining regions within the species. Specifically, although previous genetic mapping of sex determination in a northern population of *F. vesca* subsp. *bracteata* identified a nuclear locus with a dominant allele for male sterility (Tennessen *et al.* 2013), there is equivocal evidence for a fitness advantage sufficient to maintain females in this population under nuclear determination alone (Li *et al.* 2012; Dalton *et al.* 2013). Moreover, recent phylogeographic studies revealed genetic differentiation across the range of *F. vesca* subsp. *bracteata*, with one combination of chlorotype and mitotype more closely related to octoploids than other(s) (Njuguna *et al.* 2013; Govindarajulu *et al.* 2015; Stanley *et al.* 2015). Geographic differentiation was partly due to variation in a novel mitochondrial open reading frame (ORF) with sequence similarity to known CMS genes, although a sterilizing function

has not been confirmed (Govindarajulu *et al.* 2015; Stanley *et al.* 2015). These points combined with the fact that the location of the sex determining region of *F. vesca* subsp. *bracteata* identified in Tennessen *et al.* (2013) was on a chromosome that is different from either of the chromosomal locations of the sex determining region in the two octoploid congeners (*i.e.*, chromosome 4 (LG4) *vs.* a chromosome in homeologous group VI (LG VI-Av or VI-B2), Goldberg *et al.* 2010; Spigler *et al.* 2011; Tennessen *et al.* 2014) raises the possibility of additional genetic contributors to sex determination in other populations of the diploid species.

We used targeted-sequence capture (Tennessen *et al.* 2013, 2014) to efficiently fine map the sex determination region using a self-pollinated hermaphrodite plant from a population at the southeastern edge of *F. vesca* subsp. *bracteata*'s range within the USA. From this we identified a second locus affecting sexual phenotype on a different chromosome (LG6) than that previously identified from a female collected in the northwestern portion of the range (LG4; Tennessen *et al.* 2013). We conducted intra- and inter-population crosses. In addition to phenotyping the progeny, we sequenced informative single nucleotide polymorphisms (SNPs) near the target sex-determining regions to determine whether the second locus interacted with the locus previously identified. In doing, so we reveal interactions of at least two nuclear loci and population variation in sex determination. From this we propose a novel mechanistic pathway for sex determination in the diploid species, and extend it as a hypothesis for sex determination in the octoploid species that descended from the diploid species studied here.

MATERIALS AND METHODS

Study system
F. vesca subsp. *bracteata* (Rosaceae) is a diploid ($2n = 2\times = 14$) North American wild strawberry that inhabits moderately damp soils of shady forest edges, meadows, and riverbanks. Its range extends from British Columbia to southern Mexico and from the Pacific coast to the Rocky Mountains and Sierra Madre (Staudt 1999; Stanley *et al.* 2015). Variation in cytoplasmic haplotypes has been revealed across this range with western populations dominated by one chlorotype and those from the Rockies and Sierra Madre by another with a zone of admixture in between (Njuguna *et al.* 2013; Stanley *et al.* 2015). In addition, there is finer scale and more complex pattern of mitochondrial variation (Stanley *et al.* 2015).

The species has a gynodioecious sexual system (Ahmadi and Bringhurst 1989; Li *et al.* 2012) and populations vary in the frequency of females from 0% (no females) to 46% female (Stanley *et al.* 2015). Sex types produce similar numbers of flowers, seeds and plantlets on stolons (Li *et al.* 2012). Hermaphrodites are self-compatible, highly selfing and morphologically distinguishable from females by their slightly longer stamens and the presence of viable pollen grains in their anthers (Li *et al.* 2012; Dalton *et al.* 2013).

The only known nuclear locus that affects sex expression was identified from a cross between a female and a hermaphrodite from Oregon (OR-MRD; Tennessen *et al.* 2013). The dominant allele for male sterility resides in a 338-kb region of chromosome 4 that is predicted to house 57 genes, although none in protein families known to control male sterility were found based on the reference genome *F. vesca* Hawaii4 annotations (Tennessen *et al.* 2013). However, Govindarajulu *et al.* (2015) identified a mitochondrial ORF (*atp8-orf225*) with molecular characteristics similar to known CMS genes. It is present and variable in *F. vesca* subsp. *bracteata* as well as other taxa across the genus *Fragaria* (Govindarajulu *et al.* 2015; Stanley *et al.* 2015), but its functional role has yet to be confirmed.
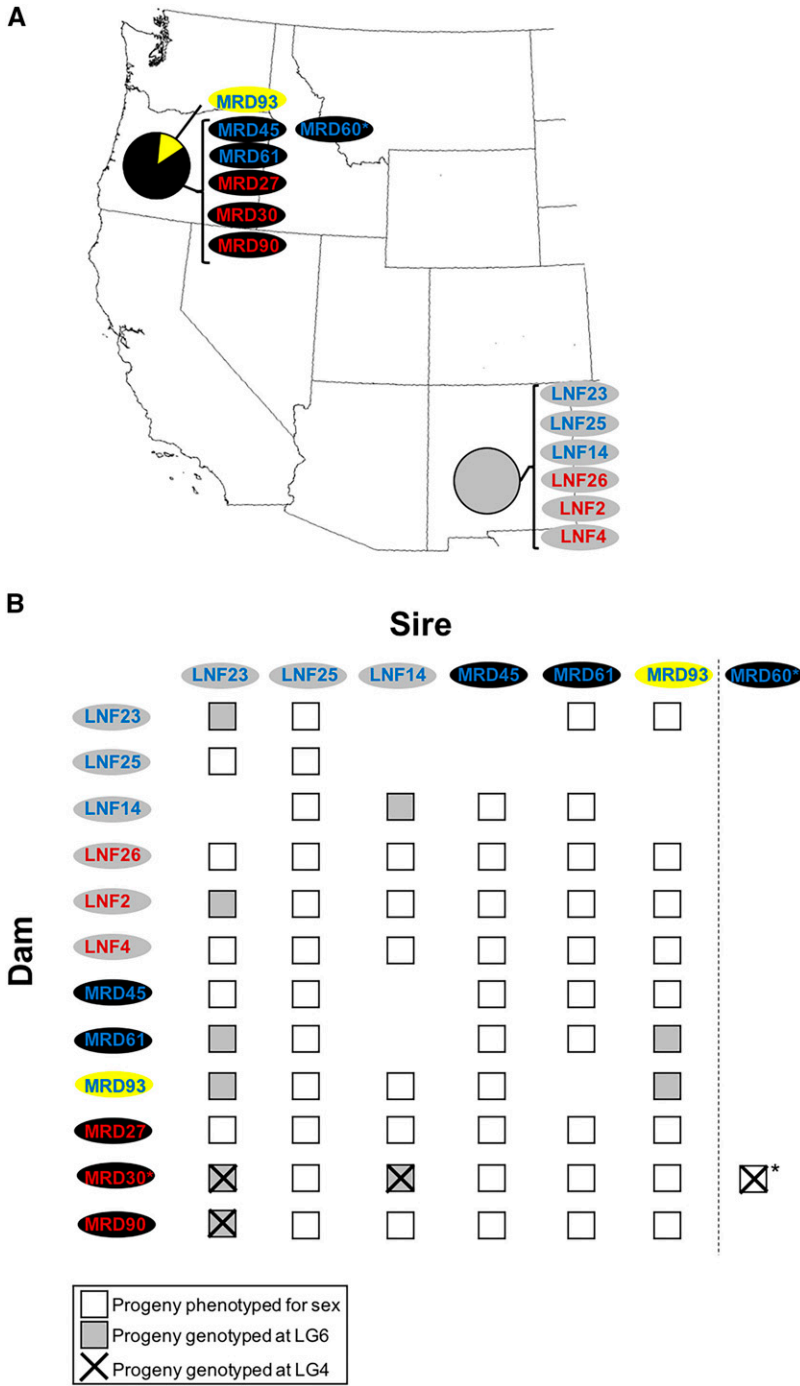
Figure 1 (A) Geographic source location of *F. vesca* subsp. *bracteata* plants used in the study and (B) the crossing design, phenotyping and genotyping strategy. Color of plant identity reflects the plant's sexual phenotype (blue = hermaphrodite; red = female). In addition, mitochondrial haplotype of each parent plant at the putative cytoplasmic male sterility (CMS) gene is represented by the oval color (yellow = B; black = C, and gray = F; see Table S1). Plants denoted with asterisk were also used as parents in the study of Tennessen *et al.* 2013. All crosses of sires × dams were performed but only those with progeny surviving for scoring phenotypic sex are represented with a box. Those genotyped at LG6 or LG4 loci are denoted by shading or 'X'. The single cross of Tennessen *et al.* 2013 is also represented.

Study materials

F. vesca subsp. *bracteata* from two populations separated by 2000 km were the focus of the current study. NM-LNF is a high elevation 'sky island' population in the southern portion of the range (in the Lincoln National Forest, New Mexico, altitude 2664 m; N 32° 58′ 6.1′′, W 105° 44′ 44.7′′). OR-MRD is on the highest peak in the Oregon Coast Range (Marys Peak in the Siuslaw National Forest, Oregon, altitude 521 m; N 44.8° 29′ 18.4′′, W 123° 32′ 14.7′′). The latter population is also within the zone of cytoplasmic introgression seen in *F. vesca* subsp. *bracteata* between the Coast Range / Cascades and Northern Rocky Mountains (Staudt 1999), although no individuals of OR-MRD showed admixture in their nuclear genomes (Stanley *et al.* 2015). The frequency of females in OR-MRD is 30% and in NM-LNF is 46% (Stanley *et al.* 2015).

