1   Multilocus sex determination revealed in two populations of gynodioecious wild strawberry, *Fragaria*

2   *vesca* subsp. *bracteata*

3

4

5   T-L. Ashman[*], J. A. Tennessen[†], R. Dalton[*], R. Govindarajulu[*], M. Koski[*], and A. Liston[‡]

6   [*]Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania, 15260 USA

7   [†]Department of Integrative Biology, Oregon State University, Corvallis, Oregon 97331

8   [‡]Department of Botany and Plant Pathology, Oregon State University, Corvallis, Oregon, 97331 USA

9

10

11   **Abstract**

12   Gynodioecy, the coexistence of females and hermaphrodites, occurs in 20% of angiosperm families and

13   often enables transitions between hermaphroditism and dioecy. Clarifying mechanisms of sex

14   determination in gynodioecious species can thus illuminate sexual system evolution. Genetic

15   determination of gynodioecy, however, can be complex and is not fully characterized in any wild species.

16   We used targeted sequence capture to genetically map a novel nuclear contributor to male sterility in a

17   self-pollinated hermaphrodite of *Fragaria vesca* subsp. *bracteata* from the southern portion of its range.

18   To understand its interaction with another identified locus and possibly additional loci, we performed

19   crosses within and between two populations separated by 2000 km, phenotyped the progeny and

20   sequenced candidate markers at both sex-determining loci. The newly mapped locus contains a high

21   density of pentatricopeptide repeat genes, a class commonly involved in restoration of fertility caused by

22   cytoplasmic male sterility. Examination of all crosses revealed three unlinked epistatically interacting loci

23   that determine sexual phenotype and vary in frequency between populations. *Fragaria vesca* subsp.

24   *bracteata* represents the first wild gynodioecious species with genomic evidence of both cytoplasmic and

25   nuclear genes in sex determination. We propose a model for the interactions between these loci and new

26   hypotheses for the evolution of sex determining chromosomes in the subdioecious and dioecious

27   *Fragaria.*

28

29   **Running title:** Sex determination in wild strawberry

30   **Key words:** male sterility, dioecy, Fragaria, gynodioecy, sex determination

31   **Author for correspondence**: Tia-Lynn Ashman Department of Biological Sciences, 4249 Fifth Ave,

32   Pittsburgh, PA 15260; email: tia1@pitt.edu) fax: 412-624-4759; phone: 412-624-0984

33

INTRODUCTION

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; Dufay et al. 2007), or may transition back to hermaphroditism (Delph et al. 2007; Goldberg et al. unpublished). 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., Dufay et al. 2009). Restorers are diverse, but most commonly are 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). 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

58   populations and there can be more than one nuclear locus (e.g., Van Damme et al. 2004; Garraud et al.

59   2011). Reversions to hermaphroditism can occur within populations when CMS or nuclear LOF genes are

60   lost or when both CMS and restorers become fixed (e.g., *Arabidopsis thaliana,* Gobron et al. 2013). As a

61   consequence, disparate populations of a widespread species that have been exposed to dissimilar selection

62   regimes or differentially subjected to drift (population size or founder events) can be differentiated for sex

63   determination mechanisms. Such variation has been revealed in species with cyto-nuclear gynodioecy

64   (e.g., *Lobelia siphiltica,* Dudle et al. 2001), but also can be seen in species with nuclear sex determination

65   or sex chromosomes (e.g., Dufresnes et al. 2014). In these cases, transitions in sex determining regions (or

66   chromosomes) are thought to be driven by additional selective forces, such as sexually antagonistic

67   selection, meiotic drivers and genetic load (e.g., Blaser et al. 2014; Ubeda et al. 2015).

68   Model organisms and agriculturally important taxa have been instrumental in characterizing sex

69   determination pathways and have revealed that transitions between sexual systems can involve new genes,

70   new alleles or entirely novel means of sex determination (e.g., Vicoso and Bachtrog 2013; Akagi et al.

71   2014). In fact, the same mechanism can be regulated by different genes, and thus developmental networks

72   can evolve rapidly (Cui et al. 2012; Wang et al. 2013). Understanding sex determination in wild plants is

73   particularly valuable because it is in these species that data can be linked to ecological processes that are

74   prevailing drivers of sexual system variation (Frank 1989; Jacobs and Wade 2003). Gynodioecious

75   species are often closely related to hermaphroditic or dioecious ones (Dufaÿ et al. 2014), so

76   characterization of their sex determination can also provide insight into genetic, developmental and

77   evolutionary dynamics of sexual systems (Spigler and Ashman 2012; Russell and Pannell 2015).

78   We explored variation in sex determination in the widespread North American diploid strawberry

79   *F. vesca* subsp. *bracteata* for two main reasons. First, it is gynodioecious and the maternal donor to the

80   sexually dimorphic octoploid *Fragaria* (dioecious *F. chiloensis* and subdioecious *F. virginiana*) (Njuguna

81   et al. 2013; Tennessen et al. 2014; Govindarajulu et al. 2015) which were hybridized to produce the

82   cultivated hermaphroditic *F. ×ananassa*. As such it represents both a tractable and evolutionarily

83   appropriate genetic model for teasing apart sex differentiation in this crop and its wild relatives (Liston et

84   al. 2014). Second, several lines of evidence allude to variation in sex determining regions within the

85   species. Specifically, although previous genetic mapping of sex determination in a northern population of

86   *F. vesca* subsp. *bracteata* identified a nuclear locus with a dominant allele for male sterility (Tennessen et

87   al. 2013), there is equivocal evidence for a fitness advantage sufficient to maintain females in this

88   population under nuclear determination alone (Li et al. 2012; Dalton et al. 2013). Moreover, recent

89   phylogeographic studies revealed genetic differentiation across the range of *F. vesca* subsp. *bracteata*,

90   with one combination of chlorotype and mitotype more closely related to octoploids than other(s)

91   (Njuguna et al. 2013; Govindarajulu et al. 2015; Stanley et al. 2015). Geographic differentiation was

92   partly due to variation in a novel mitochondrial open reading frame (ORF) with sequence similarity to

93   known CMS genes, although a sterilizing function has not been confirmed (Govindarjulu et al. 2015;

94   Stanley et al 2015). These points combined with the fact that the location of the sex determining region of

95   *F. vesca* subsp. *bracteata* identified in Tennessen et al. (2013) was on a chromosome that is different from

96   either of the chromosomal locations of the sex determining region in the two octoploid congeners (i.e.,

97   chromosome 4 (LG4) *vs.* a chromosome in homeologous group VI (LG VI-Av or VI-B2), Goldberg et al.

98   2010; Spigler et al. 2011; Tennessen et al. 2014) raises the possibility of additional genetic contributors to

99   sex determination in other populations of the diploid species.

100   We used targeted-sequence capture (Tennessen et al 2013; 2014) to efficiently fine map the sex

101   determination region using a self-pollinated hermaphrodite plant from a population at the southeastern

102   edge of *F. vesca* subsp. *bracteata*'s range within the USA. From this we identified a second locus

103   affecting sexual phenotype on a different chromosome (LG6) than that previously identified from a

104   female collected in the northwestern portion of the range (LG4; Tennessen et al. 2013). We conducted

105   intra- and inter-population crosses. In addition to phenotyping the progeny, we sequenced informative

106   SNPs near the target sex-determining regions to determine whether the second locus interacted with the

107   locus previously identified. In doing, so we reveal interactions of at least two nuclear loci and population

108   variation in sex determination. From this we propose a novel mechanistic pathway for sex determination

109   in the diploid species, and extend it as a hypothesis for sex determination in the octoploid species that

110   descended from the diploid species studied here.

MATERIALS AND METHODS

112   **Study system**

113   *Fragaria vesca* subsp. *bracteata* (Rosaceae) is a diploid ($2n = 2\times = 14$) North American wild

114   strawberry that inhabits moderately damp soils of shady forest edges, meadows, and riverbanks. Its range

115   extends from British Columbia to southern Mexico and from the Pacific coast to the Rocky Mountains

116   and Sierra Madre (Staudt 1999; Stanley et al. 2015). Variation in cytoplasmic haplotypes has been

117   revealed across this range with western populations dominated by one chlorotype and those from the

118   Rockies and Sierra Madre by another with a zone of admixture in between (Njuguna et al. 2013; Stanley

119   et al. 2015). In addition, there is finer scale and more complex pattern of mitochondrial variation (Stanley

120   et al. 2015).

121   The species has a gynodioecious sexual system (Ahmadi and Bringhurst 1989; Li et al. 2012) and

122   populations vary in the frequency of females from 0 (no females) to 46% female (Stanley et al. 2015).

123   Sex types produce similar numbers of flowers, seeds and plantlets on stolons (Li et al. 2012).

124   Hermaphrodites are self-compatible, highly selfing and morphologically distinguishable from females by

125   their slightly longer stamens and the presence of viable pollen grains in their anthers (Li et al. 2012;

126   Dalton et al. 2013).

127   The only known nuclear locus that affects sex expression was identified from a cross between a

128   female and a hermaphrodite from Oregon, USA (OR-MRD; Tennessen et al. 2013). The dominant allele

129   for male sterility resides in a 338-kb region of chromosome 4 that is predicted to house 57 genes, although

130 none in protein families known to control male sterility were found based on the reference genome *F.*

131 *vesca* Hawaii4 annotations (Tennessen et al. 2013). However, Govindarajulu et al. (2015) identified a

132 mitochondrial ORF (*atp8-orf225*) with molecular characteristics similar to known CMS genes. It is

133 present and variable in *F. vesca* subsp. *bracteata* as well as other taxa across the genus *Fragaria*

134 (Govindarajulu et al. 2015; Stanley et al. 2015), but its functional role has yet to be confirmed.

135 **Study materials**

136 *Fragaria vesca* subsp. *bracteata* from two populations separated by 2000 km were the focus of the

137 current study. NM-LNF is a high elevation 'sky island' population in the southern portion of the range (in

138 the Lincoln National Forest, New Mexico, altitude 2664 m; N 32° 58′ 6.1″, W 105° 44′ 44.7″). OR-MRD

139 is on the highest peak in the Oregon Coast Range (Marys Peak in the Siuslaw National Forest, Oregon,

140 altitude 521 m; N 44.8° 29′ 18.4″, W 123° 32′ 14.7″). The latter population is also within the zone of

141 cytoplasmic introgression seen in *F. vesca* subsp. *bracteata* between the Coast Range / Cascades and

142 Northern Rocky Mountains (Staudt 1999; Stanley et al. 2015). The frequency of females in OR-MRD is

143 30% and in NM-LNF is 46% (Stanley et al. 2015).

144   Geography, flower and fruit morphology, and sexual system indicate both populations are *F. vesca*

145 subsp. *bracteata* (Staudt 1999; T-L. Ashman, unpublished). Likewise phylogenomic analysis shows these

146 populations form a clade for the majority of their chromosomes (Tennessen et al. 2014) and are always

147 separate from *F. vesca* subsp. *vesca* Hawaii 4 (source of the published genome sequence, Shulaev et al,

148 2010) which itself forms a clade with *F. vesca* subsp. *americana* (Njuguna et al 2013). These two

149 populations, however, differed in their cytotypes as is common in *F. vesca* subsp. *bracteata*: OR-MRD

150 has one chlorotype (2) and two mitotypes (90% C and 10% B), whereas NM-LNF is fixed for different

151 chlorotype and mitotype (1 and F, respectively; Stanley et al. 2015).

152 **Crosses and sexual phenotyping**

153   Plants were collected from the wild and maintained in the greenhouse at the University of

154   Pittsburgh. Three hermaphrodites (OR-MRD61, OR-MRD45, OR-MRD93) and three females (OR-

155   MRD30 [used in Tennessen et al. 2013], OR-MRD27, OR-MRD90) from OR-MRD were used as parents.

156   From NM-LNF, three hermaphrodites (NM-LNF23 [used in Tennessen et al. 2014], NM-LNF25, and

157   NM-LNF14) and three females (NM-LNF2, NM-LNF4, NM-LNF26) were used as parents. Cytotypes of

158   the parents reflected their abundance in wild populations (Stanley et al. 2015) and are given in Table S1.

159   The three *atp8-orf225* variants translate to different proteins, although their function is unknown (Stanley

160   et al. 2015).

161   All females were crossed with pollen from both inter- and intra-population hermaphrodites, and all

162   hermaphrodites were self-pollinated. In addition, crosses between the hermaphrodites were performed

163   within and between populations. Crosses were performed by hand-pollinating 1-4 flowers per dam with

164   pollen collected from single sires. When hermaphrodites served as dams their flowers were emasculated

165   prior to pollination. Seeds were harvested and stored at -20° C until sowing. Some crosses produced few

166   seeds (and/or few germinated), notably hermaphrodite by hermaphrodite interpopulation crosses

167   involving NM-LNF as the dam and OR-MRD as the sire, and some hermaphrodite by hermaphrodite

168   within population crosses (see results Table 1). We sowed an average of 31 seeds (range 1-56) per family

169   into 98 well trays in a mix of Fafard #4, sand, and Sunshine Germination Mix. Germination was, on

170   average, high (75%), although some specific crosses performed poorly (see results). Approximately one

171   month after germination we transplanted progeny into 150cc pots of soil (two parts Fafard #4 to 1 part

172   sand) and randomized them on benches in a greenhouse where we maintained conditions of 14-15 hr days

173   21°/15.5° C (day/night) temperatures.

174   To induce flowering we subjected plants to a cold treatment (8 hr days at 8°C/16hr dark at 4°C) in

175   a growth chamber for three to four weeks, and then returned them to the greenhouse where they received

176   50ppm 10:30:20 N:P:K fertilizer (Peter's Professional® Bloom Booster) weekly. This was repeated up to

177 five times to maximize flowering. A high rate of flowering was achieved (on average 86% progeny per

178 family). Upon flowering we scored male function for at least two flowers per plant. Given the weak

179 sexual dimorphism in this species (Li et al 2012), we recorded both whether or not the anthers were

180 shedding pollen in the greenhouse and whether viable pollen grains were produced. We confirmed pollen

181 production and viability microscopically by fixing anthers with Alexander's stain (Kearns and Inouye

182 1993), and later observing them under a light microscope. Plants with no viable pollen production were

183 scored as sterile whereas those with viable pollen production were scored as fertile. In total, 1364 plants

184 flowered and the number of individuals phenotyped per family ranged from 1 to 53 (mean = 21). Progeny

185 sex ratios were scored from flowering plants in each cross and for family size >10 sex ratio was tested for

186 deviations from expected (see below) using $\chi 2$ tests.

187 **Genetic and genomic analysis**

188   To identify the sex-determining locus in NM-LNF we used the linkage map published in

189 Tennessen et al. (2014). This map was created using OneMap (Margarido et al. 2007) and the 1825

190 polymorphic sites identified from the target captured sequences of 41 progeny of self-pollinated NM-

191 LNF23 (Tennessen et al. 2014). This map had the expected seven linkage groups spanning a total of 326

192 cM and is described in full in Tennessen et al (2014) where it was used to produce a new *Fragaria vesca*

193 genome assembly ('Fvb'). Based on past experience (Tennessen et al. 2013) this sample size is sufficient

194 to identify genomic regions of major effect on male function. We coded male sterility as a recessive

195 Mendelian locus (e.g., sterile *rr* and fertile *R-*; justified because a hermaphroditic parent could not harbor

196 a dominant male sterility allele), and mapped it along with the targeted capture genotypes using OneMap

197 (Margarido et al. 2007) to determine its position on the linkage map. From this map we identified a

198 candidate region in coupling with male sterility on LG6 (see results).

199   Once the male function region was identified, we designed primers for PCR amplification and

200 Sanger sequencing of three SNPs from target capture sequences on LG6: one just upstream

201 (Fvb6_34763k), one within (Fvb6_35142k), and one downstream of the male sterility region

202 (Fvb6_36607k) (Table S2; Figure 2). Primers were designed with Primer-BLAST

203 (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) and positioned to amplify additional nearby SNPs

204 segregating in OR-MRD30, as well as the NM-LNF23 SNPs. To genotype the region housing the sex

205 expression locus on LG4, we used primers described previously using the *F. vesca* Hawaii 4 assembly but

206 that were re-aligned to the new assembly 'Fvb' (Fvb4_30092k) (Tennessen et al. 2014; Table S2).

207 Sequencing of targeted sex determining regions on LG4 and LG6 was performed on seven parents

208 and 301 progeny from 13 select crosses. The crosses were selected based on the availability of

209 segregating informative SNPs in the targeted regions in the parents (Table 2 I) and sexual phenotypes of

210 their progeny. We sequenced progeny from three crosses for LG4 markers (OR-MRD30×NM-LNF23,

211 OR-MRD30×NM-LNF14, and OR-MRD90×NM-LNF23) and progeny from ten crosses for LG6 markers

212 (OR-MRD93self, NM-LNF14self, NM-LNF2× NM-LNF23, NM-LNF2×OR-MRD93, OR-

213 MRD30×NM-LNF23, OR-MRD30×NM-LNF14, OR-MRD90×NM-LNF23, OR-MRD93×NM-LNF23,

214 OR-MRD61×NM-LNF23, and OR-MRD61×OR-MRD93).

215 For each cross we extracted DNA from 100-150 mg of fresh or 20-30 mg of silica-dried leaf tissue

216 per plant. We used a CTAB DNA protocol modified as in Govindarajulu et al. (2012) and PCR

217 amplification was performed with 1X standard reaction buffer (New England Biolabs), 100uM of each

218 dNTP, 0.5uM of each forward and reverse primer, 1.5 units of Standard Taq polymerase (New England

219 Biolabs) and 1.5 ul of genomic DNA in a 20 ul reaction. The amplification conditions were: 2.5 minutes

220 at 95°C, followed by 35 cycles of 95°C for 30 seconds, 55 and 56°C for 30 seconds, and 72°C for 60

221 seconds, and a final extension at 72°C for eight minutes (Table S2). The amplified products were Sanger

222 sequenced and aligned using Sequencher ver 4.8 (Gene Codes Corp, Ann Arbor, Michigan) and SNPs

223 were scored. For all crosses, we analyzed all segregating SNPs observed within each sequenced PCR

224 product, whether or not they were the same segregating SNPs observed in the linkage maps. Association

225 between segregating markers and sexual phenotypes was assessed in the progeny sets using Fisher's exact

226 tests.

227 We identified potential candidate genes in the region matching male sterility using the Hybrid

228 GeneMark Predictions (Shulaev et al. 2010) available from GDR: Genome Database for Rosaceae

229 (https://www.rosaceae.org/species/fragaria/fragaria_vesca/genome_v1.1) supplemented with four

230 previously unrecognized genes (Darwish et al. 2015). For functional annotation we relied on PLAZA 3.0

231 (Proost et al. 2015). We examined levels of gene expression in a floral development transcriptome

232 (Hollender et al. 2014), especially at stages 8-10 which are important in pollen development (Hollender et

233 al. 2012). To test for PPR enrichment in our genomic region of interest, we identified 653 genes annotated

234 as "Pentatricopeptide repeat-containing protein" (PPR) and counted them in 1Mb nonoverlapping

235 windows across the genome. We also examined the genomic location of the 98 PPR genes that have a

236 putative ortholog (as determined by PLAZA 3.0) with the 26 fertility restorer-like (RFL) genes in

237 *Arabidopsis thaliana* (Fujii et al. 2011).

238 RESULTS

239 **Novel sex determining region discovered**

240 Of the 53 total selfed progeny of NM-LNF23, 9 were female and 44 were hermaphrodite (1:3

241 ratio; $\chi 2 = 2.7$; $P = 0.10$; Table 1A). Of the 41 offspring that were part of the mapping population, 8 were

242 female and 33 were hermaphrodite (1:3 ratio; $\chi 2 = 0.07$; $P = 0.79$). Using these data we unambiguously

243 mapped male sterility in these plants to a region near the 3´ end of chromosome LG6 (Figure 1).

244 Specifically, at 14 sites on 10 targeted sequence capture probes between Fvb6_34958975 and

245 Fvb6_36048692 (Figure 2A; Table 2 I), we observed a perfect match to male function. These perfectly

246 matching sites include the region Fvb6_35142k (also Sanger genotyped in other crosses; Table 2 II-A), at

247 which two SNPs (Fvb6_35142280 = Sanger site 280; and Fvb6_ 35142453 = Sanger site 453)

248 cosegregated perfectly with sex type ($P = 0.0001$). At these two sites, all 8 females are homozygous for

249   one of the two parental haplotypes ("G_T" at the Fvb6_35142k sites; Table 2 I), while all 33

250   hermaphrodites were either heterozygous or homozygous for the other parental haplotype ("T_C or

251   "G/T_C/T" at the Fvb6_35142k sites), consistent with recessive male sterility (LOD = 8.8; Figure 1).

252   From this we infer the genotype of NM-LNF23 to be *Rr* at the male function locus and its female progeny

253   to be *rr* and hermaphrodite progeny to be *RR* or *Rr* (Table 3). Adjacent to these markers, the nearest

254   mismatching markers are upstream at Fvb6_34839229 and downstream at Fvb6_36607138, and thus the

255   male sterility locus must occur in the 1.769Mb region between these two markers (Figure 2A). The SNPs

256   just outside this region, including in regions Fvb6_34763k and Fvb6_36607k (Table 2 II-A) also

257   segregated significantly with sex type ($P < 0.001$), although at Fvb6_34839229 and farther upstream, a

258   single female mismatched, while at Fvb6_36607138 and father downstream, one females and two

259   hermaphrodites mismatched.

260   The 1.768Mb region on Fvb6 that shows a perfect match to male-sterility contains 361 genes in

261   the *Fragaria vesca* Hawaii 4 reference genome (Table S3). Including 182 that are upregulated in anthers,

262   pollen or microspores, with one male gametophyte-specific gene, and several F-box proteins that have

263   been seen to be upregulated in meiotic anthers at stage 9 (Hollender et al. 2014). Other gene classes

264   showing expression changes in developing anthers included Kelch repeat and Leucine-Rich Repeat (LRR)

265   proteins. This region also includes 15 PPR genes (Table S3), 12 of which occur in the 1Mb span between

266   Fvb6_35Mb and Fvb6_36Mb, and 10 of which occur in the 0.5Mb span between Fvb6_35Mb and

267   Fvb6_35.5Mb. This cluster of PPRs is unusually dense in this region relative to the rest of the genome

268   (mean genomic PPR density = 2.9 per Mb). In fact, only one other genomic location, on Fvb5, contains a

269   higher density cluster of PPRs (Figure 1). Although none of these PPRs are orthologs of the *Raphanus*

270   *Rfo* fertility restorer and *Arabidopsis* fertility restorer-like (RFL) genes (Fujii et al. 2011), four are in the

271   PLAZA gene family HOM03D000002, and one of these (gene04450) is predicted to be mitochondrial

272   targeted (Table S4). This gene family contains the *Raphanus Rfo* gene and 25 of the 26 *Arabidopsis* RFL

273   genes, as well as PPRs at two fertility-restorer loci recently identified in a hybrid cross almond×peach

274   (Donoso et al. 2015). The RF1 locus on peach LG2 has 1 of 4 PPRs in this gene family, while the RF2

275   locus on peach LG6 has a dense cluster of 12 HOM03D000002 PPRs in 843.5 kbp. Five of these are

276   considered orthologs of the *Arabidopsis* RFL genes (Table S4).