Geography, flower and fruit morphology, and sexual system indicate both populations are *F. vesca* subsp. *bracteata* (Staudt 1999; T.-L. Ashman, unpublished). Likewise phylogenomic analysis shows these populations form a clade for the majority of their chromosomes (Tennessen *et al.* 2014) and are always separate from *F. vesca* subsp. *vesca* Hawaii 4 (source of the published genome sequence, Shulaev *et al.* 2010) which itself forms a clade with *F. vesca* subsp. *americana* (Njuguna *et al.* 2013). These two populations, however, differed in their cytotypes as is common in *F. vesca* subsp. *bracteata*: OR-MRD has one chlorotype (2) and two mitotypes (90% C and 10% B), whereas

■ Table 1 Phenotypic sex ratios from intra- (within population) and inter-population (between populations) crosses of gynodioeicous *F. vesca* subsp. *bracteata*

Cross Type	Dam	Sire	Total Progeny Sexed	Female	Hermaphrodite	Expected Ratio	χ^2 Value	P Value
A. Within NM-LNF Crosses								
H-SELF	LNF23	LNF23	53	9	44	1:3	2.70	0.10
	LNF25	LNF25	53	17	36	1:3	1.42	0.23
	LNF14	LNF14	30	15	15	1:3[a]	4.6	0.03
F x H	LNF26	LNF23	12	4	8	1:1	1.33	0.25
	LNF26	LNF25	50	29	21	1:1	1.28	0.26
	LNF26	LNF14	10	6	4	1:1	0.40	0.53
	LNF2	LNF23	20	11	9	1:1	0.20	0.65
	LNF2	LNF25	42	20	22	1:1	0.10	0.76
	LNF2	LNF14	17	7	10	1:1	0.53	0.47
	LNF4	LNF23	40	17	23	1:1	0.90	0.34
	LNF4	LNF25	29	11	18	1:1	1.69	0.19
	LNF4	LNF14	7	3	4	1:1	0.14	0.71
H x H	LNF23	LNF25	4	1	3	1:3	NA	NA
	LNF25	LNF23	22	5	17	1:3	0.06	0.80
	LNF14	LNF23	27	5	22	1:3	0.61	0.43
	LNF14	LNF25	2	1	1	1:3	NA	NA
	LNF25	LNF14	0*	NA	NA	NA	NA	NA
	LNF23	LNF14	0*	NA	NA	NA	NA	NA
B. Within OR-MRD Crosses								
H-SELF	MRD45	MRD45	16	3	13	1:3	0.33	0.56
	MRD61	MRD61	43	0	43	1:3[a]	14.33	0.00
	MRD93	MRD93	39	8	31	1:3	0.42	0.52
F x H	MRD27	MRD45	18	10	8	1:1	0.22	0.64
	MRD27	MRD61	31	11	20	1:1	2.61	0.11
	MRD27	MRD93	19	11	8	1:1	0.47	0.49
	MRD30	MRD45	11	5	6	1:1	0.09	0.76
	MRD30	MRD61	23	11	12	1:1	0.04	0.83
	MRD30	MRD93	19	8	11	1:1	0.47	0.49
	MRD90	MRD45	25	13	12	1:1	0.04	0.84
	MRD90	MRD61	28	10	18	1:1	2.29	0.13
	MRD90	MRD93	38	15	23	1:1	1.68	0.19
H x H	MRD45	MRD61	11	0	11	0:1	0.00	NA
	MRD61	MRD45	14	1	13	0:1	0.00	NA
	MRD61	MRD93	20	0	20	0:1	0.00	NA
	MRD93	MRD61	0*	NA	NA	NA	NA	NA
	MRD45	MRD93	7	0	7	1:3	2.33	NA
	MRD93	MRD45	4	0	4	1:3	NA	NA
C. Between NM-LNF and OR-MRD								
F x H	MRD27	LNF23	17	13	4	1:1[a]	4.76	0.03
	MRD27	LNF25	35	18	17	1:1	0.03	0.87
	MRD27	LNF14	14	10	4	1:1	2.57	0.11
	MRD30	LNF23	26	15	11	1:1	0.62	0.43
	MRD30	LNF25	33	20	13	1:1	1.48	0.22
	MRD30	LNF14	16	9	7	1:1	0.25	0.62
	MRD90	LNF23	44	24	20	1:1	0.36	0.55
	MRD90	LNF25	46	20	26	1:1	0.78	0.38
	MRD90	LNF14	39	17	22	1:1	0.64	0.42
	LNF26	MRD45	13	0	13	0:1	0.00	NA
	LNF26	MRD61	15	0	15	0:1	0.00	NA
	LNF26	MRD93	21	15	6	1:1[a]	3.86	0.05
	LNF2	MRD45	18	0	18	0:1	0.00	NA
	LNF2	MRD61	5	0	5	0:1	0.00	NA
	LNF2	MRD93	28	21	7	1:1[a]	7.00	0.01
	LNF4	MRD45	14	0	14	0:1	0.00	NA
	LNF4	MRD61	15	0	15	0:1	0.00	NA
	LNF4	MRD93	52	35	17	1:1[a]	6.23	0.01
H x H	MRD45	LNF23	29	0	29	0:1	0.00	NA
	MRD45	LNF25	17	0	18	0:1	0.00	NA
	MRD45	LNF14	0*	NA	NA	NA	NA	NA
	MRD61	LNF23	25	0	25	0:1	0.00	NA

(continued)

Cross Type	Dam	Sire	Total Progeny Sexed	Female	Hermaphrodite	Expected Ratio	χ^2 Value	P Value
	MRD61	LNF25	15	0	15	0:1	0.00	NA
	MRD61	LNF14	NA	NA	NA	NA	NA	NA
	MRD93	LNF23	27	6	21	1:3	0.11	0.08
	MRD93	LNF25	3	0	3	1:3	NA	NA
	MRD93	LNF14	0*	NA	NA	NA	NA	NA
	LNF23	MRD45	0*	NA	NA	NA	NA	NA
	LNF23	MRD61	1	0	1	0:1	NA	NA
	LNF23	MRD93	4	2	2	1:3	NA	NA
	LNF25	MRD45	0*	NA	NA	NA	NA	NA
	LNF25	MRD61	0*	NA	NA	NA	NA	NA
	LNF25	MRD93	0*	NA	NA	NA	NA	NA
	LNF14	MRD45	4	2	2	1:3	NA	NA
	LNF14	MRD61	3	0	3	0:1	NA	NA
	LNF14	MRD93	0*	NA	NA	NA	NA	NA

Three types of crosses are presented: selfed hermaphrodites (H-self), female dam by hermaphrodite sire (F x H), and hermaphrodite dam by hermaphrodite sire (H x H) from within (A) N-LNF or (B) OR-MRD populations, or between NM-LNF and OR-MRD populations (C). For each cross the dam, sire, and total number of progeny scored for sexual phenotype are given, as well as the number of female and hermaphrodite progeny, the predicted sex ratio based on Table 3 genotypes of the parents. χ^2 statistics and *P* values from χ^2 goodness of fit tests. *Crosses performed yielded no seed or none germinated. NA, Not applicable [too few progeny (family size <10) to conduct a statistical test].
a Alternative hypothesis for expected sex ratio tested and reported in text

NM-LNF is fixed for different chlorotype and mitotype (1 and F, respectively; Stanley *et al.* 2015).

Crosses and sexual phenotyping

Plants were collected from the wild and maintained in the greenhouse at the University of Pittsburgh. Three hermaphrodites (OR-MRD61, OR-MRD45, OR-MRD93) and three females [OR-MRD30 (used in Tennessen *et al.* 2013), OR-MRD27, OR-MRD90] from OR-MRD were used as parents. From NM-LNF, three hermaphrodites [NM-LNF23 (used in Tennessen *et al.* 2014), NM-LNF25, and NM-LNF14] and three females (NM-LNF2, NM-LNF4, NM-LNF26) were used as parents. Cytotypes of the parents reflected their abundance in wild populations (Stanley *et al.* 2015). These are represented in Figure 1A and details are given in Supporting Information, Table S1. Both parents used in the mapping study of Tennessen *et al.* (2013) were the majority mitotype in OR-MRD ('C') and are also represented in Figure 1A. The three *atp8-orf225* variants translate to different proteins, although their function is unknown (Stanley *et al.* 2015).

All females were crossed with pollen from both inter- and intra-population hermaphrodites, and all hermaphrodites were self-pollinated. In addition, crosses between the hermaphrodites were performed within and between populations. Crosses were performed by hand-pollinating one to four flowers per dam with pollen collected from single sires. When hermaphrodites served as dams their flowers were emasculated prior to pollination. Seeds were harvested and stored at −20° until sowing. Some crosses produced few seeds (and/or few germinated), notably hermaphrodite by hermaphrodite inter-population crosses involving NM-LNF as the dam and OR-MRD as the sire, and some hermaphrodite by hermaphrodite within population crosses (see *Results*, Table 1). We sowed an average of 31 seeds (range 1–56) per family into 98-well trays in a mix of Fafard #4, sand, and Sunshine Germination Mix. Germination was, on average, high (75%), although some specific crosses performed poorly (see *Results*). Approximately 1 month after germination we transplanted progeny into 150cc pots of soil (two parts Fafard #4 to 1 part sand) and randomized them on benches in a greenhouse where we maintained conditions of 14–15 hr days 21°/15.5° (day/night) temperatures.

To induce flowering we subjected plants to a cold treatment (8-hr days at 8°/16 hr dark at 4°) in a growth chamber for 3–4 weeks, and then returned them to the greenhouse where they received 50ppm 10:30:20 N:P:K fertilizer (Peter's Professional Bloom Booster) weekly. This was repeated up to five times to maximize flowering. A high rate of flowering was achieved (on average 86% progeny per family). Upon flowering we scored male function for at least two flowers per plant. Given the weak sexual dimorphism in this species (Li *et al.* 2012), we recorded both whether or not the anthers were shedding pollen in the greenhouse and whether viable pollen grains were produced. We confirmed pollen production and viability microscopically by fixing anthers with Alexander's stain (Kearns and Inouye 1993), and later observing them under a light microscope. Plants with no viable pollen production were scored as sterile whereas those with viable pollen production were scored as fertile. In total, 1364 plants flowered and the number of individuals phenotyped per family ranged from 1 to 53 (mean = 21). Progeny sex ratios were scored from flowering plants in each cross and for family size >10 sex ratio was tested for deviations from expected (see below) using χ^2 tests.

Genetic and genomic analysis

To identify the sex-determining locus in NM-LNF we used the linkage map published in Tennessen *et al.* (2014). This map was created using OneMap (Margarido *et al.* 2007) and the 1825 polymorphic sites identified from the target captured sequences of 41 progeny of self-pollinated NM-LNF23 (Tennessen *et al.* 2014). This map had the expected seven linkage groups spanning a total of 326 cM and is described in full in Tennessen *et al.* (2014) where it was used to produce a new *F. vesca* genome assembly ('Fvb'). Based on past experience (Tennessen *et al.* 2013) this sample size is sufficient to identify genomic regions of major effect on male function. We coded male sterility as a recessive Mendelian locus (*e.g.*, sterile *rr* and fertile *R*–; justified because a hermaphroditic parent could not harbor a dominant male sterility allele), and mapped it along with the targeted capture genotypes using OneMap (Margarido *et al.* 2007) to determine its position on the linkage map. From this map we identified a candidate region in coupling with male sterility on LG6 (see *Results*, Figure 2).