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

278   The progeny of selfed hermaphrodite NM-LNF25 showed a 1:3 female to hermaphrodite sex ratio

279   ($\chi2 =1.42$; $P =0.23$) which is consistent with heterozygosity at a male function locus (Rr) (Table 1A;

280   Table 3). The progeny of selfed hermaphrodite NM-LNF14, however, deviated significantly from a 1:3

281   female to hermaphrodite sex ratio ($\chi2 =10.00$; $P =0.002$); leading us to propose a third, yet unmapped,

282   locus affecting male function (locus LGx). The progeny sex ratio from selfed hermaphrodite NM-LNF14

283   is consistent with a 9:7 sex ratio ($\chi2 =0.47$; $P =0.49$) that could result from selfing of a plant heterozygous

284   at two male function loci (Rr and Tt) (Table 1A; Table 3). Sequencing the progeny from this cross for the

285   Fvb6_35142k locus and evaluating the positions 35142280 and 35142453, we find that of the 15

286   hermaphrodite offspring, all either G, T or G/T, C/T at these two sites, but none of them are T_C. This

287   makes sense because T (at 35142280) and C (at 35142453) are in coupling with r at the R locus, so all 7

288   T_C offspring are rr homozygotes and therefore female. The other female offspring are presumably tt at

289   locus LGx. All other parents from NM-LNF are proposed to be TT at this locus (Table 3).

290   Crosses involving the NM-LNF hermaphrodites as sires each with three NM-LNF females

291   produced progeny sex ratios (1:1) consistent with hypothesized genotypes of $Rr\ TT$ for two of the

292   hermaphrodites, and $Rr\ Tt$ for one (NM-LNF14) and females all $rr\ TT$ (all $\chi2 < 1.6$, $P > 0.19$; Table 1A).

293   The reciprocal crosses between NM-LNF25 and NM-LNF23 also produced progeny sex ratios (1:3)

294   consistent with hypothesized genotypes of $Rr\ TT$ for both of these hermaphrodites, though one of these

295   crosses had very low seed set ($P > 0.80$; Table 1A; Table 3). The reciprocal crosses of NM-LNF14 with

296   NM-LNF23 and NM-LNF25 also produced few seeds. The one cross that produced sufficient seeds NM-

297   LNF14xLNF23 segregated in a manner consistent with the putative genotypes (1:3; $P = 0.43$).

298   We sequenced 20 progeny from female NM-LNF2 by hermaphrodite NM-LNF23 cross for the

299   Fvb6_34763k and Fvb6_35142k markers and found three SNP that segregated with sexual phenotype

300   (Table 2 II-A). At Fvb6_35142k positions 35142280 and 35142453 the *R* haplotype is T_C, and the *r*

301   haplotype is G_T, all 9 hermaphrodites are T_C and 9 of the 11 females are G_T/T_C (P <0.0003). SNP

302   position 34763440 of Fvb6_34763k also segregates with sex (P< 0.0001), and is consistent with only the

303   same two female types mismatching, although the genotype is missing for one putatively mismatched

304   female at Fvb6_34763k (Table 2 II-A). Although it would be desirable to genotype progeny in the Fvb6

305   34.8-36.6 Mbp region from the NM-LNF25 crosses, this individual was not heterozygous for any of our

306   SNP markers in this region (see Table 2 I).

307   In sum, there is a clear indication from the combined phenotypic and genotypic data that the Fvb6

308   34.8-36.6 Mbp region of the genome influences sex expression when either a female or a hermaphrodite

309   from NM-LNF population is the maternal parent. However, there apparently is a third unmapped locus

310   (LG*x*) influencing sex phenotype.

311   **Sex expression in OR-MRD crosses**

312   To evaluate whether hermaphrodites from OR-MRD also carried sex determining loci we

313   evaluated progeny sex ratios from three self-pollinated hermaphrodites. One (OR-MRD93) showed a

314   pattern consistent with a 1:3 sex ratio ($\chi2$ <0.5; P >0.50), but not 0:1 ($\chi2$ =50.6; P <0.001), one (OR-

315   MRD61) deviated significantly from 1:3 ($\chi2$ =14.3; P <0.001) but fit a 0:1 ratio (P >0.30), and one (OR-

316   MRD45) fit both 1:3 and 0:1 equally well (both $\chi2$ <0.5; P >0.30), (Table 1B). From this we inferred the

317   genotype of OR-MRD93 as a *Rr TT* heterozygote and OR-MRD61 as a *RR TT* homozygote (Table 3).

318   Crosses conducted between OR-MRD61 and the other two OR-MRD hermaphrodites produced nearly

319   exclusively hermaphrodite progeny (there was one anomalous female; Table 1B), corroborating the

320   inferred genotype of OR-MRD61 as *RR TT*. The other hermaphrodite by hermaphrodite crosses,

321   however, did not produce any (OR-MRD93 x OR-MRD61) or many viable seeds (between OR-MRD93

322 and OR-MRD45) (Table 1C). And although too few seeds (4 and 7 hermaphrodite progeny) were

323 produced from OR-MRD93 x OR-MRD45 reciprocal crosses to differentiate between 1:1 or 1:3 ratios,

324 crosses between OR-MRD45 and NM-LNF plants (see discussion below; Table 1C,D) clarify the

325 genotype of OR-MRD45 as *RR* at the LG6 locus. In addition, because the OR-MRD45 self-cross

326 produced a few female progeny (Table 1B), we deduce the genotype of OR-MRD45 as *RR Tt* (Table 3).

327   To determine whether sex of progeny from these hermaphrodites is determined by the same region

328 as in NM-LNF hermaphrodites we genotyped progeny from OR-MRD93self at the LG6 markers (Table 2

329 II-B). SNPs at position 34763379 on Fvb6_34763k and at position 35142492 on Fvb6_35142k both

330 segregated perfectly with sexual phenotype (*P*<0.0001) with females being T_G and hermaphrodites

331 being G/T-/A/G heterozygotes or G_A homozygotes. Thus, we have evidence here that the male sterility

332 region on LG6 also influences sexual phenotype in the OR-MRD population when hermaphrodites are the

333 dam. The fact that the same genomic region determines sex in both OR-MRD93 with A mitotype and

334 NM-LNF with F mitotype (Table S1) suggests that the same *R* locus is responsible for sex phenotype in

335 both cytoplasmic backgrounds. However, different PPR genes among the 15 at this locus could be the

336 functional restorer in different cytoplasmic backgrounds.

337   When all three OR-MRD hermaphrodites were used as sires on three different females from OR-

338 MRD, all produced 1:1 female to hermaphrodite ratios (all χ2 <2.63; *P* >0.11; Table 1B). Such patterns

339 are consistent with the OR-MRD30×OR-MRD60 cross mapped in Tennessen et al. (2013) where a

340 dominant allele in coupling with male sterility on the LG4 chromosome determines sex. Females were

341 thus inferred to be *MSmf* whereas hermaphrodites to be *mfmf* at this locus (Table 3).

342   Finally, genotyping the hermaphrodite progeny from the OR-MRD61 x OR-MRD93 cross showed

343 no SNP segregating with sexual phenotype as expected from a *Rr* x *RR* cross (Table 2 II-B; Table 3). It

344 also suggests that these sex determiners are allelic and function both on the A mitotype and C mitotype

345   cytoplasm (OR-MRD93 and OR-MRD61, respectively) in the absence of the influence of *MS* at the LG4

346   locus.

347   **Inter-population crosses to evaluate interactions between LG4 and LG6 sex determining regions**

348   To evaluate whether maternal control of sex at LG4 in OR-MRD extended to NM-LNF sires we

349   performed inter-population crosses on all three OR-MRD females. Eight out of the nine crosses produced

350   1:1 sex ratios ($\chi 2 < 2.57$; $P > 0.22$); the remaining cross (OR-MRD27×NM-LNF23) was skewed toward

351   females ($\chi 2 = 4.77$; $P = 0.03$) (Table 1C).

352   For three of these crosses we genotyped progeny for LG6 and LG4 markers (Table 2 II-C). None

353   of the segregating SNPs on LG6 segregated with sexual phenotype in any cross ($P = 0.67$-1), but all three

354   had SNPs at LG4 Fvb4_30092k positions 30092575 and 30092582 segregating with sex in the progeny (P

355   <0.0001). All female progeny were heterozygotes G/C_A/C whereas hermaphrodites were G_G

356   homozygotes in families of OR-MRD30×NM-LNF23 and OR-MRD90×NM-LNF23 (Table 2 II-C). In

357   the OR-MRD30×NM-LNF14 family all female progeny were A homozygotes whereas hermaphrodites

358   were A/C heterozygotes at the LG4 Fvb4_30092k position 30092575. Taken together these results are

359   consistent with epistasic dominance of the *MS* allele at the LG4 locus over the *R* allele at the LG6 locus in

360   determining sexual phenotype when the maternal parent is a female from OR-MRD.

361   To evaluate whether the sex-determiners at the LG6 locus of hermaphrodites from OR-MRD and

362   NM-LNF are allelic (e.g., alleles at the same locus) and determine sex in the absence of the *MS* allele, we

363   assessed progeny from OR-MRD hermaphrodite dams with NM-LNF hermaphrodite sires (Table 1C). No

364   seed was produced when NM-LNF14 was a father, but crosses with either NM-LNF23 and NM-LNF25 as

365   a sire produced progeny sex ratios that conformed to expectations ($\chi 2 < 2.6$; $P > 0.22$ Table 1C) based on

366   inferred genotypes (Table 3). These also provide additional support for OR-MRD45 as *RR* (and not *Rr*, as

367   all 3 crosses deviated significantly from 1:1 $P < 0.0001$).

368   To follow up on these phenotypic findings, we genotyped the progeny from OR-MRD93 and OR-

369   MRD61when pollinated by NM-LNF23. In the first cross, all four SNPs on Fvb_35142k segregated

370   significantly with sex type (P =0.0002-0.05; Table 2 IIC), indicating that R alleles in both parents

371   influence male fertility. In the second cross, as expected if OR-MRD61 is RR, none of the LG6 SNPs

372   segregated with sex and all progeny are hermaphrodite.

373   To further explore the interaction between LG6 regions we assessed inter-population crosses

374   between NM-LNF females and OR-MRD hermaphrodites (Table 1C). When crossed to OR-MRD45 and

375   OR-MRD61, all three LNF females produced 0:1 sex ratio in the progeny as expected if the dam is rr and

376   both the sires are RR at LG6 and $mfmf$ at LG 4 (Table 3). The exception, however, was that when OR-

377   MRD93 was the sire all three NM-LNF females produced sex ratios that deviated significantly from the

378   1:1 expectation (all $P < 0.01$; Table 1C). In fact, for all three crosses the female: hermaphrodite ratios fit a

379   2:1 (all $P >0.80$), potentially indicating a lethal genotype. To explore this hypothesis we genotyped the

380   progeny from the NM-LNF2×OR-MRD93 cross and found the LG6 SNPs at position 34763379 on

381   Fvb6_34763k and at position 35142492 on Fvb6_35142k both segregated perfectly with sexual

382   phenotype ($P<0.0001$), females T_G and hermaphrodites G/T_A/G as predicted by the identified sex

383   locus. Moreover, we recover all four expected genotypes in the offspring (consider both the maternally

384   segregating genotypes (e.g. at Fvb6_34763308), and paternally segregating genotypes (e.g. at

385   Fvb6_34763379): G/T_T, T_T, G/T_G/T, and T_G/T). Thus, the data are consistent with an unlinked

386   locus causing 25% of offspring who are also Rr hermaphrodites to die. Inter-population crosses between

387   NM-LNF hermaphrodites as dams and OR-MRD hermaphrodites as sires did not produce many seeds

388   (Table 1C). However, the fact that the cross between NM-LNF14 and OR-MRD45 produced two females

389   out of four total progeny is consistent their inferred heterozygous genotypes at the unmapped locus x (e.g.

390     $mfmf\ Rr\ Tt$ and $mfmf\ RR\ Tt$, respectively; Table 3).

391   DISCUSSION

392   The work presented here combined with evidence from Tennessen et al. (2013), Stanley et al. (2015) and

393   Govindarajulu et al. (2015), represents the first genomic evidence of both CMS and nuclear genes

394   (fertility restorers) in a wild gynodioecious species. Moreover, demonstration of population variation in

395   sex determiners provides insight into the potential complexity of cyto-nuclear gynodioecy in the wild and

396   provides a road map for new hypotheses for the evolution of sex chromosomes in the related octoploid

397   *Fragaria*.

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

399   We have genetically mapped two unlinked epistatically interacting loci that determine sexual phenotype

400   in wild populations of gynodioecious *F. vesca* subsp. *bracteata*. A single dominant *R* allele at the novel

401   locus fine-mapped on LG6 restores fertility in the absence of *MS* allele at the LG4 locus previously

402   identified by Tennessen et al. (2013). This newly identified LG6 34.8-36.6 Mbp region coincides with

403   the second most dense cluster of PPR in the *F. vesca* reference genome and houses 361 genes. Four of the

404   15 PPRs (Table S3) can be considered top candidates for the fertility restorer, due to their presence in the

405   HOM03D000002 gene family that contains the radish *Rfo* gene (the only confirmed CMS-interacting

406   fertility-restorer in the rosid clade, Chen and Liu 2014), *Arabidopsis* RFL genes (Fujii et al. 2011) and

407   putative RF loci in peach (Donoso et al. 2015). One of the four *F. vesca* genes is predicted to be

408   mitochondrial targeted (Proost et al. 2015) and the other three do not have an organellar prediction. Other

409   genes potentially involved in male function (upregulated in meiotic anthers; Hollender et al. 2014) also

410   reside in this region, although not more so than the rest of the genome. Nonetheless it is intriguing that

411   several important gene domains (F-box, Kelch repeat and LRR) occur here and in the LG4 male sterility

412   region (*MSmf*) (Tennessen et al. 2013). Additionally, the action of miRNAs have been proposed as a route

413   to sex determination (Akagi et al. 2014; Fagegaltier et al. 2014), and seven genes that are miRNA targets

414   reside in the LG6 region (Table S3). Only one of these is a PPR, but it is cp-targeted and not in the

415 HOM03D000002 gene family. Furthermore, none of these seven genes are targets for the novel F-box

416 associated *Fragaria* miRNA family recently described in *F. vesca* (Xia et al. 2015).

417 A third locus (LG*x*) is also present and segregation analysis suggests it is unlinked to the LG6

418 locus (i.e., 9:7 ratio in the progeny of LNF14self, Table 1,2). Extensive mapping of genome-derived

419 sequences around the sex-determining regions in several crosses revealed that the *MS* allele was

420 epistatically dominant to *R* at LG6 and *T* at LG*x*. In fact, segregation results suggest that all loci interact,

421 but functional studies will be needed to determine if and how they interact molecularly and biochemically.

422 Toward this end, we propose a conceptual model of these gene interactions (Figure 3) that draws

423 on evidence from the current study and Tennessen et al. (2013) along with knowledge of the presence of a

424 CMS-like *atp8-orf225* in the mitochondria (Table S1; Stanley et al. 2015; Govidarajulu et al. 2015). We

425 hypothesize that a CMS locus that disrupts pollen development by production of a toxic protein or energy

426 deficiencies (reviewed by Chen and Liu 2014) exists ancestrally. Male function is then restored by one or

427 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

428 dominance of the restoring allele are consistent with data and theory of restorers of CMS (Delph et al.

429 2007; Chen and Liu 2014). The LG4 locus, which is not associated with annotated PPRs (Tennessen et al.

430 2013) is an inhibitor and a single dominant allele *MS* is sufficient to block the action of the LG6 restorer.

431 Although of unknown function the LG4 locus may be a necessary cofactor required by the PPR or another

432 mitochondrial sorting gene involved in nuclear-mitochondrial cross-talk (figure 3 in Chen and Liu 2014).

433 The *MS* allele is both dominant to *mf* and epistatically dominant to *R.* Homozygosity for the *t* allele at the

434 third unmapped locus on LG*x* also leads to disruption of pollen production, and this locus could represent

435 an independently evolved restorer (*T* allele), as cyto-nuclear theory predicts successive 'waves' of

436 restoration (reviewed in Deph et al. 2007), and multiple restorer systems are common, with some loci at

437 or near fixation (e.g., Garraud et al. 2011; Caruso and Case 2013). This is especially true when there is

438 polymorphism in CMS (e.g., Dudle et al. 2001; Garraud et al. 2011) as there appears to be in *F. vesca*

439   subsp. *bracteata*.   (Table S1; Stanley et al. 2015). If so, the products of the LG4 locus might also be able

440   to block its action. Alternatively, the LG*x* locus could be a novel recessive male sterility locus that acts

441   independently of the others (as seen in Irkaeva et al. 1993). By this same token, the LG6 locus may also

442   be a recessive male sterility locus that interacts with products of the LG4 locus unrelated to CMS, but this

443   possibly seems less likely given the general predominance of multiple nuclear-cytoplasmic interactions in

444   gynodioecy and the specific evidence of both PPRs and a CMS-like *atp8-orf225* in gynodioecious *F.*

445   *vesca* subsp. *bracteata*.

446   The layers of interaction in this model may seem complex, but likely still reflect a simplification

447   of the web of genetic mechanisms involved in cyto-nuclear gynodioecy in the wild. Much of what we

448   know about these cyto-nuclear interactions comes from hybrids of hermaphroditic wild species (e.g.,

449   *Mimulus guttattus*, Barr and Fishman 2011; *Arabidopsis lyrata,* Aalto et al. 2013) or crop species where

450   restorers are studied in fixed genetic backgrounds (reviewed in Chen and Liu 2014). In fact it has been

451   proposed that one advantage of crop systems for understanding elements of cyto-nuclear gynodioecy is

452   the reduced variation (Delph et al. 2007). However, these simplified systems may not represent the reality

453   of sex determination in wild gynodioecious species which are subject to multiple and varied forces of

454   drift, selection and gene flow. Our genomic evidence of both cytoplasmic and nuclear players join a small

455   number of studies that have genetically mapped nuclear sex-determining loci in wild gynodioecious plants

456   (Touzet et al. 2004), or have performed extensive crosses to demonstrate the existence of both nuclear and

457   cytoplasmic contributors (e.g., Van Damme et al. 2004, Garraud et al. 2011). The next steps will involve

458   functionally verifying the genetic interactions. Several types of approaches will be required for this and to

459   fully test the working model we have proposed-- including finer mapping of nuclear sex loci and

460   transformation of both candidate nuclear restorer genes and the CMS-like *atp8-orf225* in the

461   mitochondrial genome (see Chen and Liu 2014).

462   **Evolutionary inferences from population differentiation in sex determination**

463   We observed population variation in the frequency of sex determiners, leading especially to

464   differences in the genotypic composition of females. All examined females in OR-MRD have dominant

465   male sterility (*MSmf R_ TT* genotypes), whereas all the females in NM-LNF have only recessive male

466   sterility (*mfmf rr TT* genotypes). It should be noted that *mfmf rr TT* females are possible from selfed

467   hermaphrodites in OR-MRD (i.e., OR-MRD93), but would be expected to be rare given the low

468   frequency of *r* in this population. More generally, if our sampling is representative of the frequency of

469   genotypes in the populations, then *r* is more frequent (1:11 vs. 9:3; $P < 0.01$) in NM-LNF than OR-MRD,

470   while *MS* is more frequent in OR-MRD than NM-LNF, though the latter is not statistically significant (3:9

471   vs. 0:12; $P > 0.05$). This spatial genetic structure of sex determiners could provide insights into past

472   evolutionary dynamics. Cyto-nuclear coevolution can lead to multiple restorer loci and polymorphism at

473   both nuclear and mitochondrial genes (e.g., Garruad et al. 2011), and local processes are expected to

474   result in spatial structure for restorer and CMS haplotypes (Bailey and McCauley 2005). So it is

475   especially intriguing that the nuclear differences mirror the strong population differentiation in the

476   mitotypes between the study populations. The F mitotype is only found in NM-LNF, whereas the B and C

477   mitotypes found in OR-MRD are more widespread (Stanley et al. 2015). If, as our results suggest, LG6

478   and LG*x* are restorer loci then *R* and *T* can restore all three mitotypes, and this would imply a greater

479   spatial range for restorers than mitotypes. Such a pattern has been inferred from crosses with differing

480   cytotypes in *Plantago coronopus* (Van Damme et al. 2004) and *Silene nutans* (Garruad et al. 2011). In

481   contrast, the dominant *MS* allele at LG4 appears restricted to the northwest (OR-MRD, in this study and

482   Tennessen et al. 2014; and possibly 'HP' population in northern California, Ahmadi and Bringhurst

483   1991), perhaps indicating that it evolved after the LG6 and LG*x* loci, consistent with our multi-layered

484   hypothesis for sex determination (Figure 3). The interactions between these genes likely contribute to

485   high variation in sex ratios across space in *F. vesca* subsp. *bracteata* (0-46% females; Stanley et al. 2015).

486   More crosses between populations of varying distances across the range (e.g., Bailey and McCauley

487   2005) and/or targeted resequencing of the sex determining regions would shed light on the wider

488 geographic context for the players in the *F. vesca* subsp. *bracteata* sex determination system and our sex-

489 determination hypothesis. This combined with fitness consequences of specific genotypes is needed to

490 test evolutionary hypotheses for maintenance of multiple restorers and mitotypes in this species (e.g.,

491 Caruso et al. 2012).

492 The genetic differentiation of NM-LNF and OR-MRD is also evident in Stanley et al. (2015)

493 where STRUCTRE analysis based on nuclear markers distinguished OR-MRD as belonging to cluster 1

494 while NM-LNF to cluster 2 (more closely allied with subspecies *F. vesca* subsp. *americana*). Similar

495 distinction was seen in the chloroplast. Such differentiation could explain not only the differences in

496 prevailing sex determiners but also the asymmetry in success of the interpopulation HxH crosses. Crosses

497 with a OR-MRD hermaphrodite as a sire on a NM-LNF hermaphrodite as a dam produced fewer seeds

498 than the reciprocal, potentially reflecting cyto-nuclear incompatibilities (Rieseberg and Blackman 2010)

499 that interact with the sex determiners because similar deficits were not see when females were the dams.