Once the male function region was identified, we designed primers for PCR amplification and Sanger sequencing of three SNPs from target capture sequences on LG6: one just upstream (Fvb6_34763k), one within (Fvb6_35142k), and one downstream (Fvb6_36607k) of the male sterility region (Table S2; Figure 3). Primers were designed with Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) and positioned to amplify additional nearby SNPs segregating in OR-MRD30, as well as the NM-LNF23 SNPs. To genotype the region housing the sex expression locus on LG4, we used primers described previously using the *F. vesca* Hawaii 4 assembly but that were realigned to the new assembly 'Fvb' (Fvb4_30092k) (Tennessen *et al.* 2014; Table S2).

Sequencing of targeted sex determining regions on LG4 and LG6 was performed on seven parents and 301 progeny from 13 select crosses (Figure 1B). The crosses were selected based on the availability of segregating informative SNPs in the targeted regions in the parents (Table 2I) and sexual phenotypes of their progeny. We sequenced progeny from three crosses for LG4 markers (OR-MRD30×NM-LNF23, OR-MRD30×NM-LNF14, and OR-MRD90×NM-LNF23) and progeny from 10 crosses for LG6 markers (OR-MRD93self, NM-LNF14self, NM-LNF2×NM-LNF23, NM-LNF2×OR-MRD93, OR-MRD30×NM-LNF23, OR-MRD30×NM-LNF14, OR-MRD90× NM-LNF23, OR-MRD93×NM-LNF23, OR-MRD61×NM-LNF23, and OR-MRD61×OR-MRD93).

For each cross we extracted DNA from 100 to 150 mg of fresh or 20–30 mg of silica-dried leaf tissue per plant. We used a CTAB DNA protocol modified as in Govindarajulu *et al.* (2013) and PCR amplification was performed with 1X standard reaction buffer (New England Biolabs), 100 μM of each dNTP, 0.5 μM of each forward and reverse primer, 1.5 units of Standard *Taq* polymerase (New England Biolabs) and 1.5 μl of genomic DNA in a 20-μl reaction. The amplification conditions were: 2.5 min at 95°, followed by 35 cycles of 95° for 30 sec, 55 and 56° for 30 sec, and 72° for 60 sec, and a final extension at 72° for 8 min (Table S2). The amplified products were Sanger sequenced and aligned using Sequencer ver 4.8 (Gene Codes Corp, Ann Arbor, MI) and SNPs were scored. For all crosses, we analyzed all segregating SNPs observed within each sequenced PCR product, whether or not they were the same segregating SNPs observed in the linkage maps. Association between segregating markers and sexual phenotypes was assessed in the progeny sets using Fisher's exact tests.

We identified potential candidate genes in the region matching male sterility using the Hybrid GeneMark Predictions (Shulaev *et al.* 2010) available from GDR: Genome Database for Rosaceae (https://www. rosaceae.org/species/fragaria/fragaria_vesca/genome_v1.1) supplemented with four previously unrecognized genes (Darwish *et al.* 2015). For functional annotation we relied on PLAZA 3.0 (Proost *et al.* 2015). We examined levels of gene expression in a floral development transcriptome (Hollender *et al.* 2014), especially at stages 8–10, which are important in pollen development (Hollender *et al.* 2012). To test for PPR enrichment in our genomic region of interest, we identified 653 genes annotated as "Pentatricopeptide repeat-containing protein" (PPR) and counted them in 1 Mb nonoverlapping windows across the genome. We also examined the genomic location of the 98 PPR genes that have a putative ortholog (as determined by PLAZA 3.0) with the 26 fertility restorer-like (RFL) genes in *Arabidopsis thaliana* (Fujii *et al.* 2011).

Data availability

Illumina sequencing data in FASTQ format have been described previously (Tennessen *et al.* 2014) and are uploaded to the NCBI SRA (Bioproject Accession PRJNA263688). Primers for additional genotyping are listed in Table S2. Sequences of all marker regions can be
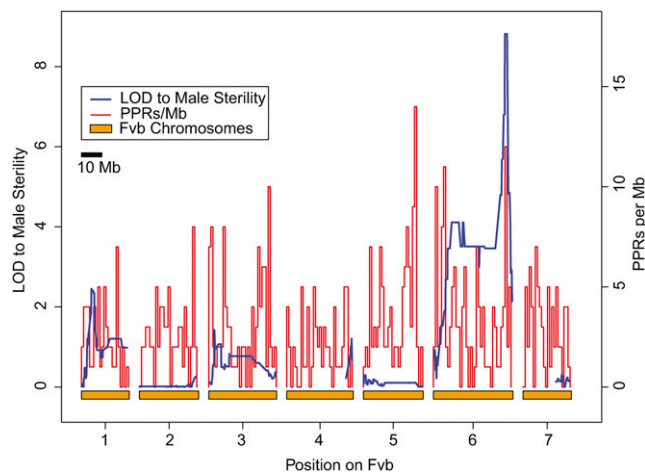
Figure 2 Map of male sterility in genome of hermaphrodite *F. vesca* subsp. *bracteata* from New Mexico (NM-LNF23). The seven chromosomes based on the Fvb reference genome (Tennessen *et al.* 2014) are denoted by orange bars along the *x*-axis and LOD scores associated with male function (blue line; left hand *y*-axis) and pentatricopeptide repeat (PPR) gene density (PPR/Mb, red line; right hand *y*-axis) on the *y*-axes. A significant LOD score (>3) only occurs on LG 6, peaking at 8.8 for markers between 35.0 and 36.0 Mb. This region overlaps one of the two densest clusters of PPR genes in the genome, the other occurring on LG5.

determined from the *F. vesca* reference genome v. 2.0 ("Fvb", https:// www.rosaceae.org/species/fragaria_vesca/genome_v2.0.a1) and the genotypes listed in Table 2.

RESULTS

Novel sex determining region discovered

Of the 53 total selfed progeny of NM-LNF23, nine were female and 44 were hermaphrodite (1:3 ratio; $\chi^2 = 2.7$; $P = 0.10$; Table 1A). Of the 41 offspring that were part of the mapping population, eight were female and 33 were hermaphrodite (1:3 ratio; $\chi^2 = 0.07$; $P = 0.79$). Using these data we unambiguously mapped male sterility in these plants to a region near the 3′ end of chromosome LG6 (Figure 2). Specifically, at 14 sites on 10 targeted sequence capture probes between Fvb6_34958975 and Fvb6_36048692 (Figure 3A; Table 2I), we observed a perfect match to male function. These perfectly matching sites include the region Fvb6_35142k (also Sanger genotyped in other crosses; Table 2II-A), at which two SNPs (Fvb6_35142280 = Sanger site 280; and Fvb6_ 35142453 = Sanger site 453) cosegregated perfectly with sex type ($P = 0.0001$). At these two sites, all eight females were homozygous for one of the two parental haplotypes ("G_T" at the Fvb6_35142k sites; Table 2I), while all 33 hermaphrodites were either heterozygous or homozygous for the other parental haplotype ("T_C or "G/T_C/T" at the Fvb6_35142k sites), consistent with recessive male sterility (LOD = 8.8; Figure 2). From this we infer the genotype of NM-LNF23 to be *Rr* at the male function locus and its female progeny to be *rr* and hermaphrodite progeny to be *RR* or *Rr* (Table 3). Adjacent to these markers, the nearest mismatching markers are upstream at Fvb6_34839229 and downstream at Fvb6_36607138, and thus the male sterility locus must occur in the 1.769 Mb region between these two markers (Figure 3A). The SNPs just outside this region, including in regions Fvb6_34763k and Fvb6_36607k (Table 2II-A) also segregated significantly with sex type ($P < 0.001$), although at Fvb6_34839229 and farther upstream, a single female mismatched, while at Fvb6_36607138 and farther downstream, one female and two hermaphrodites mismatched.
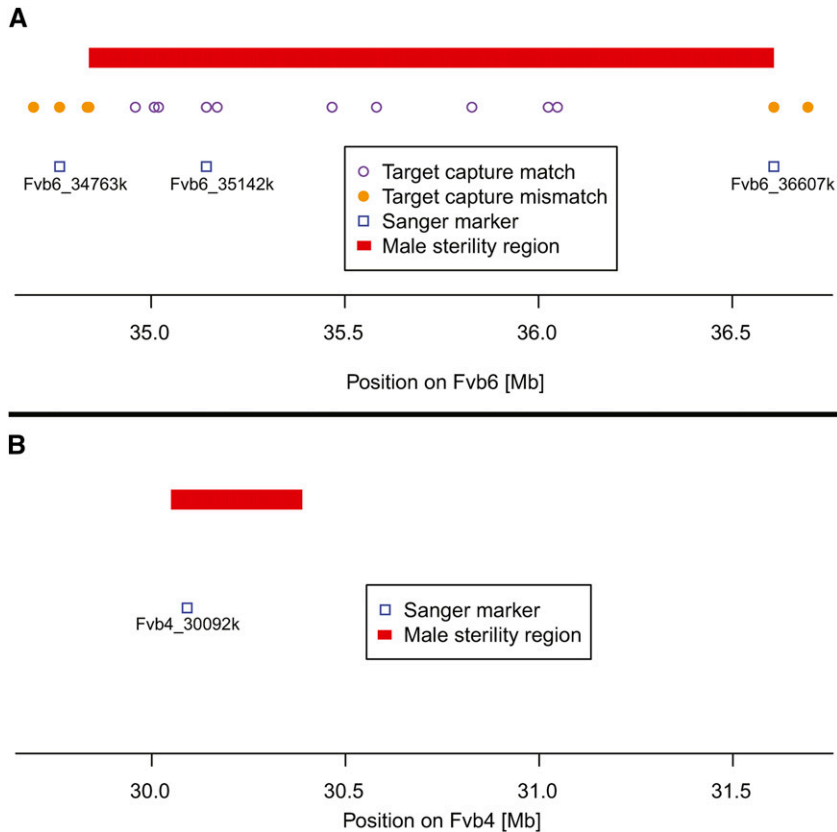
Figure 3 Male sterility genomic regions. (A) The LG6 locus (*R/r*). The male sterility region is defined as the span including ten targeted sequence markers that perfectly match male sterility (see text). Three markers that were genotyped to confirm location and explore segregation with male sterility in other crosses with Sanger sequencing are noted by blue boxes (Fvb6_34763k, Fvb6_35142k and Fvb6_36607k). (B) The LG4 (*MS/mf*) locus. The male sterility region was mapped in Tennessen *et al.* (2013). Locations of Sanger markers used in this study are indicated by a blue box (Fvb4_30092k).