500 Alternatively these could reflect exposure of costs associated with restoration that may be complex

501 (Caruso et al. 2012). It is notable in this regard that NM-LNF14 performed poorly as a sire and a dam, and

502 also was the only plant inferred to be heterozygous at both the putative fertility-restorer loci (Table 3).

503 **Relationship to sex determination octoploids and other Fragaria**

504 The present results can be considered in the context of the known sex determining regions in the

505 two octoploid descendants of *F. vesca* subsp. *bracteata*. In both *F.viginiana* and *F. chiloensis,* linked

506 male and female sterility loci reside on a chromosome in homeologous group VI (LG VI-Av in *F.*

507 *chiloensis,* and VI-B2 *F.viginiana;* Goldberg et al. 2010; Spigler et al. 2011; Tennessen et al. 2014). Thus,

508 it is intriguing to identify a locus on LG6 in *F. vesca* subsp. *bracteata* that also affects sexual expression.

509 The fact that the sex determining region in *F. vesca* subsp. *bracteata* is near the 3' end of LG6, as is the

510 sex determining region of *F. chiloensis* (Goldberg et al. 2010), raises the possibility they are homologous.

511 However, the *F. vesca* subsp. *bracteata* LG6 locus (*Rr*) codes for recessive male sterility, whereas the LG

512 VI-Av locus in *F. chiloensis* has dominant male sterility. This difference alone might suggest they are not

513 the same locus, but it could also reflect a turnover in the sex determining chromosome. Heterogametic

514 transitions are theoretically possible and can transition from male heterogamety to female heterogamety

515 when release from mutational load is followed by sexually antagonistic selection (Blaser et al. 2014).

516 If, on the other hand, the *F. vesca* subsp. *bracteata* LG6 locus and the LG VI-Av sex determining

517 locus in *F. chiloensis* are entirely unrelated, the discovery of the *F. vesca* subsp. *bracteata* LG6 locus

518 provides additional evidence that LG6 is predisposed to be a sex determining chromosome, as initially

519 proposed by Spigler et al. (2011) with respect to *F. viginiana* and *F. chiloensis*. Some autosomes are

520 thought to be prone to hosting sex determining regions because they house many genes that can affect

521 male function, or because a chromosome that already has been involved in sex determination is more

522 likely to seize back this role in the future (Graves and Peichel 2010, Blaser et al. 2014). This could be

523 facilitated by transposition of genes already involved in sex expression (e.g., Hughes et al. 2015), making

524 the LG4 locus a candidate for such transposition during the origin of the octoploids producing a gene

525 complex in *F. chiloensis*. Finer mapping and deep resequencing of the sex determining region in *F.*

526 *chiloensis* is underway to test these ideas. In either event, the present results suggest great potential for

527 sexual lability in *Fragaria,* as male sterility evolves frequently, independently and via different genetic

528 mechanisms. Indeed, recent phylogenetic character state transition analysis detects frequent transitions to

529 and from separate sexes (dioecy) in the *Fragaria* (Goldberg et al. unpublished). The present study allows

530 us to speculate that this lability is facilitated by ancient evolution of CMS and repeated evolution of

531 restorers and/or other suppressors of male function, and turnovers in sex determining chromosomes.

532   ACKNOWLEDGEMENTS

536 TLA and DEB 1020271 to AL), UPitt Mellon and NSF Predoctoral Fellowships to MK, and Norman H.

537 Horowitz Fellowship, Pennsylvania Space Grant Consortium Research Scholarship and Fellowships

538 funded by Howard Hughes Medical Institute to RD.

539 <div align="center">REFERENCES</div>

540 Aalto, E. A., H.-P. Koelewijn, and O. Savolainen. 2013. Cytoplasmic male sterility contributes to hybrid

541   incompatibility between subspecies of *Arabidopsis lyrata*. G3 **3**:1727-1740.

542 Ahmadi, H. and R. S. Bringhurst 1989. Genetics of sex expression in *Fragaria* species. American Journal

543   of Botany **78**:504-514.

544 Akagi, T., I. M. Henry, R. Tao, and L. Comai. 2014. A Y-chromosome–encoded small RNA acts as a sex

545   determinant in persimmons. Science **346**:646-650.

546 Bachtrog, D., J. E. Mank, C. L. Peichel, M. Kirkpatrick, S. P. Otto, T.-L. Ashman, M. W. Hahn, J. Kitano,

547   I. Mayrose, R. Ming, N. Perrin, L. Ross, N. Valenzuela, and J. C. Vamosi. 2014. Sex

548   determination: why so many ways of doing it? PLoS biology **12**:e1001899.

549 Bailey, M. F. and L. F. Delph. 2007. A field guide to models of sex-ratio evolution in gynodioecious

550   species. Oikos **116**:1609-1617.

551 Bailey, M. F. and D. E. McCauley. 2005. Offspring sex ratio under inbreeding and outbreeding in a

552   gynodioecious plant. Evolution **59**:287-295.

553 Barr, C. M. and L. Fishman. 2011. Cytoplasmic male sterility in *Mimulus* hybrids has pleiotropic effects

554   on corolla and pistil traits. Heredity **106**:886-893.

555 Blaser, O., S. Neuenschwander, and N. Perrin. 2014. Sex-chromosome turnovers: the hot-potato model.

556   The American Naturalist **183**:140-146.

557 Caruso, C. M. and A. L. Case. 2013. Testing models of sex ratio evolution in a gynodioecious plant:

558   female frequency covaries with the cost of male fertility restoration. Evolution **67**:561-566.

559 Caruso, C. M., A. L. Case, and M. F. Bailey. 2012. The evolutionary ecology of cytonuclear interactions

560   in angiosperms. Trends in Plant Science **17**:638-643.

561 Charlesworth, B. and D. Charlesworth. 1978. A model for the evolution of dioecy and gynodioecy. The

562   American Naturalist **112**:975-997.

563 Chen, L. and Y.-G. Liu. 2014. Male sterility and fertility restoration in crops. Annual Review of Plant

564   Biology **65**:579-606.

565 Crossman, A. and D. Charlesworth. 2014. Breakdown of dioecy: Models where males acquire cosexual

566   functions. Evolution **68**:426-440.

567 Cui, X., Q. Wang, W. Yin, H. Xu, Z. A. Wilson, C. Wei, S. Pan, and D. Zhang. 2012. PMRD: a curated

568   database for genes and mutants involved in plant male reproduction. BMC Plant Biology **12**:215-

569   224.

570 Dalton, R., M. Koski, and T.-L. Ashman. 2013. Maternal sex effects and inbreeding depression under

571   varied environmental conditions in gynodioecious *Fragaria vesca* subsp. *bracteata.* Annals of

572   Botany **112**:613-621.

573 Darwish, O., Shahan, R., Liu, Z., Slovin, J. P., & Alkharouf, N. W. (2015). Re-annotation of the

574   woodland strawberry (*Fragaria vesca*) genome. BMC genomics, 16(1), 29.

575 Delph, L. F., P. Touzet, and M. F. Bailey. 2007. Merging theory and mechanism in studies of gynodioecy.

576   Trends in Ecology & Evolution **22**:17-24.

577 Diggle, P. K., V. S. Di Stilio, A. R. Gschwend, E. M. Golenberg, R. C. Moore, J. R. Russell, and J. P.

578   Sinclair. 2011. Multiple developmental processes underlie sex differentiation in angiosperms.

579   Trends in Genetics **27**:368-376.

580 Donoso, J. M., I. Eduardo, R. Picañol, I. Batlle, W. Howad, M. J. Aranzana, and P. Arús. 2015. High-

581   density mapping suggests cytoplasmic male sterility with two restorer genes in almond× peach

582   progenies. Horticulture Research **2**. 1-10.

583 Dudle, D. A., P. Mutikainen, and L. F. Delph. 2001. Genetics of sex determination in the gynodioecious

584   species *Lobelia siphilitica*: evidence from two populations. Heredity **86**:265-276.

585 Dufaÿ, M., P. Champelovier, J. Käfer, J.-P. Henry, S. Mousset, and G. Marais. 2014. An angiosperm-

586   wide analysis of the gynodioecy–dioecy pathway. Annals of Botany **114**:539-548.

587 Dufaÿ, M., J. Cuguen, J.-F. Arnaud, and P. Touzet. 2009. Sex ratio variation among gynodioecius

588   populations of sea beet: Can it be explained by negative frequency-dependent selection? Evolution

589   **63**:1483-1497.

590 Dufaÿ, M., P. Touzet, S. Maurice, and J. Cuguen. 2007. Modelling the maintenance of male-fertile

591   cytoplasm in a gynodioecious population. Heredity **99**:349-356.

592 Dufresnes, C., Y. Bertholet, J. Wassef, K. Ghali, R. Savary, B. Pasteur, A. Brelsford, B. Rozenblut

593   Kościsty, M. Ogielska, and M. Stöck. 2014. Sex-chromosome differentiation parallels postglacial

594   range expansion in European tree frogs (*Hyla arborea*). Evolution **68**:3445-3456.

595 Ehlers, B. K. and T. Bataillon. 2007. 'Inconstant males' and the maintenance of labile sex expression in

596   subdioecious plants. New Phytologist **174**:194-211.

597 Fagegaltier, D., A. König, A. Gordon, E. C. Lai, T. R. Gingeras, G. J. Hannon, and H. R. Shcherbata.

598   2014. A genome-wide survey of sexually dimorphic expression of Drosophila miRNAs identifies

599   the steroid hormone-induced miRNA let-7 as a regulator of sexual identity. Genetics **198**:647-668.

600 Frank, S. 1989. The evolutionary dynamics of cytoplasmic male sterility. The American Naturalist

601   **133**:345-376

602 Fujii, S., C. S. Bond, and I. D. Small. 2011. Selection patterns on restorer-like genes reveal a conflict

603   between nuclear and mitochondrial genomes throughout angiosperm evolution. Proceedings of the

604   National Academy of Sciences **108**:1723-1728.

605 Garraud, C., B. Brachi, M. Dufaÿ, P. Touzet, and J. Shykoff. 2011. Genetic determination of male sterility

606   in gynodioecious *Silene nutans*. Heredity **106**:757-764.

607   Gobron, N., C. Waszczak, M. Simon, S. Hiard, S. Boivin, D. Charif, A. Ducamp, E. Wenes, and F. Budar.

608   2013. A cryptic cytoplasmic male sterility unveils a possible gynodioecious past for *Arabidopsis*

609   *thaliana*. PLOS ONE DOI: 10.1371/journal.pone.0062450.

610   Goldberg, E. E., S. P. Otto, J. Vamosi, I. Mayrose, R. Ming, and T.-L. Ashman. in prep. Straying from the

611   pathway: Species-level phylogenetic analysis reveals transitions to, and from, dioecy. MS in prep.

612   Goldberg, M., R. Spigler, and T-L. Ashman. 2010. Comparative genetic mapping points to different sex

613   chromosomes in sibling species of wild strawberry (Fragaria). Genetics **186**:1425-1433.

614   Govindarajulu, R., M. Parks, J. A. Tennessen, A. Liston, and T.-L. Ashman. 2015. Comparison of

615   nuclear, plastid, and mitochondrial phylogenies and the origin of wild octoploid strawberry

616   species. American Journal of Botany **102**:544-554.

617   Graves, J. and C. Peichel. 2010. Are homologies in vertebrate sex determination due to shared ancestry or

618   to limited options? Genome Biology **11**:205.

619   Hollender, C. A., A. C. Geretz, J. Slovin, and Z. Liu. 2012. Flower and early fruit development in a

620   diploid strawberry, *Fragaria vesca*. Planta **235**:1123-1139.

621   Hollender, C. A., C. Kang, O. Darwish, A. Geretz, B. F. Matthews, J. P. Slovin, N. W. Alkharouf, and Z.

622   Liu. 2014. Floral transcriptomes in woodland strawberry uncover developing receptacle and anther

623   gene networks. Plant Physiology **165**:1062-1075.

624   Hughes, J. F., H. Skaletsky, N. Koutseva, T. Pyntikova, and D. C. Page. 2015. Sex chromosome-to-

625   autosome transposition events counter Y-chromosome gene loss in mammals. Genome Biology

626   **16**:104-113.

627   Irkaeva, N. M., L. V. Lebedeva, and D. S. Shaziyatova. 1993. Genetics of male sterility in the strawberry.

628   Genetika **29**:1485-1491.

629   Jacobs, M. S. and M. J. Wade. 2003. A synthetic review of the theory of gynodioecy. The American

630   Naturalist **161**:837-851.

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

632   Niwot, Colorado.

633   Klaas, A. L. and M. S. Olson. 2006. Spatial distributions of cytoplasmic types and sex expression in

634   Alaskan populations of *Silene acaulis*. International Journal of Plant Sciences **167**:179-189.

635   Lewis, D. 1941. Male sterility in natural populations of hermaphrodite plants. New Phytologist **40**:56-63.

636   Li, J., M. H. Koski, and T.-L. Ashman. 2012. Functional characterization of gynodioecy in *Fragaria*

637   *vesca* ssp. *bracteata* (Rosaceae). Annals of Botany **109**:545-552.

638   Liston, A., R. Cronn, and T.-L. Ashman. 2014. Fragaria: A genus with deep historical roots and ripe for

639   evolutionary and ecological insights. American Journal of Botany **101**:1686-1699.

640   Margarido, G., A. Souza, and A. Garcia. 2007. OneMap: software for genetic mapping in outcrossing

641   species. Hereditas **144**:78-79.

642   Maurice, S. and T. H. Fleming. 1995. The effect of pollen limitation on plant reproductive systems and

643   the maintenance of sexual polymorphisms. Oikos **74**:55-60.

644   Njuguna, W., A. Liston, R. C. Cronn, T.-L. Ashman, and N. V. Bassil. 2013. Insights into phylogeny, sex

645   function and age of *Fragaria* based on whole chloroplast genome sequencing. Molecular

646   Phylogenetics and Evolution **66**:17-29.

647   Proost, S., M. Van Bel, D. Vaneechoutte, Y. Van de Peer, D. Inzé, B. Mueller-Roeber, and K.

648   Vandepoele. 2015. PLAZA 3.0: an access point for plant comparative genomics. Nucleic Acids

649   Research **43**: D974-D981.

650   Renner, S. S. 2014. The relative and absolute frequencies of angiosperm sexual systems: Dioecy,

651   monoecy, gynodioecy, and an updated online database. American Journal of Botany **101**:1588-

652   1596.

653   Rieseberg, L. H. and B. K. Blackman. 2010. Speciation genes in plants. Annals of Botany **106:** 439-455.

654   Russell, J. and J. Pannell. 2015. Sex determination in dioecious *Mercurialis annua* and its close diploid

655   and polyploid relatives. Heredity **114**: 262-271

656 Schultz, S. T. 1994. Nucleo-cytoplasmic male sterility and alternative routes to dioecy. Evolution **48**:

657   1933-1945.

658 Shulaev, V., D. J. Sargent, R. N. Crowhurst, T. C. Mockler, O. Folkerts, A. L. Delcher, P. Jaiswal, K.

659   Mockaitis, A. Liston, and S. P. Mane, et al. 2010. The genome of woodland strawberry (*Fragaria*

660   *vesca*). Nature Genetics **43**: 109-116.

661 Spigler, R. B. and T-L. Ashman. 2012. Gynodioecy to dioecy: are we there yet? Annals of Botany **109**:

662   531- 543.

663 Spigler, R. B., K. S. Lewers, and T-L. Ashman. 2011. Genetic architecture of sexual dimorphism in a

664   subdioecious plant with a proto-sex chromosome. Evolution **65**:1114-1126.

665 Stanley, L., N. J. Forrester, R. Govindarajulu, A. Liston, and T-L. Ashman. 2015. Geographic patterns of

666   genetic variation in three genomes of North American diploid strawberries with special reference

667   to *Fragaria vesca* subsp. *bracteata*. Botany **93**: 1–16. dx.doi.org/10.1139/cjb-2015-0031

668 Staudt, G. 1999. Systematics and geographic distribution of the American strawberry species. Taxonomic

669   studies in the genus *Fragaria* (Rosaceae: Potentilleae). University of California Press, Berkeley.

670 Tennessen, J. A., R. Govindarajulu, T.-L. Ashman, and A. Liston. 2014. Evolutionary origins and

671   dynamics of octoploid strawberry subgenomes revealed by dense targeted capture linkage maps.

672   Genome Biology and Evolution **6**: 3295-3313.

673 Tennessen, J. A., R. Govindarajulu, A. Liston, and T.-L. Ashman. 2013. Targeted sequence capture

674   provides insight into genome structure and genetics of male sterility in a gynodioecious diploid

675   strawberry, *Fragaria vesca* ssp. *bracteata* (Rosaceae). G3 **3**: 1341-1351.

676 Touzet, P., N. Hueber, A. Bürkholz, S. Barnes, and J. Cuguen. 2004. Genetic analysis of male fertility

677   restoration in wild cytoplasmic male sterility G of beet. Theoretical and Applied Genetics **109**:

678   240-247.

679 Ubeda, F., M. M. Patten, and G. Wild. 2015. On the origin of sex chromosomes from meiotic drive. Proc.

680   R. Soc. B http://dx.doi.org/10.1098/rspb.2014.1932.

681   Van Damme, J., M. Hundscheid, S. Ivanovic, and H. Koelewijn. 2004. Multiple CMS–restorer gene

682   polymorphism in gynodioecious *Plantago coronopus*. Heredity **93**: 175-181.

683   Vicoso, B. and D. Bachtrog. 2013. Reversal of an ancient sex chromosome to an autosome in Drosophila.

684   Nature **499**:332-335.

685   Wang, K., X. Peng, Y. Ji, P. Yang, Y. Zhu, and S. Li. 2013. Gene, protein, and network of male sterility

686   in rice. Frontiers in Plant Science **4**:1-10.

687   Xia, R., S. Ye, Z. Liu, B. Meyers, and Z. Liu. 2015. Novel and recently evolved miRNA clusters regulate

688   expansive F-box gene networks through phasiRNAs in wild diploid strawberry. Plant Physiology

689   http://dx.doi.org/10.1104/pp.15.00253

690

691 **Tables**

692 Table 1 Phenotypic sex ratios from intra- (within population) and inter-population (between populations)

693 crosses of gynodioeicous *Fragaria vesca* subsp. *bracteata*. Three types of crosses are presented: selfed

694 hermaphrodites (H-self), female dam by hermaphrodite sire (F x H) and hermaphrodite dam by

695 hermaphrodite sire (H x H) from within (A) N-LNF or (B) OR-MRD populations or between NM-LNF

696 and OR-MRD populations (C). For each cross the dam, sire and total number of progeny scored for sexual

697 phenotype are given, as well as the number of female and hermaphrodite progeny, the predicted sex ratio

698 based on Table 3 genotypes of the parents. χ^2 statistics and P values from χ^2 goodness of fit tests.

699 *Crosses performed yielded no seed or none germinated. ^Alternative hypothesis for expected sex ratio

700 tested and reported in text. Na: Too few progeny (family size <10) to conduct a statistical test.

701

A. Within NM-LNF crosses

Cross type	Dam	Sire	Total progeny sexed	Female	Hermaphrodite	Expected ratio	χ^2 value	P-value
H-SELF	LNF23	LNF23	53	9	44	1:3	0.10	2.70
	LNF25	LNF25	53	17	36	1:3	1.42	0.23
	LNF14	LNF14	30	15	15	1:3^	10.00	0.002
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^	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^	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^	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^	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^	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
	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

702

703

704

705

706

707

708

709

710

711

712

713

714

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

I. Parents

Groups: LG6 — Fvb6_34763k (259–760), Fvb6_35142k (087–492), Fvb6_36607k (138–180); LG4 — Fvb4_30092k (575–582)

Genotype ID	Sex	34763 259	34763 308	34763 320	34763 379	34763 440	34763 505	34763 655	34763 760	35142 087	35142 280	35142 453	35142 492	36607 138	36607 180	30092 575	30092 582
NM-LNF23	H	G	G/T	A	T	G/T	G/T	C	T	A/G	G/T	C/T	G	A/G	C	C	C
NM-LNF14	H	G	T	A	T	G/T	G	C	A/T	A	G/T	C/T	G	G	C	A/C	G/T
NM-LNF2	F	G	G/T	A	T	G	G/T	C	A/T	A/G	T	C	G	A/G	C	C	C
OR-MRD30	F	G	T	A/G	T	G	G	G/C	T	A/G	T	C	G	G	C/T	G/C	A/C
OR-MRD90	F	A/G	T	A	T	G	G	C	T	G	T	C	G	G	C	G/C	A/C
OR-MRD93	H	G	T	A	G/T	G	G	C	T	G	T	C	A/G	G	C	A	G
OR-MRD61	H	.	.	.	.	.		.	.	A/G	T	C	G	G	C	A	G

II. Crosses

Dam × Sire	Sex	No. of progeny	34763 259	34763 308	34763 320	34763 379	34763 440	34763 505	34763 655	34763 760	35142 087	35142 280	35142 453	35142 492	36607 138	36607 180	30092 575	30092 582
A. Within NM-LNF																		
NM-LNF23 × NM-LNF23	F	5	-	**G**	-	-	**G**	**T**	-		.	**G**	**T**	-		**A**		
	F	1	-	**G**	-	-	**G**	**T**	-		.	**G**	**T**	-		**A/G**		
	F	1	-	**G**	-	-	**G**	**T**	-		.	**G**	**T**	-		**.**		
	F	1	-	**G/T**	-	-	**G/T**	**G/T**	-		.	**G**	**T**	-		**A**		
	H	19	-	**G/T**	-	-	**G/T**	**G/T**	-		.	**G/T**	**C/T**	-		**A/G**		
	H	2	-	**G/T**	-	-	**G/T**	**G/T**	-		.	**G/T**	**C/T**	-		**A**		
	H	1	-	**G/T**	-	-	**G/T**	**G/T**	-		.	**G/T**	**C/T**	-		**G**		
	H	9	-	**T**	-	-	**T**	**G**	-		.	**T**	**C**	-		**G**		
	H	1	-	**T**	-	-	**G/T**	**G**	-		.	**T**	**C**	-		**G**		
	H	1	-	**T**	-	-	**T**	**G**	-		.	**T**	**C**	-		**A/G**		
NM-LNF14 × NM-LNF14	F	4	-	-	-	-	*G/T*	-	-	*A/T*	-	*G/T*	*C/T*	-				
	F	4	-	-	-	-	*G*	-	-	*A*	-	*G*	*T*	-				
	F	7	-	-	-	-	*T*	-	-	*T*	-	*T*	*C*	-				
	H	10	-	-	-	-	*G/T*	-	-	*A/T*	-	*G/T*	*C/T*	-				