The 1.768 Mb region on Fvb6 that shows a perfect match to male-sterility contains 361 genes in the *F. vesca* Hawaii 4 reference genome (Table S3). These include 182 that are upregulated in anthers, pollen or microspores, with one male gametophyte-specific gene, and several F-box proteins that have been seen to be upregulated in meiotic anthers at stage 9 (Hollender *et al.* 2014). Other gene classes showing expression changes in developing anthers included Kelch repeat and Leucine-Rich Repeat (LRR) proteins. This region also includes 15 PPR genes (Table S3), 12 of which occur in the 1 Mb span between Fvb6_35Mb and Fvb6_36Mb, and 10 of which occur in the 0.5 Mb span between Fvb6_35Mb and Fvb6_35.5Mb. This cluster of PPRs is unusually dense in this region relative to the rest of the genome (mean genomic PPR density = 2.9 per Mb). In fact, only one other genomic location, on Fvb5, contains a higher density cluster of PPRs (Figure 3). Although none of these PPRs are orthologs of the *Raphanus Rfo* fertility restorer and *Arabidopsis* RFL genes (Fujii *et al.* 2011), four are in the PLAZA gene family HOM03D000002, and one of these (gene04450) is predicted to be mitochondrial targeted (Table S4). This gene family contains the *Raphanus Rfo* gene and 25 of the 26 *Arabidopsis* RFL genes, as well as PPRs at two fertility-restorer loci recently identified in a hybrid cross almond × peach (Donoso *et al.* 2015). The RF1 locus on peach LG2 has one of four PPRs in this gene family, while the RF2 locus on peach LG6 has a dense cluster of 12 HOM03D000002 PPRs in 843.5 kbp. Five of these are considered orthologs of the *Arabidopsis* RFL genes (Table S4).

Sex expression in additional NM-LNF crosses and confirmation of LG6 region of influence

The progeny of selfed hermaphrodite NM-LNF25 showed a 1:3 female to hermaphrodite sex ratio ($\chi^2 = 1.42$; $P = 0.23$), which is consistent with heterozygosity at a male function locus (*Rr*) (Table 1A and Table 3).

The progeny of selfed hermaphrodite NM-LNF14, however, deviated significantly from a 1:3 female to hermaphrodite sex ratio ($\chi^2 = 4.6$; $P = 0.03$), leading us to propose a third, yet unmapped, locus affecting male function (locus LG*x*). The progeny sex ratio from selfed hermaphrodite NM-LNF14 is consistent with a 9:7 sex ratio ($\chi^2 = 0.47$; $P = 0.49$) that could result from selfing of a plant heterozygous at two male function loci (*Rr* and *Tt*), although one needs to acknowledge this is based on a small set of progeny (Table 1A and Table 3). Sequencing the progeny from this cross for the Fvb6_35142k locus and evaluating the positions 35142280 and 35142453, however, we find that of the 15 hermaphrodite offspring, all are either G, T or G/T, C/T at these two sites, but none of them are T_C. This makes sense because T (at 35142280) and C (at 35142453) are in coupling with *r* at the *R* locus, so all seven T_C offspring are *rr* homozygotes and therefore female, consistent with the proposed two locus model. The other female offspring are presumably *tt* at locus LG*x*. In conclusion, the fact that the *Rr* locus is correlated with the NM-LNF14 offspring phenotypes, but does not perfectly explain them, joins the phenotypic segregation results in supporting our conclusion that *Rr* and another locus (*Tt*) are acting jointly. All other parents from NM-LNF are proposed to be *TT* at this locus (Table 3).

Crosses involving the NM-LNF hermaphrodites as sires each with three NM-LNF females produced progeny sex ratios (1:1) consistent with hypothesized genotypes of *Rr TT* for two of the hermaphrodites, and *Rr Tt* for one (NM-LNF14) and females all *rr TT* (all $\chi^2 < 1.6$, $P > 0.19$; Table 1A). The reciprocal crosses between NM-LNF25 and NM-LNF23 also produced progeny sex ratios (1:3) consistent with hypothesized genotypes of *Rr TT* for both of these hermaphrodites, though one of these crosses had very low seed set ($P > 0.80$; Table 1A and Table 3). The reciprocal crosses of NM-LNF14 with NM-LNF23 and NM-LNF25 also produced few seeds. The one cross that produced sufficient

■ **Table 2 Single nucleotide polymorphic sites segregating at the putative sex determining regions on linkage groups 6 and 4 (LG6 and LG4) (see Figure 3, A and B) in two populations (OR-MRD, NM-LNF) of gynodioecious *F. vesca* subsp. *bracteata***

I. Parents

| | | LG6 | | | | | | | | | | | | | | LG4 | |
| | | Fvb6_34763k | | | | | | | | Fvb6_35142k | | | | Fvb6_36607k | | Fvb4_30092k | |
Genotype ID	Sex	34763259	34763308	34763320	34763379	34763440	34763505	34763655	34763760	35142087	35142280	35142453	35142492	36607138	36607180	30092575	30092582
NM-LNF23	H	G	G/T	A	T	G/T	C	T	A/G	G/T	C/T	—	—	A/G	C	C	C
NM-LNF14	H	G	T	A	T	G	C	A/T	A	G/T	C/T	—	—	G	C	A/C	G/T
NM-LNF2	F	G	G/T	A	T	G/T	C	A/T	A/G	T	C	—	—	A/G	C	C	C
OR-MRD30	F	G	T	A/G	T	G	G/C	T	A/G	T	C	—	—	G	C/T	G/C	A/C
OR-MRD90	F	A/G	T	A	T	G	C	T	G	T	C	—	—	G	C	G/C	A/C
OR-MRD93	H	G	T	A	G/T	G	C	T	G	T	C	A/G	G	G	C	A	G
OR-MRD61	H	.	.	.	.	.	.	.	A/G	T	C	G	G	G	C	A	G

II. Crosses

A. Within NM-LNF

Dam × Sire	Sex	No. of Progeny	34763259	34763308	34763320	34763379	34763440	34763505	34763655	34763760	35142087	35142280	35142453	35142492	36607138	36607180	30092575	30092582
NM-LNF23 × NM-LNF23	F	5	—	G	—	—	G	—	—	—	G	T	—	—	A	—	—	—
	F	1	—	G	—	—	T	—	—	—	G	T	—	—	A/G	—	—	—
	F	1	—	G	—	—	T	—	—	—	G	T	—	—	.	—	—	—
	F	1	—	G/T	—	—	G/T	—	—	—	G/T	T	—	—	A	—	—	—
	H	19	—	G/T	—	—	G/T	—	—	—	G/T	C/T	—	—	A/G	—	—	—
	H	2	—	G/T	—	—	G/T	—	—	—	G/T	C/T	—	—	A	—	—	—
	H	1	—	G/T	—	—	G/T	—	—	—	G/T	C/T	—	—	G	—	—	—
	H	9	—	T	—	—	G	—	—	—	T	C	—	—	G	—	—	—
	H	1	—	T	—	—	G	—	—	—	T	C	—	—	G	—	—	—
	H	1	—	T	—	—	G	—	—	—	T	C	—	—	A/G	—	—	—
NM-LNF14 × NM-LNF14	F	4	—	—	—	—	—	—	A/T	—	G/T	C/T	—	—	—	—	—	—
	F	4	—	—	—	—	—	—	A	—	G	T	—	—	—	—	—	—
	F	7	—	—	—	—	—	—	T	—	T	C	—	—	—	—	—	—
	H	10	—	—	—	—	—	—	A/T	—	G/T	C/T	—	—	—	—	—	—
	H	5	—	—	—	—	—	—	A	—	G	T	—	—	—	—	—	—
NM-LNF2 × NM-LNF23	F	7	—	G/T	—	—	G/T	—	A/T	—	G/T	C/T	—	—	A/G	—	—	—
	F	2	—	G	—	—	T	—	T	—	G	T	—	—	A	—	—	—
	F	1	—	G/T	—	—	G/T	—	T	—	G/T	C/T	—	—	A	—	—	—
	F	1	—	.	—	—	.	—	.	—	.	.	—	—	.	—	—	—
	H	7	—	G/T	—	—	G/T	—	T	—	G/T	C/T	—	—	G	—	—	—
	H	2	—	T	—	—	G	—	A/T	—	T	C	—	—	A/G	—	—	—

B. Within OR-MRD

Dam × Sire	Sex	No. of Progeny	34763259	34763308	34763320	34763379	34763440	34763505	34763655	34763760	35142087	35142280	35142453	35142492	36607138	36607180	30092575	30092582
OR-MRD93 × OR-MRD93	F	8	—	T	—	—	—	—	—	—	—	—	G	—	—	—	—	—
	H	28	—	G/T	—	—	—	—	—	—	—	—	A/G	—	—	—	—	—
	H	3	—	G	—	—	—	—	—	—	—	—	A	—	—	—	—	—
OR-MRD61 × OR-MRD93	F	2	—	—	—	—	—	—	—	—	—	—	G	—	—	—	—	—
	H	7	—	—	—	—	—	—	—	—	—	—	A/G	—	A/G	—	—	—
	H	2	—	—	—	—	—	—	—	—	—	—	A/G	—	G	—	—	—
	H	1	—	—	—	—	—	—	—	—	—	—	G	—	G	—	—	—

(continued)