			1	2	3	4	5	6	7	8	9	10	11	12	13	14
NM-LNF2 × NM-LNF23	H	5	-	-	-	-	*G*	-	-	*A*	-	*G*	*T*	-		
	F	7	-	*G/T*	-	-	**G**	*G/T*	-	*A/T*	-	**G/T**	**C/T**	-		
	F	2	-	*G*	-	-	**G**	*T*	-	*T*	-	**G/T**	**C/T**	-		
	F	1	-	*G/T*	-	-	**G/T**	*G/T*	-	*T*	-	**T**	**C**	-		
	F	1	-	.	-	-	**.**	.	-	.	-	**T**	**C**	-		
	H	7	-	*G/T*	-	-	**G/T**	*G/T*	-	*T*	-	**T**	**C**	-		
	H	2	-	*T*	-	-	**G/T**	*G*	-	*A/T*	-	**T**	**C**	-		

B. Within OR- MRD

			5				12
OR-MRD93 × OR-MRD93	F	8	**T**				**G**
	H	28	**G/T**				**A/G**
	H	3	**G**				**A**
OR-MRD61 × OR-MRD93	H	2				-	*A/G*
	H	7				*A/G*	*G*
	H	2				*G*	*A/G*
	H	1				*G*	*G*

C. Between OR-MRD and NM-LNF

			1	2	3	4	5	6	7	8	9	10	11	12	13	14
OR-MRD30 × NM-LNF23	F	1	-	*G/T*	*A*	-	*G/T*	*C*	-	*G/T*	*C/T*	-	*A/G*	*C/T*	**G/C**	**A/C**
	F	1	-	*G/T*	*A*	-	*G/T*	*C*	-	.	.	-	.	.	**.**	**.**
	F	1	-	*G/T*	*A*	-	*G/T*	*C*	-	*G/T*	*C/T*	-	*A/G*	*C*	**G/C**	**A/C**
	F	3	-	*G/T*	*A/G*	-	*G/T*	*G/C*	-	*G/T*	*C/T*	-	*A/G*	*C/T*	**G/C**	**A/C**
	F	1	-	*G/T*	*A/G*	-	*G/T*	*G/C*	-	*G/T*	*C/T*	-	.	.	**G/C**	**A/C**
	F	1	-	*G/T*	*A/G*	-	*G/T*	*G/C*	-	*G/T*	*C/T*	-	*G*	*C/T*	**G/C**	**A/C**
	F	1	-	*T*	*A*	-	*G*	*C*	-	*T*	*C*	-	*A/G*	*C*	**G/C**	**A/C**
	F	2	-	*T*	*A*	-	*G*	*C*	-	*T*	*C*	-	*G*	*C*	**G/C**	**A/C**
	F	2	-	*T*	*A/G*	-	*G*	*G/C*	-	*T*	*C*	-	*G*	*C/T*	**G/C**	**A/C**
	F	1	-	*T*	*A/G*	-	*G*	*G/C*	-	*T*	.	-	.	.	**G/C**	**A/C**
	H	2	-	*G/T*	*A*	-	*G/T*	*C*	-	*G/T*	*C/T*	-	*A/G*	*C*	**G**	**G**
	H	4	-	*G/T*	*A/G*	-	*G/T*	*G/C*	-	*G/T*	*C/T*	-	*A/G*	*C/T*	**G**	**G**
	H	1	-	*T*	*A/G*	-	*G*	*G/C*	-	*T*	*C*	-	*G*	*C/T*	**G**	**G**
	H	2	-	*T*	*A/G*	-	*G*	*G/C*	-	*T*	*C*	-	*A/G*	*C/T*	**G**	**G**
	H	1	-	*T*	*A*	-	*G*	*C*	-	.	.	-	.	.	**G**	**G**

			1	2	3	4	5	6	7	8	9	10	11	12	13	14
OR-MRD30 × NM-LNF14	F	1	-	-	*A*	-	*G*	-	*C*	*T*	*A/G*	*T*	.	-	**A**	*G/T*
	F	1	-	-	*A*	-	*G*	-	*C*	*A/T*	*A/G*	*G/T*	.	-	**A**	*G*
	F	1	-	-	*A*	-	*G/T*	-	*C*	*T*	*A/G*	*T*	.	-	**A**	*G/T*
	F	1	-	-	*A*	-	*G/T*	-	*C*	*T*	*A/G*	*T*		-	**A**	*G*
	F	1	-	-	*A*	-	*G/T*	-	*C*	*T*	*A/G*	*T*	*C*	-	**A**	*G*
	F	1	-	-	*A/G*	-	*G/T*	-	*G/C*	*T*	*A*	*T*	*C*	-	**A**	*G*
	F	1	-	-	*A/G*	-	*G/T*	-	*G/C*	.	*A*	*T*	.	-	**A**	*G/T*
	F	1	-	-	*A/G*	-	*G*	-	*G/C*	*A/T*	.	.	.	-	**A**	*G/T*
	F	1	-	-	*A*	-	*G/T*	-	*C*	.	*A/G*	*T*	.	-	**A**	*G/T*
	H	3	-	-	*A*	-	*G/T*	-	*C*	*T*	*A/G*	*T*	*C*	-	**A/C**	*C/T*
	H	1	-	-	*A*	-	*G*	-	*C*	*A/T*	*A/G*	*G/T*	.	-	**A/C**	*C/T*
	H	1	-	-	*A*	-	*G*	-	*C*	.	*A*	*G/T*	*C*	-	**A/C**	*C/T*
	H	1	-	-	*A/G*	-	*G*	-	*G/C*	*A/T*	*A*	.	.	-	**A/C**	*C/T*

Cross																A/C	G/C
	H	1	-	-	*A/G*	-	*G/T*	-	*G/C*	*T*	.	.	.	-		**A/C**	**G/C**
OR-MRD90 × NM-LNF23	F	5	*A/G*	*G/T*			*G*	*G/T*								**G/C**	**A/C**
	F	5	*G*	*G/T*			*G*	*G/T*								**G/C**	**A/C**
	F	10	*A/G*	*T*			*G/T*	*G*								**G/C**	**A/C**
	F	4	*G*	*T*			*G/T*	*G*								**G/C**	**A/C**
	H	3	*G*	*G/T*			*G*	*G/T*								**G**	**G**
	H	7	*A/G*	*G/T*			*G*	*G/T*								**G**	**G**
	H	3	*G*	*T*			*G/T*	*G*								**G**	**G**
	H	7	*A/G*	*T*			*G/T*	*G*								**G**	**G**
OR-MRD93 × NM-LNF23	F	5										**A/G**	**G/T**	**C/T**	**G**		
	F	1										**A/G**	**G/T**	**C/T**	**A/G**		
	H	4										**A/G**	**G/T**	**C/T**	**A/G**		
	H	9										**G**	**T**	**C**	**A/G**		
	H	5										**G**	**T**	**C**	**G**		
OR-MRD61 × NM-LNF23	H	1										*A/G*	*T*	*C*			
	H	9										*G*	*T*	*C*			
	H	7										*A*	*G/T*	*C/T*			
	H	6										*A/G*	*G/T*	*C/T*			
NM-LNF2 × OR-MRD93	F	2		*G/T*	**T**		*G/T*			*T*					**G**	*G*	
	F	10		*G/T*	**T**		*G/T*			*T*					**G**	*A/G*	
	F	9		*T*	**T**		*G*			*A/T*					**G**	*G*	
	H	3		*G/T*	**G/T**		*G/T*			*T*					**A/G**	*A/G*	
	H	1		*G/T*	**G/T**		*G/T*			*T*					**A/G**	*G*	
	H	3		*T*	**G/T**		*G*			*A/T*					**A/G**	*G*	

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

716 Plant identity, phenotypic sex (F = female; H = hermaphrodite) and putative genotype at loci mapped to linkage group 4 and 6 (LG4

717 and LG6), or unmapped but inferred from progeny segregation ratios (Locus *x*). At LG4 the *MS* allele codes for male sterility and is

718 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

719 thus, codes for male sterility). At the unmapped locus *x*, the *T* codes male fertility and is dominant to *t* which codes for male sterility.

			Inferred genotypes		
Population	**Plant ID**	**Sex**	**Locus LG4**	**Locus LG6**	**Locus *x***
OR-MRD	OR-MRD93	H	*mfmf*	*Rr*	*TT*
	OR-MRD61	H	*mfmf*	*RR*	*TT*
	OR-MRD45	H	*mfmf*	*RR*	*Tt*
	OR-MRD30	F	*MSmf*	*RR*	*TT*
	OR-MRD27	F	*MSmf*	*RR*	*TT*
	OR-MRD90	F	*MSmf*	*RR*	*TT*
NM-LNF	NM-LNF23	H	*mfmf*	*Rr*	*TT*
	NM-LNF25	H	*mfmf*	*Rr*	*TT*
	NM-LNF14	H	*mfmf*	*Rr*	*Tt*
	NM-LNF2	F	*mfmf*	*rr*	*TT*
	NM-LNF4	F	*mfmf*	*rr*	*TT*
	NM-LNF26	F	*mfmf*	*rr*	*TT*

720

721

722

723

724 **Figure legends**

725 Figure 1. Map of male sterility in genome of hermaphrodite *F. vesca* subsp. *bracteata* from New

726 Mexico (NM-LNF23).The seven chromosomes based on the Fvb reference genome (Tennessen

727 et al. 2014) are denoted by orange bars along the *x*-axis and LOD scores associated with male

728 function (blue line; left hand *y* –axis) and PPR gene density (PPR/Mb, red line; right hand *y* –

729 axis) on the *y* –axes. A significant LOD score (>3) only occurs on LG 6, peaking at 8.8 for

730 markers between 35.0-36.0 Mb. This region overlaps one of the two densest clusters of PPR

731 genes in the genome, the other occurring on LG5.

732 Figure 2. Male sterility genomic regions. (A). The LG6 locus (*R/r*). The male sterility region is

733 defined as the span including 10 targeted sequence markers that perfectly match male sterility

734 (see text). Three markers that were genotyped to confirm location and explore segregation with

735 male sterility in other crosses with Sanger sequencing are noted by blue boxes (Fvb6_34763k,

736 Fvb6_35142k and Fvb6_36607k). (B). The LG4 (*MS/mf*) locus. The male sterility region was

737 mapped in Tennessen et al. (2013). Locations of Sanger markers used in this study are indicated

738 by a blue box (Fvb4_30092k).