				LG6											LG4
			Fvb6_34763k						Fvb6_35142k				Fvb6_36607k		Fvb4_30092k	

C. Between OR-MRD and NM-LNF

Cross	Sex	No.	Fvb6_34763k					Fvb6_35142k				Fvb6_36607k		Fvb4_30092k	
OR-MRD30 × NM-LNF23	F	1	—	G/T	A	—	G/T					—	C	A/G	C/T
	F	1	—	G/T	A	—	G/T					—	C	·	C
	F	1	—	G/T	A	—	G/T					—	C	A/G	C
	F	3	—	G/T	A/G	—	G/T					—	G/C	A/G	C/T
	F	1	—	G/T	A/G	—	G/T					—	G/C	·	·
	F	1	—	G/T	A/G	—	G/T					—	G/C	G	C/T
	F	1	—	T	A	—	G					—	C	A/G	C
	F	2	—	T	A	—	G					—	C	G	C
	F	2	—	T	A/G	—	G					—	G/C	G	C/T
	F	1	—	T	A/G	—	G					—	G/C	·	·
	H	2	—	G/T	A	—	G/T					—	C	A/G	C
	H	4	—	G/T	A/G	—	G/T					—	G/C	A/G	C/T
	H	1	—	T	A/G	—	G					—	G/C	G	C/T
	H	2	—	T	A/G	—	G					—	G/C	A/G	C/T
	H	1	—	T	A	—	G					—	C	·	·
OR-MRD30 × NM-LNF14	F	1	—	G	A	—	G		—	T	A/G	C			
	F	1	—	G	A	—	G		A/T	G/T	A/G	C			
	F	1	—	G/T	A	—	G/T		T	T	A/G				
	F	1	—	G/T	A	—	G/T		T	T	A/G				
	F	1	—	G/T	A	—	G/T		T	T	A/G	C			
	F	1	—	G/T	A/G	—	G/T		T	A	A	C			
	F	1	—	G/T	A/G	—	G/T			A	A	C			
	F	3	—	G	A/G	—	G		A/T	·	·				
	H	1	—	G/T	A	—	G/T			A/G	A/G	C			
	H	1	—	G/T	A	—	G/T		T	A/G	A/G				
	H	1	—	G	A	—	G		A/T	G/T	A	C			
	H	1	—	G	A/G	—	G			A/G	A				
	H	1	—	T	A/G	G/T	G		T	G/T	A				
	H	1	—	T	A/G	G/T	G			·	·				
OR-MRD90 × NM-LNF23	F	5	A/G	G/T			G		A/G	G/T	C/T	G	A/C		
	F	5	G	G/T			G		A/G	G/T	C/T	A/G	A/C		
	F	10	A/G	T			G/T		A/G	G/T	C/T	A/G	A/C		
	F	4	G	T			G/T		G	T	C	G	A/C		
	H	3	G	G/T			G		G	T	C	G	A/C		
	H	7	A/G	G/T			G		A/G	T	C	G	A/C		
	H	3	G	T			G/T						G	G/C	
	H	7	A/G	T			G/T						G	G	
OR-MRD93 × NM-LNF23	F	5	A/G										G/C	A/C	
	F	1	G										G/C	A/C	
	H	4	A/G										G/C	A/C	
	H	9	G										G/C	A/C	
	H	5	A/G										G/C	A/C	
OR-MRD61 × NM-LNF23	H	1	A/G										G/C	G	

(continued)

Table 2, continued

		LG6			LG4
		Fvb6_34763k	Fvb6_35142k	Fvb6_36607k	Fvb4_30092k
H	9	G/T	G T C	**G**	
H	7	G/T	A G/T C/T	**G** A/G	
H	6	T	A/G G/T C/T	**G** G	
F	2	G/T		**G** A/G	
F	10	G/T		**A/G** G	
F	9	T	A/T	**A/G** G	
H	3	G/T	T	**A/G** G	
H	1	G/T	T		
H	3	**G/T**	A/T		
NM-LNF2 × OR-MRD93					

Sexual phenotypes (F = female, H = hermaphrodite) and genotypes for parents (I) and progeny (II) following intra- (A, B) and inter-population (C) crosses. Genotypes in progeny are only shown for screened positions that were polymorphic in the parents. Dashes denote nonsegregating SNPs, "." indicate missing data, whereas empty cells reflect positions that were not sequenced. SNPs in bold are correlated with sex, while SNPs in italics are not. SNP, Single nucleotide polymorphism.

seeds NM-LNF14×LNF23 segregated in a manner consistent with the putative genotypes (1:3; $P = 0.43$).

We sequenced 20 progeny from female NM-LNF2 by hermaphrodite NM-LNF23 cross for the Fvb6_34763k and Fvb6_35142k markers and found three SNP that segregated with sexual phenotype (Table 2II-A). At Fvb6_35142k positions 35142280 and 35142453 the R haplotype is T_C, and the r haplotype is G_T, all nine hermaphrodites are T_C and nine of the 11 females are G_T/T_C ($P < 0.0003$). SNP position 34763440 of Fvb6_34763k also segregates with sex ($P < 0.0001$), and is consistent with only the same two female types mismatching, although the genotype is missing for one putatively mismatched female at Fvb6_34763k (Table 2II-A). Although it would be desirable to genotype progeny in the Fvb6 34.8-36.6 Mbp region from the NM-LNF25 crosses, this individual was not heterozygous for any of our SNP markers in this region (see Table 2I).

In sum, there is a clear indication from the combined phenotypic and genotypic data that the Fvb6 34.8-36.6 Mbp region of the genome influences sex expression when either a female or a hermaphrodite from NM-LNF population is the maternal parent. However, there apparently is a third unmapped locus (LGx) influencing sex phenotype.

Sex expression in OR-MRD crosses

To evaluate whether hermaphrodites from OR-MRD also carried sex determining loci we evaluated progeny sex ratios from three self-pollinated hermaphrodites. One (OR-MRD93) showed a pattern consistent with a 1:3 sex ratio ($\chi^2 < 0.5$; $P > 0.50$), but not 0:1 ($\chi^2 = 50.6$; $P < 0.001$), one (OR-MRD61) deviated significantly from 1:3 ($\chi^2 = 14.3$; $P < 0.001$) but fit a 0:1 ratio ($P > 0.30$), and one (OR-MRD45) fit both 1:3 and 0:1 equally well (both $\chi^2 < 0.5$; $P > 0.30$), (Table 1B). From this we inferred the genotype of OR-MRD93 as a $Rr\ TT$ heterozygote and OR-MRD61 as a $RR\ TT$ homozygote (Table 3). Crosses conducted between OR-MRD61 and the other two OR-MRD hermaphrodites produced nearly exclusively hermaphrodite progeny (there was one anomalous female; Table 1B), corroborating the inferred genotype based on selfing of OR-MRD61 as $RR\ TT$. The other hermaphrodite by hermaphrodite crosses, however, did not produce any (OR-MRD93 × OR-MRD61) or many viable seeds (between OR-MRD93 and OR-MRD45) (Table 1C). And although too few seeds (four and seven hermaphrodite progeny) were produced from OR-MRD93 × OR-MRD45 reciprocal crosses to differentiate between 1:1 or 1:3 ratios, crosses between OR-MRD45 and NM-LNF plants (see *Discussion* below; Table 1, C and D) clarify the genotype of OR-MRD45 as RR at the LG6 locus. In addition, because the OR-MRD45 self-cross produced a few female progeny and fit a 1:3 ratio (see above; Table 1B), we deduce the genotype of OR-MRD45 as $RR\ Tt$ (Table 3). One must acknowledge, however, that so few female progeny could also reflect instability in male function or seed contaminants, so the inferred genotypes at the LGx are more tentative than those at the LG4 and LG6 loci.

To determine whether sex of progeny from these hermaphrodites is determined by the same region as in NM-LNF hermaphrodites, we genotyped progeny from OR-MRD93self at the LG6 markers (Table 2II-B). SNPs at position 34763379 on Fvb6_34763k and at position 35142492 on Fvb6_35142k both segregated perfectly with sexual phenotype ($P < 0.0001$) with females being T_G and hermaphrodites being G/T-/A/G heterozygotes or G_A homozygotes. Thus, we have evidence here that the male sterility region on LG6 also influences sexual phenotype in the OR-MRD population when hermaphrodites are the dam. The fact that the same genomic region determines sex in both OR-MRD93 with A mitotype and NM-LNF with F mitotype (Table S1) suggests that the same R locus is responsible for sex phenotype in both cytoplasmic backgrounds. However, different PPR genes among the 15 at this

■ Table 3 Inferred genotypes of parents at sex determining loci in two populations of gynodioeicous *F. vesca* subsp. *bracteata*

Population	Plant ID	Sex	Mt Haplotype Code	Locus LG4	Inferred Genotypes		
					Locus LG6	Locus LGx	
OR-MRD	OR-MRD93	H	B	*mfmf*	*Rr*	*TT*	
	OR-MRD61	H	C	*mfmf*	*RR*	*TT*	
	OR-MRD45	H	C	*mfmf*	*RR*	*Tt*	
	OR-MRD30	F	C	*MSmf*	*RR*	*TT*	
	OR-MRD27	F	C	*MSmf*	*RR*	*TT*	
	OR-MRD90	F	C	*MSmf*	*RR*	*TT*	
NM-LNF	NM-LNF23	H	F	*mfmf*	*Rr*	*TT*	
	NM-LNF25	H	F	*mfmf*	*Rr*	*TT*	
	NM-LNF14	H	F	*mfmf*	*Rr*	*Tt*	
	NM-LNF2	F	F	*mfmf*	*rr*	*TT*	
	NM-LNF4	F	F	*mfmf*	*rr*	*TT*	
	NM-LNF26	F	F	*mfmf*	*rr*	*TT*	

Plant identity, phenotypic sex (F = female; H = hermaphrodite), mitochondrial (mt) haplotype (code, see Table S1) and putative genotype at loci mapped to linkage group 4 and 6 (LG4 and LG6), or unmapped but inferred from progeny segregation ratios (Locus LGx). At LG4 the *MS* allele codes for male sterility and is dominant to *mf* which confers male fertility. At the LG6 locus, the *R* 'restores' male fertility and is dominant to *r* which does not (and thus, codes for male sterility). At the unmapped locus LGx, the *T* codes male fertility and is dominant to *t* which codes for male sterility.

locus could be the functional restorer in different cytoplasmic backgrounds.

When all three OR-MRD hermaphrodites were used as sires on three different females from OR-MRD, all produced 1:1 female to hermaphrodite ratios (all $\chi^2 < 2.63$; $P > 0.11$; Table 1B). Such patterns are consistent with the OR-MRD30×OR-MRD60 cross mapped in Tennessen *et al.* (2013) where a dominant allele in coupling with male sterility on the LG4 chromosome determines sex. Females were thus inferred to be *MSmf* whereas hermaphrodites to be *mfmf* at this locus (Table 3).

Finally, genotyping the hermaphrodite progeny from the OR-MRD61 × OR-MRD93 cross showed no SNP segregating with sexual phenotype as expected from a $Rr \times RR$ cross (Table 2II-B and Table 3). It also suggests that these sex determiners are allelic and function both on the A mitotype and C mitotype cytoplasm (OR-MRD93 and OR-MRD61, respectively) in the absence of the influence of *MS* at the LG4 locus.

Interpopulation crosses to evaluate interactions between LG4 and LG6 sex determining regions

To evaluate whether maternal control of sex at LG4 in OR-MRD extended to NM-LNF sires we performed interpopulation crosses on all three OR-MRD females. Eight out of the nine crosses produced 1:1 sex ratios ($\chi^2 < 2.57$; $P > 0.22$); the remaining cross (OR-MRD27×NM-LNF23) was skewed toward females ($\chi^2 = 4.77$; $P = 0.03$) (Table 1C).

For three of these crosses we genotyped progeny for LG6 and LG4 markers (Table 2II-C). None of the segregating SNPs on LG6 segregated with sexual phenotype in any cross ($P = 0.67$-1), but all three had SNPs at LG4 Fvb4_30092k positions 30092575 and 30092582 segregating with sex in the progeny ($P < 0.0001$). All female progeny were heterozygotes G/C_A/C whereas hermaphrodites were G_G homozygotes in families of OR-MRD30×NM-LNF23 and OR-MRD90×NM-LNF23 (Table 2II-C). In the OR-MRD30×NM-LNF14 family all female progeny were A homozygotes whereas hermaphrodites were A/C heterozygotes at the LG4 Fvb4_30092k position 30092575. Taken together these results are consistent with epistatic dominance of the *MS* allele at the LG4 locus over the *R* allele at the LG6 locus in determining sexual phenotype when the maternal parent is a female from OR-MRD.

To evaluate whether the sex-determiners at the LG6 locus of hermaphrodites from OR-MRD and NM-LNF are allelic (*e.g.*, alleles at the same locus) and determine sex in the absence of the *MS* allele, we assessed progeny from OR-MRD hermaphrodite dams with NM-LNF hermaphrodite sires (Table 1C). No seed was produced when NM-LNF14 was a father, but crosses with either NM-LNF23 and NM-LNF25 as a sire produced progeny sex ratios that conformed to expectations ($\chi^2 < 2.6$; $P > 0.22$; Table 1C) based on inferred genotypes (Table 3). These also provide additional support for OR-MRD45 as *RR* (and not *Rr*, as all three crosses deviated significantly from 1:1 $P < 0.0001$).

To follow up on these phenotypic findings, we genotyped the progeny from OR-MRD93 and OR-MRD61 when pollinated by NM-LNF23. In the first cross, all four SNPs on Fvb6_35142k segregated significantly with sex type ($P = 0.0002$–0.05; Table 2II-C), indicating that *R* alleles in both parents influence male fertility. In the second cross, as expected if OR-MRD61 is *RR*, none of the LG6 SNPs segregated with sex and all progeny are hermaphrodite.

To further explore the interaction between LG6 regions we assessed interpopulation crosses between NM-LNF females and OR-MRD hermaphrodites (Table 1C). When crossed to OR-MRD45 and OR-MRD61, all three LNF females produced a 0:1 sex ratio in the progeny as expected if the dam is *rr* and both the sires are *RR* at LG6 and *mfmf* at LG 4 (Table 3). The exception, however, was that when OR-MRD93 was the sire all three NM-LNF females produced sex ratios that deviated significantly from the 1:1 expectation (all $P < 0.01$; Table 1C). In fact, for all three crosses the female: hermaphrodite ratios fit a 2:1 (all $P > 0.80$), potentially indicating a lethal genotype. To explore this hypothesis we genotyped the progeny from the NM-LNF2×OR-MRD93 cross and found the LG6 SNPs at position 34763379 on Fvb6_34763k and at position 35142492 on Fvb6_35142k both segregated perfectly with sexual phenotype ($P < 0.0001$), females T_G and hermaphrodites G/T_A/G as predicted by the identified sex locus. Moreover, we recover all four expected genotypes in the offspring (consider both the maternally segregating genotypes (*e.g.*, at Fvb6_34763308), and paternally segregating genotypes (*e.g.*, at Fvb6_34763379): G/T_T, T_T, G/T_G/T, and T_G/T). Thus, the data are consistent with an unlinked locus causing 25% of offspring who are also *Rr* hermaphrodites to die. Interpopulation crosses between NM-LNF hermaphrodites as dams and OR-MRD hermaphrodites as sires did not produce many seeds (Table 1C). However, the fact that the cross between NM-LNF14 and OR-MRD45 produced two females out of four total progeny is consistent with their inferred heterozygous genotypes at the unmapped LGx locus (*e.g.*, *mfmf Rr Tt* and *mfmf RR Tt*, respectively; Table 3).
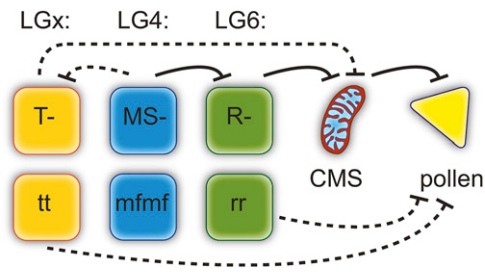
Figure 4 Conceptual model of male sterility loci in *F. vesca* subsp. *bracteata*. Black lines represent potential inhibitory interactions between genes. Dashed lines are uncharacterized genetically or posed as alternatives to the solid lines (see text). Three nuclear loci are represented by their known (LG4 and LG6) or unknown (LG*x*) location in the genome. The hypothesized involvement of these loci with a CMS locus that blocks pollen development. The LG6 locus (*R/r*) is a restorer locus at which one copy of a dominant allele is sufficient to block CMS and restore fertility; thus male sterility is recessive at this locus. The LG4 locus (*MS/mf*) is an inhibitor at which a single dominant allele is sufficient to block the LG6 restorer. Thus the *MS* allele is both dominant to *mf* and epistatically dominant to *R*. Homozygosity for the *t* allele at the third unmapped locus on LG*x* also leads to disruption of pollen production, or could act as a restorer (*T* allele) similar to the LG6 locus.

DISCUSSION

The work presented here combined with evidence from Tennessen *et al.* (2013), Stanley *et al.* (2015) and Govindarajulu *et al.* (2015), represents the first genomic evidence of both CMS and nuclear genes (fertility restorers) in a wild gynodioecious species. Moreover, demonstration of population variation in sex determiners provides insight into the potential complexity of cyto-nuclear gynodioecy in the wild and provides a road map for new hypotheses for the evolution of sex chromosomes in the related octoploid *Fragaria*.

A multilocus model of sex expression in *F. vesca* subsp. *bracteata*

We have genetically mapped two unlinked epistatically interacting loci that determine sexual phenotype in wild populations of gynodioecious *F. vesca* subsp. *bracteata*. A single dominant *R* allele at the novel locus fine-mapped on LG6 restores fertility in the absence of *MS* allele at the LG4 locus previously identified by Tennessen *et al.* (2013). This newly identified LG6 34.8–36.6 Mbp region coincides with the second most dense cluster of PPR in the *F. vesca* reference genome and houses 361 genes. Four of the 15 PPRs (Table S3) can be considered top candidates for the fertility restorer, due to their presence in the HOM03D000002 gene family that contains the radish *Rfo* gene (the only confirmed CMS-interacting fertility-restorer in the rosid clade, Chen and Liu 2014), *Arabidopsis* RFL genes (Fujii *et al.* 2011) and putative RF loci in peach (Donoso *et al.* 2015). One of these four *F. vesca* HOM03D000002 PPR genes is predicted to be mitochondrial targeted (Proost *et al.* 2015) and the other three do not have an organellar prediction. Other genes potentially involved in male function (upregulated in meiotic anthers; Hollender *et al.* 2014) also reside in this region, although not more so than the rest of the genome. Nonetheless it is intriguing that several important gene domains (F-box, Kelch repeat and LRR) occur here and in the LG4 male sterility region (*MSmf*) (Tennessen *et al.* 2013). Additionally, the action of miRNAs has been proposed as a route to sex determination (Akagi *et al.* 2014; Fagegaltier *et al.* 2014), and seven genes that are miRNA targets reside in the LG6 region (Table S3). Only one of these is a PPR, but it is chloroplast (cp)-targeted and not in the HOM03D000002 gene family. Furthermore, none of these seven genes

are targets for the novel F-box associated *Fragaria* miRNA family recently described in *F. vesca* (Xia *et al.* 2015).

A third locus (LG*x*) is also predicted and, although based on only a few crosses, segregation analysis suggests it is unlinked to the LG6 locus (*i.e.*, 9:7 ratio in the progeny of LNF14self; Table 1 and Table 2). Extensive mapping of genome-derived sequences around the sex-determining regions in several crosses revealed that the *MS* allele was epistatically dominant to *R* at LG6 and *T* at LG*x*. In fact, segregation results suggest that all loci interact, but additional crosses are necessary to confirm the existence and interaction of the LG*x* locus in particular, which was found to segregate in only a few crosses and has not been genetically mapped. Ultimately, functional studies will be needed to determine if and how the three loci interact molecularly and biochemically.

Toward this end, we propose a conceptual model of these gene interactions (Figure 4) that draws on strong evidence in the current study and Tennessen *et al.* (2013) for two nuclear loci and more limited evidence for a third, along with knowledge of the presence of a CMS-like *atp8-orf225* in the mitochondria (Table S1; Stanley *et al.* 2015; Govindarajulu *et al.* 2015). We hypothesize that a CMS locus that disrupts pollen development by production of a toxic protein or energy deficiencies (reviewed by Chen and Liu 2014) exists ancestrally. We propose male function is then restored by one or more copies of *R* at the LG6 34.8–36.6 Mbp locus. The presence of a cluster of PPRs at this locus, and the dominance of the restoring allele are consistent with data and theory of restorers of CMS (Delph *et al.* 2007; Chen and Liu 2014), but we note the current mapping precision does not rule out other possibilities. The LG4 locus, which is not associated with annotated PPRs (Tennessen *et al.* 2013) is proposed to be an inhibitor because our data shows a single dominant allele *MS* is sufficient to block the action of the LG6 restorer. Although of unknown function, the LG4 locus may be a necessary cofactor required by a PPR or another mitochondrial sorting gene involved in nuclear-mitochondrial crosstalk (Figure 3 in Chen and Liu 2014), but again a precise regulator has not been identified. Our data supports the *MS* allele as both dominant to *mf* and epistatically dominant to *R*. In our model, homozygosity for the *t* allele at the putative LG*x* locus also leads to disruption of pollen production. This locus has not been genetically mapped, but we hypothesize it could represent an independently evolved restorer (*T* allele), as cyto-nuclear theory predicts successive 'waves' of restoration (reviewed in Delph *et al.* 2007), and multiple restorer systems are common, with some loci at or near fixation (*e.g.*, Garraud *et al.* 2011; Caruso and Case 2013). Multiple restorers have been known to occur when there is polymorphism in CMS (*e.g.*, Dudle *et al.* 2001; Garraud *et al.* 2011) as there appears to be in *F. vesca* subsp. *bracteata* (Table S1; Stanley *et al.* 2015). If so, the products of the LG4 locus might also be able to block its action. Alternatively, the LG*x* locus could be a novel recessive male sterility locus that acts independently of the others (as seen in Irkaeva *et al.* 1993). By this same token, the LG6 locus may also be a recessive male sterility locus that interacts with products of the LG4 locus unrelated to CMS, but in this case this possibly seems less likely because we have specific evidence of PPRs at this locus and a CMS-like *atp8-orf225* in gynodioecious *F. vesca* subsp. *bracteata*. Additional fine mapping in these crosses and others is needed to test and refine this model.

The layers of interaction in this model may seem complex, but likely still reflect a simplification of the web of genetic mechanisms involved in cyto-nuclear gynodioecy in the wild. Much of what we know about these cyto-nuclear interactions comes from hybrids of hermaphroditic wild species (*e.g.*, *Mimulus guttattus*, Barr and Fishman 2010; *Arabidopsis lyrata*, Aalto *et al.* 2013) or crop species where restorers are studied in fixed genetic backgrounds (reviewed in Chen and Liu 2014). In fact it

has been proposed that one advantage of crop systems for understanding elements of cyto-nuclear gynodioecy is the reduced variation (Delph *et al.* 2007). However, these simplified systems may not be good models for natural processes, and cannot inform on the dynamics that govern the balanced polymorphism of wild gynodioecious species, where the ecological context, and pleiotropic effects of restorers and CMS variants govern persistence. Nevertheless, dissecting genetic mechanisms in wild species will allow comparisons with crop (or cryptic CMS) species. Our genomic evidence of both cytoplasmic and nuclear players take a step in that direction, and joins a small number of studies that have genetically mapped nuclear sex-determining loci in wild gynodioecious plants (Touzet *et al.* 2004), or have performed extensive crosses to demonstrate the existence of both nuclear and cytoplasmic contributors (*e.g.*, Van Damme *et al.* 2004, Garraud *et al.* 2011). The next steps will involve both wider sampling of plants and populations to confirm the presence of three loci, and functionally verifying their genetic interactions. Several types of approaches will be required for this and to fully test the working model we have proposed—including finer mapping of nuclear sex loci and transformation of both candidate nuclear restorer genes and the CMS-like *atp8-orf225* in the mitochondrial genome (see Chen and Liu 2014).

Evolutionary inferences from population differentiation in sex determination

We observed population variation in the frequency of sex determiners, leading especially to differences in the genotypic composition of females. All examined females in OR-MRD have dominant male sterility (*MSmf R_ TT* genotypes), whereas all the females in NM-LNF have only recessive male sterility (*mfmf rr TT* genotypes). It should be noted that *mfmf rr TT* females are possible from selfed hermaphrodites in OR-MRD (*i.e.*, OR-MRD93), but would be expected to be rare given the low frequency of *r* in this population. More generally, if our sampling is representative of the frequency of genotypes in the populations, then *r* is more frequent (1:11 *vs.* 9:3; $P < 0.01$) in NM-LNF than OR-MRD, while *MS* is more frequent in OR-MRD than NM-LNF, though the latter is not statistically significant (3:9 *vs.* 0:12; $P > 0.05$). This spatial genetic structure of sex determiners could provide insights into past evolutionary dynamics. Cyto-nuclear coevolution can lead to multiple restorer loci and polymorphism at both nuclear and mitochondrial genes (*e.g.*, Garraud *et al.* 2011), and local processes are expected to result in spatial structure for restorer and CMS haplotypes (Bailey and McCauley 2005). So it is especially intriguing that the nuclear differences mirror the strong population differentiation in the mitotypes between the study populations. The F mitotype is only found in NM-LNF, whereas the B and C mitotypes found in OR-MRD are more widespread (Stanley *et al.* 2015). If, as our results suggest, LG6 and LGx are restorer loci then *R* and *T* can restore all three mitotypes, and this would imply a greater spatial range for restorers than mitotypes. Such a pattern has been inferred from crosses with differing cytotypes in *Plantago coronopus* (Van Damme *et al.* 2004) and *Silene nutans* (Garraud *et al.* 2011). In contrast, the dominant *MS* allele at LG4 appears restricted to the northwest (OR-MRD, in this study and Tennessen *et al.* 2013; and possibly 'HP' population in northern California, Ahmadi and Bringhurst 1991), perhaps indicating that it evolved after the LG6 and LGx loci, consistent with our multi-layered hypothesis for sex determination (Figure 4). The interactions between these genes likely contribute to high variation in sex ratios across space in *F. vesca* subsp. *bracteata* (0–46% females; Stanley *et al.* 2015). More crosses between populations of varying distances across the range (*e.g.*, Bailey and McCauley 2005) and/or targeted resequencing of the sex determining regions would shed light on the wider geographic context

for the players in the *F. vesca* subsp. *bracteata* sex determination system and our sex-determination hypothesis. This combined with fitness consequences of specific genotypes is needed to test evolutionary hypotheses for maintenance of multiple restorers and mitotypes in this species (*e.g.*, Caruso *et al.* 2012).

The genetic differentiation of NM-LNF and OR-MRD is also evident in Stanley *et al.* (2015) where STRUCTURE analysis based on nuclear markers distinguished OR-MRD as belonging to cluster 1 while NM-LNF to cluster 2 (more closely allied with subspecies *F. vesca* subsp. *americana*). Similar distinction was seen in the chloroplast. Such differentiation could explain not only the differences in prevailing sex determiners but also the asymmetry in success of the interpopulation H×H crosses. Crosses with a OR-MRD hermaphrodite as a sire on a NM-LNF hermaphrodite as a dam produced fewer seeds than the reciprocal, potentially reflecting cyto-nuclear incompatibilities (Rieseberg and Blackman 2010) that interact with the sex determiners because similar deficits were not seen when females were the dams. Alternatively these could reflect exposure of costs associated with restoration that may be complex (Caruso *et al.* 2012). It is notable in this regard that NM-LNF14 performed poorly as a sire and a dam, and also was the only plant inferred to be heterozygous at both the putative fertility-restorer loci (Table 3).

Relationship to sex determination octoploids and other *Fragaria*

The present results can be considered in the context of the known sex determining regions in the two octoploid descendants of *F. vesca* subsp. *bracteata*. In both *F. virginiana* and *F. chiloensis*, linked male and female sterility loci reside on a chromosome in homeologous group VI (LG VI-Av in *F. chiloensis*, and VI-B2 *F. virginiana*; Goldberg *et al.* 2010; Spigler *et al.* 2011; Tennessen *et al.* 2014). Thus, it is intriguing to identify a locus on LG6 in *F. vesca* subsp. *bracteata* that also affects sexual expression. The fact that the sex determining region in *F. vesca* subsp. *bracteata* is near the 3′ end of LG6, as is the sex determining region of *F. chiloensis* (Goldberg *et al.* 2010), raises the possibility they are homologous. However, the *F. vesca* subsp. *bracteata* LG6 locus (*Rr*) codes for recessive male sterility, whereas the LG VI-Av locus in *F. chiloensis* has dominant male sterility. This difference alone might suggest they are not the same locus, but it could also reflect a turnover in the sex determining chromosome. Heterogametic transitions are theoretically possible and can transition from male heterogamety to female heterogamety when release from mutational load is followed by sexually antagonistic selection (Blaser *et al.* 2014).

If, on the other hand, the *F. vesca* subsp. *bracteata* LG6 locus and the LG VI-Av sex determining locus in *F. chiloensis* are entirely unrelated, the discovery of the *F. vesca* subsp. *bracteata* LG6 locus provides additional evidence that LG6 is predisposed to be a sex determining chromosome, as initially proposed by Spigler *et al.* (2011) with respect to *F. virginiana* and *F. chiloensis*. Some autosomes are thought to be prone to hosting sex determining regions because they house many genes that can affect male function, or because a chromosome that already has been involved in sex determination is more likely to seize back this role in the future (Graves and Peichel 2010; Blaser *et al.* 2014). This could be facilitated by transposition of genes already involved in sex expression (*e.g.*, Hughes *et al.* 2015), making the LG4 locus a candidate for such transposition during the origin of the octoploids producing a gene complex in *F. chiloensis*. Finer mapping and deep resequencing of the sex determining region in *F. chiloensis* is underway to test these ideas. In either event, the present results suggest great potential for sexual lability in *Fragaria*, as male sterility evolves frequently, independently and via different genetic mechanisms. Indeed,

recent phylogenetic character state transition analysis detects frequent transitions to and from separate sexes (dioecy) in *Fragaria* (Goldberg *et al.* unpublished results). The present study allows us to speculate that this lability is facilitated by ancient evolution of CMS (as evidenced by the widespread existence of the mitochondrial ORF (*atp8-orf225*) in the genus *Fragaria*; Govindarajulu *et al.* 2015; Stanley *et al.* 2015), and repeated evolution of restorers and/or other suppressors of male function, and turnovers in sex determining chromosomes.

LITERATURE CITED

Aalto, E. A., H.-P. Koelewijn, and O. Savolainen, 2013 Cytoplasmic male sterility contributes to hybrid incompatibility between subspecies of *Arabidopsis lyrata*. G3 (Bethesda) 3: 1727–1740.

Ahmadi, H., and R. S. Bringhurst, 1991 Genetics of sex expression in *Fragaria* species. Am. J. Bot. 78: 504–514.

Akagi, T., I. M. Henry, R. Tao, and L. Comai, 2014 A Y-chromosome-encoded small RNA acts as a sex determinant in persimmons. Science 346: 646–650.

Bachtrog, D., J. E. Mank, C. L. Peichel, M. Kirkpatrick, S. P. Otto *et al.*, 2014 Sex determination: why so many ways of doing it? PLoS Biol. 12: e1001899.

Bailey, M. F., and L. F. Delph, 2007 A field guide to models of sex-ratio evolution in gynodioecious species. Oikos 116: 1609–1617.

Bailey, M. F., and D. E. McCauley, 2005 Offspring sex ratio under inbreeding and outbreeding in a gynodioecious plant. Evolution 59: 287–295.

Barr, C. M., and L. Fishman, 2010 The nuclear component of a cytonuclear hybrid incompatibility in Mimulus maps to a cluster of pentatricopeptide repeat (PPR) genes. Genetics 184: 455–465.

Blaser, O., S. Neuenschwander, and N. Perrin, 2014 Sex-chromosome turnovers: the hot-potato model. Am. Nat. 183: 140–146.

Caruso, C. M., and A. L. Case, 2013 Testing models of sex ratio evolution in a gynodioecious plant: female frequency covaries with the cost of male fertility restoration. Evolution 67: 561–566.

Caruso, C. M., A. L. Case, and M. F. Bailey, 2012 The evolutionary ecology of cytonuclear interactions in angiosperms. Trends Plant Sci. 17: 638–643.

Charlesworth, B., and D. Charlesworth, 1978 A model for the evolution of dioecy and gynodioecy. Am. Nat. 112: 975–997.

Chen, L., and Y.-G. Liu, 2014 Male sterility and fertility restoration in crops. Annu. Rev. Plant Biol. 65: 579–606.

Crossman, A., and D. Charlesworth, 2014 Breakdown of dioecy: Models where males acquire cosexual functions. Evolution 68: 426–440.

Cui, X., Q. Wang, W. Yin, H. Xu, Z. A. Wilson *et al.*, 2012 PMRD: a curated database for genes and mutants involved in plant male reproduction. BMC Plant Biol. 12: 215–224.

Dalton, R., M. Koski, and T.-L. Ashman, 2013 Maternal sex effects and inbreeding depression under varied environmental conditions in gynodioecious *Fragaria vesca* subsp. *bracteata*. Ann. Bot. (Lond.) 112: 613–621.

Darwish, O., R. Shahan, Z. Liu, J. P. Slovin, and N. W. Alkharouf, 2015 Re-annotation of the woodland strawberry (*Fragaria vesca*) genome. BMC Genomics 16(1): 29.

Delph, L. F., P. Touzet, and M. F. Bailey, 2007 Merging theory and mechanism in studies of gynodioecy. Trends Ecol. Evol. 22: 17–24.

Diggle, P. K., V. S. Di Stilio, A. R. Gschwend, E. M. Golenberg, R. C. Moore *et al.*, 2011 Multiple developmental processes underlie sex differentiation in angiosperms. Trends Genet. 27: 368–376.

Donoso, J. M., I. Eduardo, R. Picañol, I. Batlle, W. Howad *et al.*, 2015 High-density mapping suggests cytoplasmic male sterility with two restorer genes in almond× peach progenies. Horticulture Research 2: 1–10.

Dudle, D. A., P. Mutikainen, and L. F. Delph, 2001 Genetics of sex determination in the gynodioecious species *Lobelia siphilitica*: evidence from two populations. Heredity 86: 265–276.

Dufaÿ, M., P. Touzet, S. Maurice, and J. Cuguen, 2007 Modelling the maintenance of male-fertile cytoplasm in a gynodioecious population. Heredity 99: 349–356.

Dufaÿ, M., J. Cuguen, J.-F. Arnaud, and P. Touzet, 2009 Sex ratio variation among gynodioecious populations of sea beet: Can it be explained by negative frequency-dependent selection? Evolution 63: 1483–1497.

Dufaÿ, M., P. Champelovier, J. Käfer, J.-P. Henry, S. Mousset *et al.*, 2014 An angiosperm-wide analysis of the gynodioecy–dioecy pathway. Ann. Bot. (Lond.) 114: 539–548.

Dufresnes, C., Y. Bertholet, J. Wassef, K. Ghali, R. Savary *et al.*, 2014 Sex-chromosome differentiation parallels postglacial range expansion in European tree frogs (*Hyla arborea*). Evolution 68: 3445–3456.

Ehlers, B. K., and T. Bataillon, 2007 'Inconstant males' and the maintenance of labile sex expression in subdioecious plants. New Phytol. 174: 194–211.

Fagegaltier, D., A. König, A. Gordon, E. C. Lai, T. R. Gingeras *et al.*, 2014 A genome-wide survey of sexually dimorphic expression of Drosophila miRNAs identifies the steroid hormone-induced miRNA let-7 as a regulator of sexual identity. Genetics 198: 647–668.

Frank, S., 1989 The evolutionary dynamics of cytoplasmic male sterility. Am. Nat. 133: 345–376.

Fujii, S., C. S. Bond, and I. D. Small, 2011 Selection patterns on restorer-like genes reveal a conflict between nuclear and mitochondrial genomes throughout angiosperm evolution. Proc. Natl. Acad. Sci. USA 108: 1723–1728.

Garraud, C., B. Brachi, M. Dufaÿ, P. Touzet, and J. Shykoff, 2011 Genetic determination of male sterility in gynodioecious *Silene nutans*. Heredity 106: 757–764.

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

Govindarajulu, R., A. Liston, and T. -L. Ashman, 2013 Sex-determining chromosomes and sexual dimorphism: insights from genetic mapping of sex expression in a natural hybrid Fragaria × ananassa subsp. cuneifolia. Heredity 110: 430–438.

Govindarajulu, R., M. Parks, J. A. Tennessen, A. Liston, and T.-L. Ashman, 2015 Comparison of nuclear, plastid, and mitochondrial phylogenies and the origin of wild octoploid strawberry species. Am. J. Bot. 102: 544–554.

Graves, J., and C. Peichel, 2010 Are homologies in vertebrate sex determination due to shared ancestry or to limited options? Genome Biol. 11: 205.

Hollender, C. A., A. C. Geretz, J. Slovin, and Z. Liu, 2012 Flower and early fruit development in a diploid strawberry, *Fragaria vesca*. Planta 235: 1123–1139.

Hollender, C. A., C. Kang, O. Darwish, A. Geretz, B. F. Matthews *et al.*, 2014 Floral transcriptomes in woodland strawberry uncover developing receptacle and anther gene networks. Plant Physiol. 165: 1062–1075.

Hughes, J. F., H. Skaletsky, N. Koutseva, T. Pyntikova, and D. C. Page, 2015 Sex chromosome-to-autosome transposition events counter Y-chromosome gene loss in mammals. Genome Biol. 16: 104–113.

Irkaeva, N. M., L. V. Lebedeva, and D. S. Shaziyatova, 1993 Genetics of male sterility in the strawberry. Genetika 29: 1485–1491.

Jacobs, M. S., and M. J. Wade, 2003 A synthetic review of the theory of gynodioecy. Am. Nat. 161: 837–851.

Kearns, C. A., and D. W. Inouye, 1993 Techniques for Pollination Biologists, Univ. Press of Colorado, Niwot, Colorado.

Klaas, A. L., and M. S. Olson, 2006 Spatial distributions of cytoplasmic types and sex expression in Alaskan populations of *Silene acaulis*. Int. J. Plant Sci. 167: 179–189.

Lewis, D., 1941   Male sterility in natural populations of hermaphrodite plants. New Phytol. 40: 56–63.

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

Liston, A., R. Cronn, and T.-L. Ashman, 2014   Fragaria: A genus with deep historical roots and ripe for evolutionary and ecological insights. Am. J. Bot. 101: 1686–1699.

Margarido, G., A. Souza, and A. Garcia, 2007   OneMap: software for genetic mapping in outcrossing species. Hereditas 144: 78–79.

Maurice, S., and T. H. Fleming, 1995   The effect of pollen limitation on plant reproductive systems and the maintenance of sexual polymorphisms. Oikos 74: 55–60.

Njuguna, W., A. Liston, R. C. Cronn, T.-L. Ashman, and N. V. Bassil, 2013   Insights into phylogeny, sex function and age of *Fragaria* based on whole chloroplast genome sequencing. Mol. Phylogenet. Evol. 66: 17–29.

Proost, S., M. Van Bel, D. Vaneechoutte, Y. Van de Peer, D. Inzé *et al.*, 2015   PLAZA 3.0: an access point for plant comparative genomics. Nucleic Acids Res. 43: D974–D981.

Renner, S. S., 2014   The relative and absolute frequencies of angiosperm sexual systems: Dioecy, monoecy, gynodioecy, and an updated online database. Am. J. Bot. 101: 1588–1596.

Rieseberg, L. H., and B. K. Blackman, 2010   Speciation genes in plants. Ann. Bot. (Lond.) 106: 439–455.

Russell, J., and J. Pannell, 2015   Sex determination in dioecious *Mercurialis annua* and its close diploid and polyploid relatives. Heredity 114: 262–271.

Schultz, S. T., 1994   Nucleo-cytoplasmic male sterility and alternative routes to dioecy. Evolution 48: 1933–1945.

Shulaev, V., D. J. Sargent, R. N. Crowhurst, T. C. Mockler, O. Folkerts *et al.*, 2010   The genome of woodland strawberry (*Fragaria vesca*). Nat. Genet. 43: 109–116.

Spigler, R. B., and T.-L. Ashman, 2012   Gynodioecy to dioecy: are we there yet? Ann. Bot. (Lond.) 109: 531–543.

Spigler, R. B., K. S. Lewers, and T.-L. Ashman, 2011   Genetic architecture of sexual dimorphism in a subdioecious plant with a proto-sex chromosome. Evolution 65: 1114–1126.

Stanley, L., N. J. Forrester, R. Govindarajulu, A. Liston, and T.-L. Ashman, 2015   Geographic patterns of genetic variation in three genomes of North American diploid strawberries with special reference to *Fragaria vesca* subsp. *bracteata*. Botany 93: 573–588.

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

Tennessen, J. A., R. Govindarajulu, A. Liston, and T.-L. Ashman, 2013   Targeted sequence capture provides insight into genome structure and genetics of male sterility in a gynodioecious diploid strawberry, *Fragaria vesca* ssp. *bracteata* (Rosaceae). G3 (Bethesda) 3: 1341–1351.

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

Touzet, P., N. Hueber, A. Bürkholz, S. Barnes, and J. Cuguen, 2004   Genetic analysis of male fertility restoration in wild cytoplasmic male sterility G of beet. Theor. Appl. Genet. 109: 240–247.

Ubeda, F., M. M. Patten, and G. Wild, 2015   On the origin of sex chromosomes from meiotic drive. Proc. R. Soc. B 282: 20141932.

Van Damme, J., M. Hundscheid, S. Ivanovic, and H. Koelewijn, 2004   Multiple CMS–restorer gene polymorphism in gynodioecious *Plantago coronopus*. Heredity 93: 175–181.

Vicoso, B., and D. Bachtrog, 2013   Reversal of an ancient sex chromosome to an autosome in Drosophila. Nature 499: 332–335.

Wang, K., X. Peng, Y. Ji, P. Yang, Y. Zhu *et al.*, 2013   Gene, protein, and network of male sterility in rice. Front. Plant Sci. 4: 1–10.

Xia, R., S. Ye, Z. Liu, B. Meyers, and Z. Liu, 2015   Novel and recently evolved miRNA clusters regulate expansive F-box gene networks through phasiRNAs in wild diploid strawberry. Plant Physiol. .10.1104/pp.15.00253.

Communicating editor: S. I. Wright