739 Figure 3. Conceptual model of male sterility loci in *F. vesca* subsp. *bracteata*. Black lines

740 represent potential inhibitory interactions between genes. Dashed lines are uncharacterized

741 genetically or posed as alternatives to the solid lines (see text). Three nuclear loci are represented

742 by their known (LG4 and LG6) or unknown (LG*x*) location in the genome. The hypothesized

743 involvement of these loci with a cytoplasmic male sterility (CMS) locus that blocks pollen

744 development. The LG6 locus (*R/r*) is a restorer locus at which one copy of a dominant allele is

745 sufficient to block CMS and restore fertility; thus male sterility is recessive at this locus. The

746   LG4 locus (*MS/mf*) is an inhibitor at which a single dominant allele is sufficient to block the LG6

747   restorer. Thus the *MS* allele is both dominant to *mf* and epistatically dominant to *R*.

748   Homozygosity for the *t* allele at the third unmapped locus on LG*x* also leads to disruption of

749   pollen production, or could act as a restorer (*T* allele) similar to the LG6 locus.
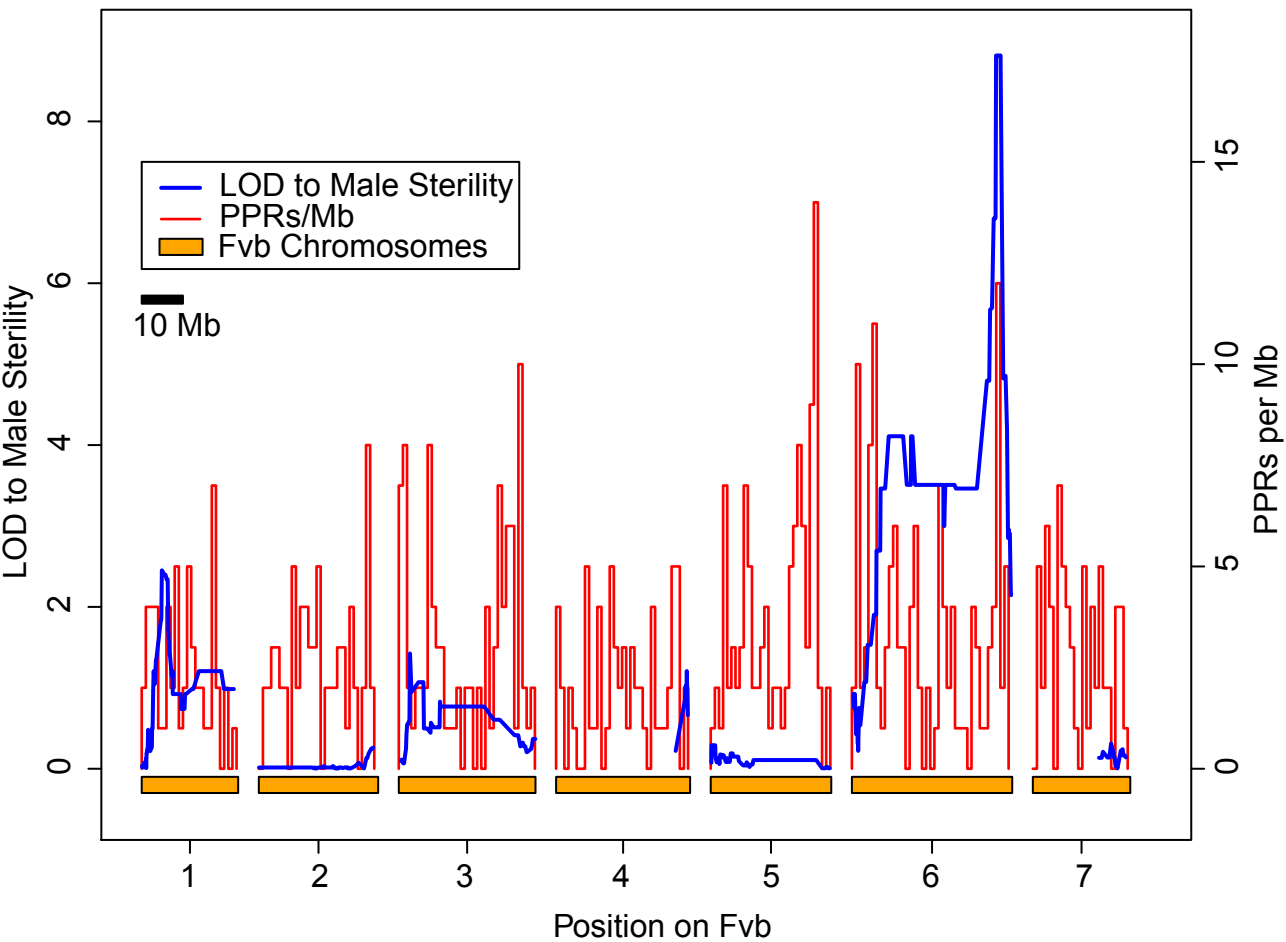
750

751
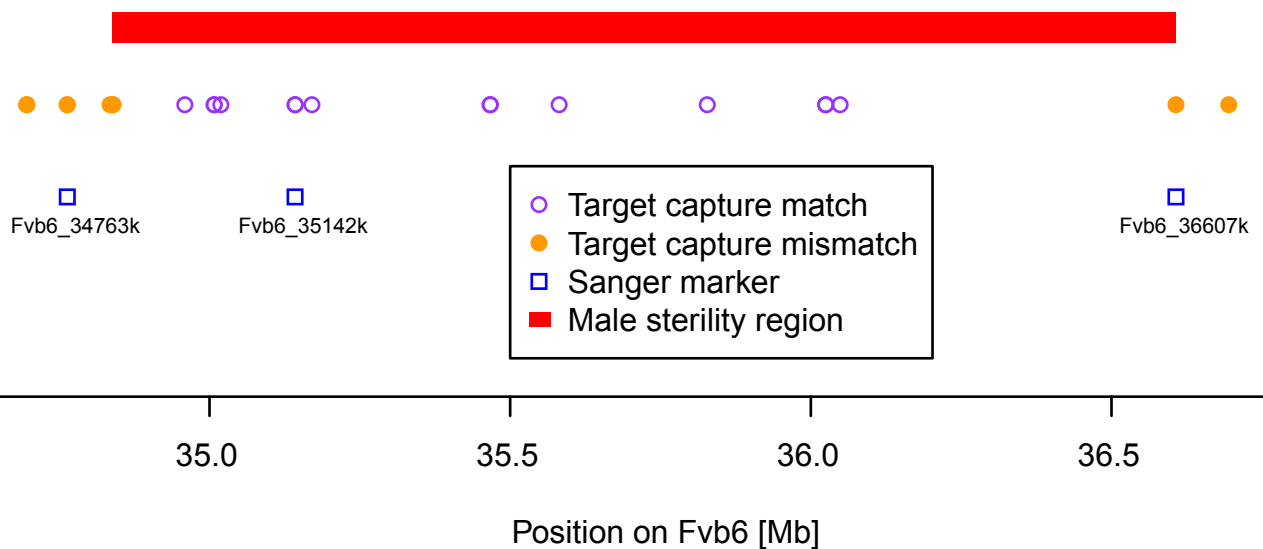
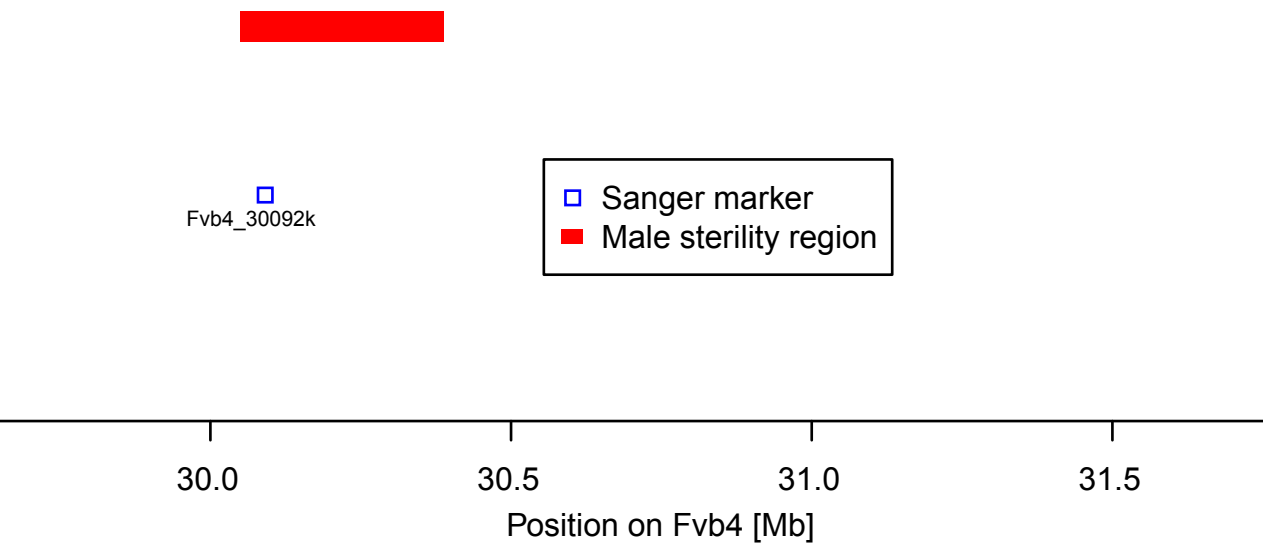752

753

754   **Supplementary Files**

755   Table S1. Cytotypes of *F. vesca* subsp. *bracteata* parents used in the crossing study. Population

756   and plant ids and sex types are given along with sequence for chloroplast and mitochondrial

757   genes. Cytotype codes follow Stanley et al. (2015).

758   Table S2   Primer sequences (forward/reverse), annealing temperatures and reference genome

759   coordinates for eight informative polymorphic sites segregating with male sterility identified

760   OR-MRD30xOR-MRD60 and NM-LNF23 self map cross population. Coordinates in FvH4 are

761   with respect to the *F. vesca* ssp. *vesca* reference genome version 1. 0 (Shulev et al. 2010), while

762   those in bold are on the scaffold scf0513158b (Tennessen et al. 2013). Coordinates in Fvb are

763   with respect to *F. vesca* subsp. *bracteata* assembly of the reference genome version 1.0 ('Fvb'

764   described in Tennessen et al. 2014).

765   Table S3. Genomic locations, functional annotations, and PLAZA 3.0 gene families of genes at

766   the Fvb6 male sterility locus in *Fragaria vesca* subsp. *bracteata*. . Notes: 1: Pentatricopeptide

767   repeats (PPR), 2: the PPR gene family that contains known and suspected fertility restorers, 3: F-

768   box proteins observed to be upregulated in meiotic anthers at stage 9, 4: miRNA target (Xia et al

769   2015), 5: unique to Darwish et al. (2015) annotation.

770

771   Table S4. Genomic locations and PLAZA 3.0 gene families of PPR genes at the peach RF1 and

772   RF2 loci (Donoso et al. 2015). The PPR gene family that contains known and suspected fertility

773   restorers denoted with * in the notes column.

774

A

Fvb6_34763k	Fvb6_35142k				Fvb6_36607k

- ○ Target capture match
- ● Target capture mismatch
- ☐ Sanger marker
- ▬ Male sterility region

Position on Fvb6 [Mb]

B

Fvb4_30092k

- ☐ Sanger marker
- ▬ Male sterility region

Position on Fvb4 [Mb]

LGx: LG4:   LG6: