THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Recombination changes at the boundaries of fully and partially sex-linked regions between closely related Silene species pairs

OPEN ACCESS

1   Recombination changes at the boundaries of fully and partially sex-linked regions

2   between closely related *Silene* species pairs.

3

4

5   **Authors**

6   JL Campos, S Qiu[1], S Guirao-Rico[2], R Bergero, D Charlesworth

7   Institute of Evolutionary Biology, School of Biological Sciences, University of

8   Edinburgh, Edinburgh, UK

9

10   [1] Current address: Department of Animal and Plant Sciences, University of Sheffield,

11   Sheffield, UK

12   [2] Current address: Department of Animal Genetics, Center for Research in

13   Agricultural Genomics (CRAG) CSIC-IRTA-UAB-UB, Edifici CRAG,

14   Campus Universitat Autònoma Barcelona, Bellaterra, Spain

15   Corresponding author: D Charlesworth

16   Institute of Evolutionary Biology, School of Biological Sciences, University of

17   Edinburgh, Edinburgh, UK

18   Email: Deborah.Charlesworth@ed.ac.uk

19   Fax: ++44 131 650 5751

20

21   **Keywords:** Sex chromosomes, recombination suppression, pseudoautosomal region,

22   partial sex-linkage

23

24   **Running title:** Evolution of sex-linkage in *Silene*

25

26   **Data Archival Location**: Dryad

27   **Word count for main text:** 6874, including all main text, acknowledgements etc, but

28   excluding references

29 **Abstract**

30 The establishment of a region of suppressed recombination is a critical change

31 during sex chromosome evolution, leading to such properties as Y (and W)

32 chromosome genetic degeneration, accumulation of repetitive sequences, and

33 heteromorphism. Although chromosome inversions can cause large regions to have

34 suppressed recombination, and inversions are sometimes involved in sex

35 chromosome evolution, gradual expansion of the non-recombining region could

36 potentially sometimes occur. We here test whether closer linkage has recently

37 evolved between the sex-determining region and several genes that are partially sex-

38 linked in *Silene latifolia*, using *S. dioica*, a closely related dioecious plants whose XY

39 sex chromosome system is inherited from a common ancestor. The *S. latifolia*

40 pseudoautosomal region (PAR) includes several genes extremely closely linked to the

41 fully Y-linked region. These genes were added to an ancestral PAR of the sex

42 chromosome pair in two distinct events probably involving translocations of

43 autosomal genome regions causing multiple genes to become partially sex-linked.

44 Close linkage with the PAR boundary must have evolved since these additions,

45 because some genes added in both events now show almost completely sex-linkage

46 in *S. latifolia*. We compared diversity patterns of five such *S. latifolia* PAR boundary

47 genes with their orthologues in *S. dioica*, including all three regions of the PAR (one

48 gene that was in the ancestral PAR, and two from each of the added regions). The

49 results suggest recent recombination suppression in *S. latifolia*, since its split from *S.*

50 *dioica*.

51

52

53 **INTRODUCTION**

54   Sex chromosome evolution involves the establishment of a region of

55 suppressed recombination between a chromosome pair that carries a sex-

56 determining locus (reviewed by Bull 1983). If a non-recombining region persists, it

57 may lead to differentiation of the Y from the homologous X chromosome (or

58 between Z and W chromosomes in ZW systems). Regions where recombination still

59 occurs between the X and the Y are called pseudo-autosomal regions (abbreviated to

60 PAR). In several species, the region of suppressed recombination has expanded in

61 successive events, each time shrinking the PAR and forming a new fully sex-linked

62 region whose alleles in the non-recombining Y or W chromosome then start

63 diverging from those of the homologous X or Z chromosome. Consequently, genes in

64 X chromosome regions physically closest to the PAR have lower sequence divergence

65 than ones distant from the PAR. These regions of different Y-X divergence are

66 termed "evolutionary strata", and were first noticed in human sex chromosome

67 sequences, and later found in all mammals (Cortez et al. 2014), and in the ZW

68 chromosomes of birds, although some paleognathous birds have retained extensive

69 PARs (reviewed in Zhou et al. 2014). Two plants that have separate sexes (dioecious

70 plants) and sex chromosomes with XY males and XX females, *S. latifolia* and *Carica*

71 *papaya*, also have strata (Bergero et al. 2007; Wang et al. 2012). In papaya, two

72 distinct strata appear to have evolved by fixation of inversions in the Y-linked region

73 (Wang et al. 2012), but in *S. latifolia* Y-X divergence may increase without

74 discontinuity with distance from the PAR (Chibalina and Filatov 2011).

75   The reasons for the repeated formation of new evolutionary strata, or for the

76 possible gradual extension of sex chromosomes' non-recombining regions, are not

77 yet fully understood, and the mechanisms involved are not completely known in

78 most species, with new mysteries emerging about the PAR boundary even in well-

79 studied species like humans (Cotter et al. 2016). One plausible reason/cause for the

80 situation driving changes in the PAR boundary (and creation of new strata) involves

81 sexually antagonistic mutations (advantageous in one sex but disadvantageous in the

82 other, often abbreviated to SA). SA mutations in genes closely linked to the fully sex-

83 linked region may establish polymorphisms, leading to selection for closer linkage.

84 For example, closer linkage between male-benefit SA alleles and the Y-linked region

85   restricts these alleles to males, and avoids harming the females (Bull 1983; Rice

86   1987; Jordan and Charlesworth 2012).

87   Recently evolved sex chromosomes, such as those in some plants, offer good

88   systems for studying how partially sex-linked regions evolve full sex linkage. In the

89   plant *S. latifolia*, recombination suppression between the fully Y- and X-linked

90   regions was initiated around 5-10 MY ago, but at least one region evolved

91   suppressed recombination subsequently, forming a younger fully Y-linked male-

92   specific, or MSY, region (Bergero et al. 2007; Chibalina and Filatov 2011). By

93   comparing the genetic map of the *S. latifolia* X chromosome with mapping results in

94   the related species *S. vulgaris*, which does not have sex chromosomes, the *S. latifolia*

95   PAR was inferred to have been formed by independent additions of two genomic

96   regions to an ancestral PAR, through translocations from other chromosomes, as

97   shown in Figure 1 (Bergero et al. 2013; Qiu et al. 2016). Several genes located near

98   the boundary with the fully sex-linked region recombine so rarely that some variants

99   are found only in males (Qiu et al. 2016); among these PAR boundary genes, some

100   were added in one addition event, and some in the other (Figure 1).

101

102   Figure 1 here

103

104   Translocations that add new regions onto a PAR, as in *S. latifolia* and its close

105   relatives, are particularly interesting in relation to the evolution of suppressed

106   recombination, as crossing over in the ancestral PAR should not be affected by an

107   addition, and should continue after the addition event, generating recombinants

108   between the MSY and the added regions. Therefore, if such recombination does not

109   occur, it was probably suppressed after the rearrangement occurred. Such an

110   addition occurred in the ancestor of Eutherian mammals, and recombination was

111   indeed subsequently suppressed, moving the PAR-MSY boundary far from its pre-

112   addition location (reviewed in Cortez et al. 2014). However, few such cases have

113   been studied. The genes near the *S. latifolia* PAR boundary represent a much more

114   recent translocation situation than that in mammals, and are ideal for testing

115   whether the PAR-MSY boundary has remained in the same location. Close linkage

116   observed between the *S. latifolia* MSY and genes from both addition events (Qiu et

117 al. 2016) suggests recombination suppression after the translocations occurred

118 (other genes added in both translocations, have remained loosely linked to the MSY

119 boundary). However, although the translocations should not directly suppress

120 recombination in the PAR boundary region, it is important to test this alternative.

121   We therefore examined species closely related to *S. latifolia*, mainly, but not

122 exclusively, the closest relative, *S. dioica*. The *S. dioica* XY chromosomes are

123 homologous with those of *S. latifolia* (Nicolas et al. 2005), and have indistinguishable

124 morphology and arm ratios (Grabowska-Joachimiak and Joachimiak 2002). The two

125 species hybridise readily, producing fertile progeny, consistent with non-rearranged

126 chromosomes, and there is evidence for ongoing gene flow (Muir et al. 2012; Hu and

127 Filatov 2015). If the translocation events outlined above occurred in a common

128 ancestor of these species and directly suppressed recombination between the added

129 regions and the MSY region, the PAR boundary genes should also show close linkage

130 to the MSY in *S. dioica*. If, however, closer linkage with the Y-linked region has

131 evolved in *S. latifolia*, linkage should be looser, or absent, in *S. dioica*.

132   Three other closely related *Silene* species, *S. diclinis, S. marizii* and *S. heufellii*,

133 are also dioecious. Reliable inference of the order in which they split will require

134 large numbers of gene sequences, which are not currently available other than for *S.*

135 *latifolia* (Rautenberg et al. 2010). However, *S. latifolia* forms hybrids more readily

136 with *S. dioica* than *S. diclinis* (Prentice 1978), suggesting that *S. diclinis* is an outgroup

137 to the two other species. Genes that are fully sex-linked in *S. latifolia* are also fully

138 sex-linked in *S. dioica* and *S. diclinis* (Laporte et al. 2005; Nicolas et al. 2005; Kaiser et

139 al. 2009; Muir et al. 2012), supporting the view that the sex chromosomes of these

140 species are homologous. *S. marizii* and *S. heufellii* have been less studied, but,

141 consistent with the genetic evidence just outlined, the sex chromosomes of *S. marizii*

142 are similar in morphology to those of *S. latifolia* and *S. dioica*, whereas *S. diclinis* has

143 undergone a Y-autosome reciprocal translocation (Howell et al. 2009). As explained

144 below, we demonstrated that one gene, *SlCyp* in the younger *S. latifolia* sex

145 chromosome stratum, with an estimated K_S value between the X- and Y-linked

146 sequences of 0.067 (Bergero et al. 2007), also has variants found only in males of *S.*

147 *dioica, S. diclinis* and *S. marizii*, indicating that this recently evolved stratum is sex-

148 linked in all of them. However, *S. latifolia* PAR genes have not yet been mapped in

149 the outgroup species.

150   We here study outgroups to ask (i) whether the *S. latifolia* PAR boundary genes

151 are as well closely linked to the MSY in the outgroup species, and to ask the related

152 question (ii) whether results suggest that the translocations that caused these genes

153 to become partially sex-linked in *S. latifolia* directly caused restricted recombination,

154 versus recombination being suppressed subsequently. Rather than genetically

155 mapping these genes, we used a highly sensitive population genetic approach,

156 testing for associations between SNPs in the PAR boundary loci and the male-

157 determining region, using the subdivision measure K_{ST} between males and females.

158 K_{ST} is the average F_{ST} per site, computed from sequence data (Hudson et al. 1992),

159 and reflects linkage disequilibrium due to population subdivision (Charlesworth et al.

160 1997), such as Y-X differentiation resulting from absent or very rare recombination.

161 F_{ST} has already been proposed as the best way to test for associations between

162 alleles of partially sex-linked genes and a fully sex-linked locus (Qiu et al. 2013;

163 Kirkpatrick and Guerrero 2014; Qiu et al. 2016). Importantly, this can potentially

164 detect recombination even if it is too rare to be detectable in families, as is the case

165 in *S. latifolia* for the PAR boundary loci studied here (Qiu et al. 2016). Our analysis

166 suggests recently decreased recombination in *S. latifolia*. Did the translocations

167 directly cause recombination suppression? Based on evidence of partial sex linkage

168 in *S. dioica* for the two genes we studied in the first translocation, subsequent

169 recombination suppression is implied between the *S. latifolia* MSY and the added

170 region. The second translocation could, however, have occurred in *S. latifolia* during

171 the very short evolutionary time separating this species from *S. dioica*, so it remains

172 possible that this event directly suppressed recombination (though we present

173 arguments that this is unlikely).

174

175 **METHODS**

176 Genes and plant samples

177 In *S. latifolia*, no recombinants were detected between the genes studied here and

178 the MSY region in male meiosis, and, in natural populations of *S. latifolia*, they all

179 show marked sex differences in allele frequencies, including some variants found

180   only in males, indicating very close linkage to the MSY (Qiu et al. 2016). Gene *cs3597*

181   is part of a putative ancestral PAR, as, in *S. vulgaris*, it maps to the same linkage

182   group (SvLG12) as the genes that are fully sex-linked in *S. latifolia* (Qiu et al. 2016).

183   The four other *S. latifolia* PAR-MSY boundary region genes studied here map to two

184   other linkage *S. vulgaris* groups. Specifically, the sequences of genes *E559* and *E521*

185   were added along with 4 other genes that map to linkage group SvLGSmall and map

186   far from the MSY in *S. latifolia*, while, of the 7 *S. latifolia* PAR genes that map to

187   SvLG9, only *cs935* and *E523* are in the boundary region (Qiu et al. 2016). We also

188   sequenced a gene, *cs4991*, located slightly more distal to these PAR boundary genes.

189   No recombinants were seen in males in the family in which *cs4991* was mapped;

190   however, it maps very close to gene *E352* in the X chromosome, which yielded many

191   recombinants (6/58) in male meiosis of another *S. latifolia* family, so the

192   recombination frequency in the *S. latifolia* population as a whole could be several

193   percent (Qiu et al. 2016). The orthologue of *cs4991* has not been mapped in *S.*

194   *vulgaris*, but its map location in *S. latifolia* suggests that it was added to the PAR

195   along with other SvLG9 genes, after the SvLGSmall genes were added. All these

196   genes segregate as single-copy loci in at least one *S. latifolia* full-sib family, with two

197   alleles in both sexes (Bergero et al. 2013; Qiu et al. 2016).

198   We sequenced orthologues of the *S. latifolia* PAR genes in *S. dioica* males and

199   females, using plants grown from seeds collected from 11 different locations

200   distributed across Europe (see Supplementary Table S1 and Figure S1, which also

201   shows the localities from which the *S. latifolia* samples were collected; these are

202   described in detail in Qiu, Bergero and Charlesworth, 2013). The samples from each

203   species were collected from many different populations for two reasons. First, this

204   ensures that the results are representative of the species generally. Second, our goal

205   was to test for associations between sequence variants and the sex-determining

206   region. Such a "scattered sample", with few individuals per population, minimizes

207   false inferences caused by random associations within populations (Städler et al.

208   2009).

209   Plants were grown in Edinburgh and sexed once they flowered. The sexes were

210   also confirmed by PCR amplification of intron 2 of *SlCyp*, a gene in the younger of the

211   two evolutionary strata in this species (Bergero et al. 2007). The sexes assigned by

212   this marker agreed perfectly with those observed at flowering, indicating that none

213   of our males carries a recombinant genotype for the sex chromosomes.

214   In total, we obtained sequences for most of the genes from 12 alleles from *S.*

215   *dioica* females, and 20 from males, so as to represent alleles associated with the Y as

216   well as the X chromosome. The *S. latifolia* samples with which we compared these

217   new results included 38-44 alleles sequenced from females and 40-42 from males

218   (Qiu et al. 2016). Smaller samples of *S. marizii* and *S. diclinis* (see Supplementary

219   Table S1) were also studied for two genes, *E559* and *E523*, one from each addition

220   event.

221

222   **DNA extraction, PCR reactions, and cloning**

223   Genomic DNA for sequencing was extracted from leaves using the DNeasy

224   Plant Mini Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions.

225   PCR amplifications were then performed with primers given in Qiu et al. (2016),

226   using Phire Hot-Start DNA Polymerase (Thermo Fisher Scientific, Paisley, UK) in a

227   Finnzymes' Piko cycler, with the following conditions: 1 cycle of initial denaturation

228   at 98° for 30 sec, 10 cycles of DNA denaturation at 98° for 5 sec, primer annealing

229   varying from 60 to 70° for 5 sec, and DNA amplification at 72° for 30 sec, 25 cycles at

230   98° for 5 sec, 60° for 5 sec, 72° for 60 sec, and finally 1 cycle at 72° for 5 min. The PCR

231   products were cleaned with ExoSAP-IT (Affymetrix, High Wycombe, UK) and Sanger

232   sequenced on an ABI 3730 capillary sequencer (Applied Biosystems, Warrington,

233   UK). Single nucleotide polymorphisms (SNPs) were examined after direct sequencing.

234   DNAs producing PCR amplicons with heterozygous indels (generally in intron regions)

235   were re-amplified with the Phusion proof-reading DNA polymerase (Thermo Fisher

236   Scientific, Painsley, UK) using the same PCR conditions as above, and cloned into the

237   vector pJET 1.2/blunt (Thermo Fisher Scientific, Paisley, UK) before sequencing, six to

238   eight colonies were screened per amplicon. The resulting sequences were aligned in

239   Sequencher 4.8 (Gene Codes, Ann Arbor, MI; http://www.genecodes.com), including

240   sequences of the orthologous genes from *S. latifolia* and *S. vulgaris*, and manually

241   adjusted using Se-al v. 2.0 (http://tree.bio.ed.ac.uk/software/seal/). Alignments

242   including sequences from all species studied here have been deposited in Dryad

243   under accession numbers [**TO BE ADDED**].

244

245 **Sequence analyses**

246   We compared variant frequencies of alleles, and frequencies of

247 heterozygotes in males and females, using Chi-square tests. The polymorphism

248 analyses, including estimates of nucleotide diversity and Tajima's D (Tajima 1989),

249 and divergence estimates between the *S. latifolia* and *S. dioica* sequences were done

250 using DnaSP v.5.00.06 (Librado and Rozas 2009). Estimates of K_{ST} and tests for

251 subdivision under the null hypothesis of no subdivision were done using 1000

252 permutations of $K*_{ST}$ and another subdivision measure, *Snn* (Hudson et al. 1992)

253 implemented in DnaSP. Significance of K_{ST} was tested using the $K*_{ST}$ statistic, rather

254 than with K_{ST}, because $K*_{ST}$ has good power for small samples (Hudson et al. 1992).

255 To obtain error bars for the K_{ST} values shown in Table 1 and Supplementary Figure

256 S4, we resampled individuals 40 times from each species with replacement,

257 maintaining the same sample sizes of males and females for each gene, and re-

258 estimated K_{ST} using these new alignments for each gene and species.

259   Our samples from *S. dioica* are smaller than from *S. latifolia*, which will

260 reduce the power of our tests of significance with *S. dioica* samples. To compare

261 associations between variants and the fully sex-linked region, we therefore sub-

262 sampled from *S. latifolia*, as follows. For each gene, we constructed samples with

263 the same numbers of males and females as in the *S. dioica* sample, by randomly

264 sampling individuals separately from the two sexes to form a reduced sized *S.

265 latifolia* sequence sample. We estimated the means of Tajima's D from 1 000 such

266 sub-samples, using the batch mode of DnaSP. However, as the batch mode does not

267 estimate K_{ST} or perform tests of significance of subdivision, we used 40 sub-sampled

268 alignments to assess the effect of sample size on this measure of subdivision, and its

269 significance, in the manner just described.

270

271 RESULTS

272 **Subdivision between *S. latifolia* and *S. dioica***

273   Because there is ongoing gene flow between *S. latifolia* and *S. dioica* (Minder

274 et al. 2007; Karrenberg and Favre 2008)*,* this might affect different genome regions

275 differently (Muir et al. 2012). Before comparing associations between the partially

276   sex-linked genes and the MSY in the two species, we therefore tested whether the

277   PAR gene sequences we obtained might show less gene flow than autosomal genes.

278   Based on all site types, divergence is invariably low, and the PAR genes studied here

279   have mean net divergence, D_a, of 0.25%, and include no fixed differences between

280   the species, while shared variants are common, similar to the published results for

281   autosomal and X-linked genes (Supplementary Table S2), and consistent with

282   incomplete isolation between the two species. Net divergence, D_a, ranges from only

283   0.55% for 20 autosomal genes (Muir et al. 2012), or 0.27% in a larger sample of

284   genes recently studied in three populations of each species (Hu and Filatov 2015), to

285   0.51% for X-linked genes (Hu and Filatov 2015). K_{ST} values for the 6 PAR genes

286   studied here nevertheless indicate differentiation between *S. latifolia* and *S. dioica*

287   (Supplementary Table S2).

288   Gene flow between *S. latifolia* and *S. dioica* does not appear to be highly

289   restricted for the PAR genes studied here. One PAR gene, *E521*, had high diversity in

290   *S. latifolia*, because many sites are heterozygous in all males, reflecting diverged

291   sequences associated with the X and Y chromosome fully sex linked haplotypes (Qiu

292   et al. 2016), but the X-linked alleles showed no variants. In *S. dioica*, we detected no

293   variation in *E521*, despite testing several pairs of PCR primers to exclude the

294   possibility that a highly diverged allele is present and is not amplified. Excluding

295   *E521*, whose low diversity in *S. dioica* will increase K_{ST}, the mean K_{ST} value for the five

296   other PAR genes studied here is 0.084 (with 95% confidence intervals 0.034-0.128),

297   versus the estimated values of 0.154 for autosomal genes, or 0.381 for X-linked

298   genes (Hu and Filatov 2015).

299

300   **Subdivision of PAR boundary genes between males and females in *S. latifolia***

301   **versus *S. dioica***

302   Our results suggest that, consistent with the cytogenetic evidence outlined

303   above, the genes in the added regions closely linked to the *S. latifolia* PAR-MSY

304   boundary are also partially sex-linked in *S. dioica*. However, they appear to

305   recombine more often in *S. dioica*. A first indication of sex linkage is higher

306   nucleotide diversity in males than in females, indicating that some variants in the

307 sequences are male-specific. The difference is significant in *S. latifolia* (Mann-

308 Whitney U-test, P = 0.03), and many variants are present only in males, or with much

309 higher frequencies in males than females (Supplementary Figures S2 and S3). A

310 significantly higher frequency of sites that are heterozygous in males than in females

311 (Mann-Whitney U-test, P = 0.012), with many sites heterozygous in most males, but

312 few in females (Supplementary Figures S2-S4), shows that the sex difference in

313 diversity is not due to our sample having included some males from a population

314 with divergent sequences, consistent with recombinant alleles not being

315 geographically restricted (Qiu et al. 2016). In *S. dioica*, however, diversity is only

316 slightly higher in males (Figure 2), though male-specific variants in gene *E559*

317 (Supplementary Figure S2), and high heterozygote frequencies in males in several of

318 the genes studied (Figure 3A, Supplementary Figure S5) suggest partial sex linkage of

319 at least some of these genes in this species also. However, many variants are not

320 strongly associated with the fully sex-linked region (Supplementary Figures S2 and

321 S3); again recombinant alleles are not confined to any geographic region or sub-set

322 of populations.

323

324   Figure 2 and Figure 3 about here

325

326   Tajima's *D* values more clearly suggest sex linkage of some of these genes in

327 both species (Figure 3B). In *S. latifolia*, *D* values are positive for most PAR genes in

328 males, whereas other genome regions consistently show negative values (Qiu et al.

329 2016). For the genes studied here, males and females differ significantly (D_{males}=

330 1.28, $D_{females}$ = -1.29, Mann-Whitney U-test, *P* = 0.02), suggesting that these genes

331 are closely linked with the fully sex-linked region, so that variants are often

332 heterozygous in males. Positive Tajima's *D* values in males, and negative values in

333 females are also seen n *S. dioica*, particularly for genes *cs3597* and *cs935*, where the

334 value is statistically significant in males (Figure 3B and Table S3).

335   As explained above, analysis of subdivision between the sexes is the best way

336 to test for associations of variants with the fully sex-linked region (Kirkpatrick and

337 Guerrero 2014). For all five PAR boundary genes that could be tested in *S. latifolia,*

338 K_{ST} between sequences from males and females was high, and differed highly

339 significantly from zero (Table 1), unlike the more distally located *cs4991* gene

340 (P=0.052), or other PAR genes more loosely linked to the fully sex-linked region, or

341 autosomal loci (Qiu et al. 2010). Consistent with partial sex linkage, the significant

342 associations in *S. latifolia* are somewhat weaker than the value of one third expected

343 for fully sex-linked genes assuming a 1:1 sex ratio and fixed differences between Y-

344 and X linked alleles. In *S. dioica*, the K_{ST} values were smaller, but significant

345 subdivision between the sexes was detected for several PAR boundary genes (Table

346 1). Because variances are not available for K_{ST} values, we tested for subdivision by

347 two significance tests, and by comparing Tajima's D values between the sexes,

348 including using sub-samples of sequences from each species; for three genes, at least

349 one test suggests partial sex linkage in *S. dioica* (Supplementary Figure S4).

350

351 Table 1 about here

352

353 The lower significance levels for our tests of subdivision between the sexes in

354 *S. dioica* are not wholly due to our having sequenced fewer alleles than from *S.*

355 *latifolia*, because smaller sub-samples of sequences from the *S. latifolia* alleles, re-

356 analyzed in the same manner as for the complete data set (see Methods)

357 consistently yielded K_{ST} values much higher than in *S. dioica*, and highly significant

358 subdivision between the sexes (Table 1). For two of the genes, *cs3597* and *E523*,

359 none of our 40 sub-samples had K_{ST} value as low as the *S. dioica* value. For the other

360 genes, only a few sub-samples from *S. latifolia* had lower K_{ST} values than those

361 observed in *S. dioica*, and the means over the 40 sub-samples are invariably higher

362 (Table 1).

363 Sub-samples also maintained the higher Tajima's *D* values in males than

364 females in *S. latifolia* (Supplementary Table S3). Although the difference between

365 the sexes is smaller in the sub-samples than in the full data set, as expected, it

366 remains statistically significant (Mann-Whitney U-test, *P* = 0.03). Overall, the results

367 from the sub-samples show that there is a real biological difference between the

368 species.

369   Although only a 120 bp region of *E521* could be sequenced in *S. latifolia*

370   females, and in *S. dioica* (Supplementary Table S2), our results also support the

371   conclusion that this gene is also closely associated with the MSY in *S. latifolia*

372   (Figures 1 to 3 and Supplementary Table S3), but not In *S. dioica*, in which there is a

373   complete absence of variants, and thus no difference between the sexes.

374   As explained above, the *cs3597* gene is probably part of an ancestral PAR (Qiu

375   et al. 2016), and appears partially sex-linked in both species based on Tajima's D

376   values (Figure 3B) and K_{ST} analysis. Interestingly, even though this gene was probably

377   partially sex-linked in the common ancestor, it has a considerably smaller allele

378   frequency difference and K_{ST} between the sexes in *S. dioica* than *S. latifolia* (see

379   Table 1), and a smaller difference in the frequency of heterozygotes (Figure 3A). It

380   has therefore probably become more closely linked to the MSY in *S. latifolia*.

381

382   **Evidence that closer linkage has evolved in *S. latifolia* since splitting from *S. dioica***

383   To infer more rigorously whether the recombination state is changed in *S.*

384   *latifolia*, or in *S. dioica*, an outgroup is needed. We therefore sequenced two genes

385   in *S. marizii* and *S. diclinis*, one from each of the two additions that formed the PAR

386   (see above). Both these genes include many sites showing complete sex linkage in *S.*

387   *latifolia*, but neither of them included any fully Y-linked variants in *S. marizii* or *S.*

388   *diclinis* (Supplementary Figures S2 and S3). These genes are therefore not completely

389   sex linked in *S. marizii* and *S. diclinis*; they could be partially sex linked, but, because

390   we have only small samples of these two species (see Supplementary Table S1), we

391   have no firm evidence that the genes studied are not autosomal in these species.

392   Both the rearrangements that created the PAR in *S. latifolia*, and the one that is

393   firmly inferred in *S. dioica*, might thus have occurred after the split from *S. marizii*

394   and *S. diclinis*. Nevertheless, taken together, these results suggest that the *S. latifolia*

395   state of strong associations with the fully sex-linked region is the derived state for

396   PAR boundary genes derived from both addition events that formed the PAR, and

397   that the other three species share the less closely linked state.

398

399   **DISCUSSION**

400   Overall, the difference between *S. latifolia* versus *S. dioica* in associations of

401   sequence variants with the sexes (reflecting linkage disequilibrium between alleles of

402   PAR boundary region genes and the MSY region), suggests less recombination

403   between these PAR boundary genes in *S. latifolia* than in *S. dioica*. Subdivision

404   between X and Y chromosomes due to suppressed recombination between the XY

405   pair in males will be most evident when X and Y haplotypes are sequenced, and less

406   readily detected in samples of sequences from males and females (Kirkpatrick and

407   Guerrero 2014), so that some discrepancies between different tests for associations

408   between the MSY and gene sequences are unsurprising. However, for partially sex-

409   linked genes, the phase of variants is not known. Our approach of testing sequences

410   from males and females, without attempting to infer their phases, is conservative.

411   Our analysis suggests recently decreased recombination in *S. latifolia*, but this

412   could be either by suppression of recombination between the MSY and the region

413   carrying the genes we studied, or because the translocations that brought these

414   genes onto the XY pair occurred in the short evolutionary time since *S. latifolia* split

415   from *S. dioica*. We consider this less likely, as our analyses of subdivision between

416   males and females suggests partial sex-linkage in *S. dioica*, for at least some of the

417   PAR boundary genes that we tested, implying addition before the split between the

418   species. This is consistent with cytogenetic evidence that both the X and Y

419   chromosomes are very similar in these species (Grabowska-Joachimiak and

420   Joachimiak 2002), and considerably larger than other chromosomes, unlike the

421   situation in other *Silene* species (Siroky et al. 2001).

422   Two of the genes studied, *E523* and *cs935*, may not be sex-linked in *S. dioica*

423   (see Table 1), though subdivision between the two sexes is supported for *cs935* by a

424   significant result (P = 0.2%) with the *Snn* test of Hudson et al. (1992), and a large

425   positive Tajima's D value in males but not females, see Figure 3, suggesting partial

426   sex linkage; both these genes map to the SvLG9 linkage group in *S. vulgaris*, and

427   were probably added to the evolving *S. latifolia* PAR in the second of the two

428   inferred addition events diagrammed in Figure 1 (Qiu et al. 2016). This event could

429   potentially have occurred after the split between the two species, leaving these

430   genes autosomal in *S. dioica* (this could be tested by in situ hybridisation of genes

431   from this linkage group). Alternatively, this event pre-dated the split, but close

432   linkage has subsequently evolved very recent, and only in *S. latifolia*. Interestingly,

433   however, genes *E521* and *E559* are located on LGSmall of *S. vulgaris*, indicating that

434   they became part of a PAR in the first translocation event (Qiu et al. 2016), yet only

435   *E559* shows associations with the MSY in *S. dioica*.

436   Linkage disequilibrium, and associations reflecting it, depends on the effective

437   population size (N_e) as well as the recombination rate, and also on natural selection.

438   However, the difference we find between *S. latifolia* and *S. dioica* cannot be

439   explained by an N_e difference. The estimated silent site nucleotide diversity values

440   for autosomal genes differ very slightly between *S. latifolia* and *S. dioica*, based on

441   estimates from large numbers of genes (Hu and Filatov 2015). Diversity estimates for

442   silent sites are most appropriate for assessing effective population sizes, and the

443   largest difference in such diversity estimates so far published is 7% (in the estimated

444   value of π in Muir et al. 2012), which is too small to explain the observed difference

445   in associations of variants in the PAR boundary genes with the MSY between *S.*

446   *latifolia* and *S. dioica* (the alternative diversity estimate in the same paper,

447   Watterson's theta, yielded a slight difference in the opposite direction). The

448   difference in K_{ST} between males and females in the two species therefore probably

449   reflects different recombination rates between the MSY and the genes studied.

450   Because only one of the genes studied is part of the ancestral PAR, while the

451   other four became sex-linked through the translocation events, it is likely that

452   linkage has become closer in *S. latifolia* and that *S. dioica* remained more similar to

453   the ancestral state. This is supported by the lack of evidence for sex linkage in the

454   two outgroup species, *S. marizii* and *S. diclinis*. Whether the translocations are

455   shared between all four species studied here, or whether one addition or both is

456   present in *S. dioica*, closer linkage has nevertheless clearly evolved subsequently, at

457   least in *S. latifolia*.

458   Situations where regions have been added to pre-existing sex chromosomes

459   are of great interest, because one factor that has been proposed to favour such

460   rearrangements is sexual antagonism (Charlesworth and Charlesworth 1980; Pennell

461   et al. 2015), similar to the selection leading to new evolutionary strata on sex

462   chromosomes in the first instance (see Introduction). This view predicts that closer

463 linkage with the sex-determining region should subsequently evolve, unless the

464 rearrangement directly caused close linkage between the sex-determining region

465 and the SA factor. If closer linkage with the sex-determining region is generally found

466 to evolve, this would lend support to the SA polymorphism hypothesis. We therefore

467 next briefly review the evidence about whether such changes are, in fact, seen when

468 genome regions have been added to sex chromosomes of species other than *Silene*.

469

470 **Recombination in genome regions added to regions to sex chromosomes**

471 In XY species with no recombination in males, such as *Drosophila*, Y-autosome

472 and X-autosome translocations both immediately result in complete sex-linkage

473 (reviewed in Bachtrog 2006). While these have been important for studying the

474 consequences of suppressed recombination (reviewed in Bachtrog 2006), they are

475 uninformative regarding the evolution of recombination and the possible

476 involvement of sexual antagonism.

477 In species with recombination in the heterogametic sex, however,

478 chromosomes can initially continue recombining in both sexes, even after Y-

479 autosome translocations in XY species, or W-autosome translocations in ZW species

480 like birds, where females are the heterogametic sex; similarly, X-A translocations in

481 systems with X0 males form neo-Y chromosomes that can continue to recombine

482 with their autosomal counterparts (e.g. Castillo et al. 2010; Henzel et al. 2011). It is

483 therefore interesting to ask whether recombination suppression evolves

484 subsequently.

485

486 **Additions to fully sex-linked regions**

487 In *Silene diclinis* the neo-Y (Y_2) arm created by a Y-autosome reciprocal

488 translocation probably largely recombines with the former autosome (Howell et al.

489 2009), although population genetic studies to confirm this have not yet been done,

490 nor genetic mapping to study recombination patterns in detail. Similarly, some

491 threespine stickleback populations have a Y-A translocation involving the fully sex-

492 linked end of the Y chromosome (Natri et al. 2013). In this case, however, a region

493 adjacent to the rearrangement breakpoint shows large allele frequency differences

494   between the sexes, indicating suppressed recombination (reviewed in Natri et al.

495   2013). Sex linkage of this region, extending for least 3 Mb, could simply be due to the

496   rearrangement, as a neo-X or –Y chromosome that segregates from an enlarged Y or

497   X, but is not physically joined to it, may fail to pair near the breakpoints.

498   In yet other cases, however, recombination has probably subsequently

499   become suppressed. For example, the neo-Y chromosomes have become

500   heterochromatic in several groups of related species with shared Robertsonian

501   fusions. Examples include grasshoppers (Castillo et al. 2010) and deer (Cernohorska

502   et al. 2015).

503   Sex chromosome-autosome translocations have occurred in several dioecious

504   plants (reviewed in Ming et al. 2011), and these may be excellent for studying

505   whether recombination suppression has evolved, because the pre-rearrangement

506   state is often known in outgroup species, as in *S. diclinis*. In *Rumex acetosa*, the Y_2 is

507   heterochromatic, suggesting suppressed recombination (however, the involvement

508   of an X-autosome translocation is not currently certain in this plant, see Rejón et al.

509   1994). Several other plants with XX/XY_1Y_2 systems probably have new regions added

510   to fully X-linked regions through X-autosome translocations, so that recombination

511   could be suppressed in males in parts of the added regions as direct effects of the

512   rearrangements. In the hop species *Humulus lupulus* var. *cordifolius* and *H.*

513   *japonicus*, the Y_2 (neo-Y, former autosome) chromosome is still non-heterochromatic

514   (Grabowska-Joachimiak et al. 2011). The situation is similar in the genus *Baccharis*

515   (Hunziker et al. 2002). In *Viscum fischeri*, a chain of five Y chromosomes (Barlow and

516   Wiens 1976) suggests multiple events over considerable evolutionary time, so that

517   multiple strata, corresponding to different translocation events might exist. X-Y

518   divergence data could indicate whether the neo-Y chromosomes of these plants still

519   recombine.

520

521   **Additions to partially sex-linked regions**

522   Few cases of additions onto PARs have been studied, although possible cases

523   have been inferred in birds. In two warblers, markers from a previously autosomal

524   micro-chromosome (chicken 4a) have alleles on both the Z and W chromosomes

525   (Pala et al. 2012), and larks may be similar (Brooke et al. 2010). However, it has not

526   yet been excluded that these additions occurred onto the fully sex-linked region. This

527   could be tested by genetic maps of these species and outgroup species lacking the

528   rearrangement, and tests for complete sex-linkage of the added genes, and other

529   genes from the chicken linkage groups involved in these events, would be

530   informative.

531   Cases with strong evidence for translocations that added new regions onto a

532   PAR, as in *S. latifolia* and its close relatives, are thus particularly interesting. Other

533   than *S. latifolia*, an addition to the PAR has so far been suggested in only one other

534   plant, *Rumex hastatulus* (Grabowska-Joachimiak et al. 2015). This species is in the

535   same dioecious clade as *R. acetosa* (Navajas-Pérez et al. 2005), but the

536   rearrangements may be independent, as the *R. hastatulus* one is found only in one

537   race. This North Carolina race has an XX/XY$_1$Y$_2$ sex chromosome system, so an X-

538   autosome fusion (Smith 1964) is more likely than an addition to the PAR. The neo-Y

539   has remained euchromatic (Grabowska-Joachimiak et al. 2015), consistent with a

540   smaller divergence of its sequences from those of the neo-X, and less signs of genetic

541   degeneration, compared with the ancestral chromosome pair found in the closely

542   related XX/XY Texas race (Hough et al. 2014). Moreover, while independently

543   sequenced transcriptomes from six populations of each race found male-specific

544   variants in ~80% of the genes ascertained as sex-linked in the XY populations, this

545   was found for only 28% of the neo-Y genes, suggesting that most of the Y$_2$ still

546   recombines (Hough et al. 2014). As for the species discussed in the previous section,

547   it remains unclear whether any of the ancestrally autosomal genes have become

548   fully sex-linked.

549

550   **What might cause a difference in recombination between *S. latifolia* and *S. dioica*?**

551   The mechanism causing the apparent difference in recombination between *S.*

552   *dioica* and *S. latifolia* is currently unknown, and it is not yet clear whether a single

553   change was involved, such as lower crossing over in a single region proximal to the

554   fully sex-linked region (leaving the genetic map distances of most intervals

555   unchanged), or several changes. In *S. latifolia*, the PAR has undergone

556 rearrangements, but the current order of the PAR boundary region genes appears to

557 be the same in the X and Y (Qiu et al. 2016), which suggests that they became closely

558 linked to the *S. latifolia* MSY after the rearrangements, which therefore probably did

559 not cause the changed recombination.

560   We can also probably exclude the inversion that has been detected in the *S.*

561 *dioica* X chromosome. This inversion is not shared with the putative outgroup

562 species, *S. diclinis*, so *S. dioica* is thought to have the derived state (Nicolas et al.

563 2005); moreover, the inversion is probably confined to *S. dioica*, because the order

564 of the genes that allowed the inversion to be detected differs from that in *S. latifolia*,

565 which is the same as that in the more distant outgroup *S. vulgaris*. The extent of the

566 inversion is not yet known, as these genes belong to the younger stratum (stratum II

567 of Bergero et al. 2013), and it could extend some way into the older stratum, and/or

568 into the PAR. The difference we detect affects recombination between PAR and fully

569 Y-linked variants, and must therefore be due to a difference in male meiosis. If the

570 inversion lies wholly within the X-linked region, recombination in regions near the

571 PAR boundary would probably not be affected in males. If, however, it extends into

572 the PAR, recombination events in *S. dioica* heterozygotes might shift to more distal

573 parts of the chromosome (Henzel et al. 2011), causing genes that are closely linked

574 to the *S. latifolia* PAR boundary to show complete sex-linkage in *S. dioica*, the

575 opposite of the effect we observe. Finally, our limited results from *S. diclinis*

576 (probably lacking the inversion) show that the orthologues of the genes studied here

577 probably recombine with the MSY, as in *S. dioica*, contrasting with the closer linkage

578 in *S. latifolia.*

579   It also seems unlikely that a chromosome rearrangement such as an inversion

580 could be involved, because the PAR gene sequences appear to recombine in both *S.*

581 *dioica* and *S. latifolia*. All the PAR boundary genes we sequenced included variants

582 shared between *S. latifolia* and *S. dioica* (consistent with their low K_{ST} values

583 between these species, see above). This indicates that gene flow between the

584 species occurs in the genome region that we studied, arguing against an inversion,

585 since a rearrangement would impede gene flow. Moreover, the *S. dioica* sequences

586 containing variants shared with *S. latifolia* are not intact, unrecombined copies of *S.*

587 *latifolia* haplotypes (see Figures S1 and S2). Rather, a few sites in the *S. dioica*

588 sequences show clear signs of sex-linkage, with variants appearing either X- or Y-

589 linked, while other sites in the same genes have the variant that appears X-linked in

590 *S. latifolia*, or the Y-linked variant.

591   Overall, therefore, we conclude that the PAR boundary genes studied here

592 probably did not immediately become fully sex-linked when they were added to the

593 sex chromosomes, but that reduced recombination has subsequently evolved, so

594 that the region now recombines rarely with the Y-linked region, particularly in *S.*

595 *latifolia*.

596   What might have led to the difference in recombination between the closely

597 related species *S. latifolia* and *S. dioica*? It is possible that some general pressure

598 promotes the evolution of reduced recombination. As outlined above, one possibility

599 is that the ultimate cause is a selective pressure due to a sexually antagonistic

600 polymorphism at a locus in the genome region close to the PAR boundary (requiring

601 that the gene and the loci studied here are very closely linked to the boundary, as

602 reviewed in Qiu et al. 2016); the observed allele frequency differences between the

603 sexes are consistent with the presence of such a polymorphism in *S. latifolia*. This

604 variant might not have become established in *S. dioica*, explaining the difference.

605 However, an equally plausible alternative is that such a polymorphism arose in an

606 ancestor of both species, but a response to the selection for closer linkage occurred

607 in only one species, due to a lack of genetic variation for the recombination rate in

608 this region in the other. Alternatively, the change in recombination could be caused

609 by a non-selective force. For example, it has been hypothesized that fully sex-linked

610 regions may expand, at the expense of PARs, by an automatic process that

611 redistributes crossovers towards the distal regions of sex bivalents as the sex-specific

612 region differentiates and heteromorphism evolves, causing chromosome asymmetry

613 (Henzel et al. 2011). Again, this might differ between closely related species.

614   Understanding the proximate mechanism(s) involved may help to distinguish

615 between these different possibilities leading to suppressed recombination between

616 sex chromosomes. The presence of sexually antagonistic polymorphisms may favour

617 chromosome rearrangements or other major factors controlling recombination and

618 preventing recombination across regions carrying many genes (Charlesworth and

Charlesworth 1980), while the alternative just outlined, which could occur without SA effects being involved, might cause gradual extension of the border of the non-recombining region into the PAR. Our results establish *S. latifolia* as a species in which it may be possible to learn in detail about changes that are extending the sex chromosomes' non-recombining regions, at the expense of the PAR.

DATA ARCHIVING: The sequence data are available in Dryad under accession number [TO BE ADDED].

Supplementary information is available at (the journal's name TO BE ADDED)'s website

CONFLICT OF INTEREST: The authors declare no conflict of interest.

LITERATURE CITED

Bachtrog D (2006) A dynamic view of sex chromosome evolution. *Curr. Opin. Genet. Dev.* **16:** 578-585.

Barlow BA, and Wiens D (1976) Translocation heterozygosity and sex ratio in *Viscum fischeri*. *Heredity* **37:** 27-40.

Bergero R, Forrest A, Kamau E and Charlesworth D (2007) Evolutionary strata on the X chromosomes of the dioecious plant *Silene latifolia*: evidence from new sex-linked genes. *Genetics* **175:** 1945-1954.

Bergero R, Qiu S, Forrest A, Borthwick H and Charlesworth D (2013) Expansion of the pseudoautosomal region and ongoing recombination suppression in the *Silene latifolia* sex chromosomes. *Genetics* **194:** 673-686.

650   Brooke M, Welbergen J, Mainwaring M, Van-Der-Velde M, Harts A *et al.* (2010)

651   Widespread translocation from autosomes to sex chromosomes preserves

652   genetic variability in an endangered lark. *J. Mol. Evol.* **70:** 242-246.

653   Bull JJ, (1983) *Evolution of Sex Determining Mechanisms*. Benjamin/Cummings,

654   Menlo Park, CA.

655   Castillo E, Bidau C and Marti D (2010) Neo-sex chromosome diversity in Neotropical

656   melanopline grasshoppers (Melanoplinae, Acrididae). *Genetica* **138:** 775-786.

657   Cernohorska H, Kubickova S, Kopecna O, Vozdova M, Matthee C *et al.* (2015) Nanger,

658   Eudorcas, Gazella, and Antilope form a well-supported chromosomal clade

659   within Antilopini (Bovidae, Cetartiodactyla). *Chromosoma* **124:** 235-247.

660   Charlesworth B, Nordborg M and Charlesworth D (1997) The effects of local

661   selection, balanced polymorphism and background selection on equilibrium

662   patterns of genetic diversity in subdivided inbreeding and outcrossing

663   populations. *Genet. Res.* **70:** 155-174.

664   Charlesworth D, and Charlesworth B (1980) Sex differences in fitness and selection

665   for centric fusions between sex-chromosomes and autosomes. *Genetical*

666   *Research (Cambridge)* **35:** 205-214.

667   Chibalina M, and Filatov D (2011) Plant Y chromosome degeneration is retarded by

668   haploid purifying selection. *Curr. Biol.* **21:** 1475–1479.

669   Cortez D, Marin R, Toledo-Flores D, Froidevaux L, Liechti A *et al.* (2014) Origins and

670   functional evolution of Y chromosomes across mammals. *Nature* **508:** 488–493.

671   Cotter DJ, Brotman SM and Sayres MW (2016) Genetic diversity on the human X

672   chromosome does not support a strict pseudoautosomal boundary. *Genetics*

673   **203:** 485-492.

674   Grabowska-Joachimiak A, and Joachimiak A (2002) C-banded karyotypes of two

675   Silene species with heteromorphic sex chromosomes. *Genome* **45:** 243-252.

676   Grabowska-Joachimiak A, Kula A, Książczyk T, Chojnicka J, Sliwinska E *et al.* (2015)

677   Chromosome landmarks and autosome-sex chromosome translocations in

678   *Rumex hastatulus*, a plant with XX/XY1Y2 sex chromosome system. *Chromosome*

679   *Res.* **23:** 187-197.

680   Grabowska-Joachimiak A, Mosiolek M, Lech A and Góralski G (2011) C-Banding/DAPI

681   and *in situ* hybridization reflect karyotype structure and sex chromosome

682   differentiation in *Humulus japonicus* Siebold & Zucc. *Cytogenet Genome*

683   *Research* **132:** 203–211.

684   Henzel JV, Nabeshima K, Schvarzstein M, Turner B, Villeneuve AM *et al.* (2011) An

685   asymmetric chromosome pair undergoes synaptic adjustment and crossover

686   redistribution during *Caenorhabditis elegans* meiosis: implications for sex

687   chromosome evolution. *Genetics* **197:** 685–699.

688   Hough J, Hollister JD, Wang W, Barrett SCH and Otto SP (2014) Genetic degeneration

689   of old and young Y chromosomes in the flowering plant *Rumex hastatulus*. *Proc.*

690   *Natl. Acad. Sci. U. S. A.* **111:** 7713–7718.

691   Howell EC, Armstrong S and Filatov D (2009) Evolution of neo-sex chromosomes in

692   *Silene diclinis Genetics* **182:** 1109 - 1115.

693   Hu X, and Filatov DA (2015) The large-X effect in plants: Increased species divergence

694   and reduced gene flow on the Silene X-chromosome. *Mol. Ecol.*

695   Hudson RR, Boos DD and Kaplan NL (1992) A statistical test for detecting geographic

696   subdivision. *Mol. Bio. Evol.* **9:** 138-151.

697   Hunziker J, Wulff A and Escobar A (2002) Permanent translocation heterozygosity in

698   dioecious *Baccharis coridifolia* DC. (Asteraceae). *Hereditas* **137:** 132-139.

699   Jordan C, and Charlesworth D (2012) The potential for sexually antagonistic

700   polymorphism in different genome regions. *Evolution* **66:** 505–516.

701   Kaiser VB, Bergero R and Charlesworth D (2009) *SlCyt*, a newly identified sex-linked

702   gene, has recently moved onto the X chromosome in *Silene latifolia*

703   (Caryophyllaceae). *Mol. Bio. Evol.* **26:** 2343 - 2351.

704   Karrenberg S, and Favre A (2008) Genetic and ecological differentiation in the

705   hybridizing campions *Silene dioica* and *S. latifolia*. *Evolution* **62:** 763-773.

706   Kirkpatrick M, and Guerrero R (2014) Signatures of sex-antagonistic selection on

707   recombining sex chromosomes. *Genetics* **197:** 531-541.

708   Laporte V, Filatov DA, Kamau E and Charlesworth D (2005) Indirect evidence from

709   DNA sequence diversity for genetic degeneration of Y-chromosome in dioecious

710   species of the plant *Silene*: the *SlY4/SlX4* and *DD44*-X/*DD44*-Y gene pairs. *J. Evol.*

711   *Biol.* **18:** 337-347.

712   Librado P, and Rozas J (2009) DnaSP v5: a software for comprehensive analysis of

713   DNA polymorphism data. *Bioinformatics* **25:** 1451-1452.

714   Minder A, Rothenbuehler C and Widmer A (2007) Genetic structure of hybrid zones
715   between *Silene latifolia* and *Silene dioica* (Caryophyllaceae): evidence for
716   introgressive hybridization. *Mol. Ecol.* **16:** 2504-2516.

717   Ming R, Bendahmane A and Renner S (2011) Sex chromosomes in land plants. *Annual*
718   *Review of Plant Biology* **62:** 485-514.

719   Muir G, Dixon CJ, Harper AL and Filatov DA (2012) Dynamics of drift, gene flow and
720   selection during speciation in *Silene*. *Evolution* **66:** 1447–1458.

721   Natri HM, Shikano T and Merilä J (2013) Progressive recombination suppression and
722   differentiation in recently evolved neo-sex chromosomes. *Mol. Bio. Evol.* **30:**
723   1131-1144.

724   Navajas-Pérez R, Herran RDL, Gonzalez G, Jamilena M, Lozano R *et al.* (2005) The
725   evolution of reproductive systems and sex-determining mechanisms within
726   *Rumex* (Polygonaceae) inferred from nuclear and chloroplastidial sequence data.
727   *Mol. Bio. Evol.* **22:** 1929-1939.

728   Nicolas M, Marais G, Hykelova V, Janousek B, Laporte V *et al.* (2005) A gradual
729   process of recombination restriction in the evolutionary history of the sex
730   chromosomes in dioecious plants. *PLoS Biol.* **3:** 47-56.

731   Pala I, Naurin S, Stervander M, Hasselquist D, Bensch S *et al.* (2012) Evidence of a
732   neo-sex chromosome in birds. *Heredity* **108:** 264–272.

733   Papadopulos AST, Chester M, Ridout K and Filatov DA (2015) Rapid Y degeneration
734   and dosage compensation in plant sex chromosomes. *Proc. Natl. Acad. Sci. U. S.*
735   *A.* **112:** 13021–13026.

736   Pennell M, Kirkpatrick M, Otto S, Vamosi J, Peichel C *et al.* (2015) Y fuse? Sex
737   chromosome fusions in fishes and reptiles. *PLoS Genetics* **11:** e1005237.

738   Prentice HC (1978) Experimental taxonomy of Silene section Elisanthe
739   (Caryophyllaceae): crossing experiments. *Bot. J. Linn. Soc.* **77:** 203–216.

740   Qiu S, Bergero R and Charlesworth D (2013) Testing for the footprint of sexually
741   antagonistic polymorphisms in the pseudo-autosomal region of a plant sex
742   chromosome pair. *Genetics* **194:** 663-672.

743   Qiu S, Bergero R, Forrest A, Kaiser VB and Charlesworth D (2010) Nucleotide diversity
744   in *Silene latifoli*a autosomal and sex-linked genes. *Proc. R. Soc. B* **277:** 3283-3290.

745 Qiu S, Bergero R, Guirao-Rico S, Campos JL, T Cezard *et al.* (2016) RAD-mapping

746   reveals an evolving, polymorphic and fuzzy boundary of a plant

747   pseudoautosomal region. *Mol. Ecol.* **25:** 414–430

748 Rautenberg A, Hathaway L, Oxelman B and Prentice HC (2010) Geographic and

749   phylogenetic patterns in Silene section *Melandrium* (Caryophyllaceae) as

750   inferred from chloroplast and nuclear DNA sequences. *Mol. Phylogenet. Evol.* **57:**

751   978-991.

752 Rejón CR, Jamilena M, Ramos MG, Parker JS and Rejón MR (1994) Cytogenetic and

753   molecular analysis of the multiple sex-chromosome system of *Rumex acetosa*.

754   *Heredity* **72:** 209-215.

755 Rice WR (1987) The accumulation of sexually antagonistic genes as a selective agent

756   promoting the evolution of reduced recombination between primitive sex-

757   chromosomes. *Evolution* **41:** 911-914.

758 Siroky J, Lysak MA, Dolezel J, Kejnovsky E and B.Vyskot (2001) Heterogeneity of rDNA

759   distribution and genome size in *Silene* spp. *Chromosome Res.* **9:** 387-393.

760 Smith BW (1964) The evolving karyotype of *Rumex hastatulus*. *Evolution* **18:** 93-104.

761 Städler T, Haubold B, Merino C, Stephan W and Pfaffelhuber P (2009) The impact of

762   sampling schemes on the site frequency spectrum in nonequilibrium subdivided

763   populations. *Genetics* **182:** 205-216.

764 Tajima F (1989) Statistical method for testing the neutral mutation hypothesis.

765   *Genetics* **123:** 585-595.

766 Wang J, Na J, Yu Q, Gschwend AR, Han J *et al.* (2012) Sequencing papaya X and Yh

767   chromosomes reveals molecular basis of incipient sex chromosome evolution.

768   *Proc. Natl. Acad. Sci. U. S. A.* **109:** 13710-13715.

769 Zhou Q, Zhang J, Bachtrog D, An N, Huang Q *et al.* (2014) Complex evolutionary

770   trajectories of sex chromosomes across bird taxa. *Science* **346:** 1332.

771 **Titles and legends to figures**

772 **Figure 1**. Evolution of *S. latifolia* PAR, showing the origins of the PAR boundary genes

773 studied in this paper (simplified from Figure 4 of Qiu et al. 2016, and including gene

774 locations only for these PAR boundary genes). **A**. The ancestral state before the

775 translocation events had a sex chromosome with a PAR (green region) and a region

776 that had stopped recombining (indicated in black region, but not as separate X-linked

777 and male-specific, or MSY, regions; different strata are not indicated in this diagram,

778 but the black line includes all fully sex-linked regions). The two autosomes that later

779 became translocated, moving the whole or part of these chromosome arms onto the

780 PAR of the ancestral sex chromosome, are represented as pink and blue lines. The

781 first and second translocation events are represented in parts **B** and **C**, with pink and

782 blue indicating the linkage groups in *S. vulgaris* that carry the PAR boundary genes,

783 when this is known. The *cs4991* gene is not indicated because it could not be

784 mapped in our *S. vulgaris* family; however its location in the genetic map of the *S.*

785 *latifolia* X chromosome is among the more distal genes added in the second

786 translocation event. Other genes added in both translocations have remained loosely

787 linked to the MSY boundary, and are not shown in the diagram.

788 .

789

790 **Figure 2**. Comparisons of nucleotide diversity in males and females of *S. latifolia*

791 (black and grey) and *S. dioica* (dark and pale pink). The estimated diversity values for

792 silent sites are shown above the x axis, and the differences in diversity between

793 males and females are shown below the axis. Note that there are no values for gene

794 *E521* in *S. dioica* because no variants were found in the sequenced region of this

795 gene in this species.

796 **Figure 3**. Heterozygote frequencies at polymorphic sites (A) and Tajima's *D* values (B)

797 for all site types in males and females of *S. latifolia* (black and grey, respectively) and

798 *S. dioica* (dark and pale pink). In panel A, the differences in heterozygote frequencies

799 between males and females are also shown.

800

801

802

803 **Supplementary files**

804 **Table S1.** Locations of the populations from which the *S. dioica*, *S. diclinis* and *S.*
805 *marizii* plants were collected.

806 **Table S2.** Divergence and K_{ST} estimates for *S. dioica* and *S. latifolia.*

807 **Table S3.** Tajima's D values in *S. dioica* and *S. latifolia.* For *S. latifolia*, results are
808 shown for the complete set of sequences, and for sub-sampled sets (see text).

809 **Figure S1.** Locations from which the plant samples were collected.

810 **Figure S2.** Non-singleton variants in gene *E559* in the four *Silene* species studied.

811 **Figure S3.** Non-singleton variants in gene *E523* in the four *Silene* species studied.

812 **Figure S4.** K_{ST} values with error bars showing the ranges of values in 40 sub-samples
813 of sequences, as described in the Methods section of the main text.

814 **Figure S5**. Heterozygote frequencies at non-singleton sites in 5 genes in *S. latifolia*
815 and *S. dioica* males (blue) and females (pink). Gene *E521* is not shown, because there
816 were no polymorphic sites in *S. dioica.* The other genes are shown in the order of
817 their genetic map locations on the X chromosome, as inferred by Qiu et al. (2015),
818 with *cs3597* closest to the fully sex-linked region.

Table 1. K_{ST} values and tests of subdivision between males and females in complete sequence sets and in sub-samples from *S. latifolia* with the same size as the *S. dioica* data. For the *S. dioica* results, P values are given for the significance tests using K^*_{ST} with 1000 permutations (Hudson et al. 1992), and significance is indicated by bold text; in *S. latifolia*, all the PAR boundary genes, but not the more distal *cs4991* gene, have K_{ST} values significantly different from zero with P < 0.0001. *S. latifolia* sub-samples were not created for gene *E521*, where there are no variants in *S. dioica*, so that we cannot compare the significance of subdivision between males and females, or for *cs4991*, where neither species has a K_{ST} value significantly different from zero.

Gene	*S. vulgaris* linkage group [1]	Sequence length [2]	K_{ST} values in complete sequence data sets, and significance test results			Sub-samples from *S. latifolia* data		
			S. latifolia K_{ST}	*S. dioica*		Mean K_{ST} in 40 sub-samples	Numbers of sub-samples with $K_{ST} \leq$ *S. dioica* value	Number of significant values in sub-samples (number with P < 0.01)
				K_{ST}	P values			
cs3597	SvLG12	177	**0.365**	**0.08**	**0.023**	**0.292**	0	40 (40)
E559	SvLGSmall	382	**0.194**	**0.079**	**0.003**	**0.168**	1	40 (24)
E521	SvLGSmall	120	**0.213**	— (no variants)		—	—	—
E523	SvLG9	486	**0.202**	0.032	0.057	**0.206**	0	40 (40)
cs935	SvLG9	388	**0.071**	0.02	0.13	**0.064**	1	28 (12)
cs4991	Not known	282	0.0158	0.00262	0.354	—	—	—

[1] From Qiu et al. 2015
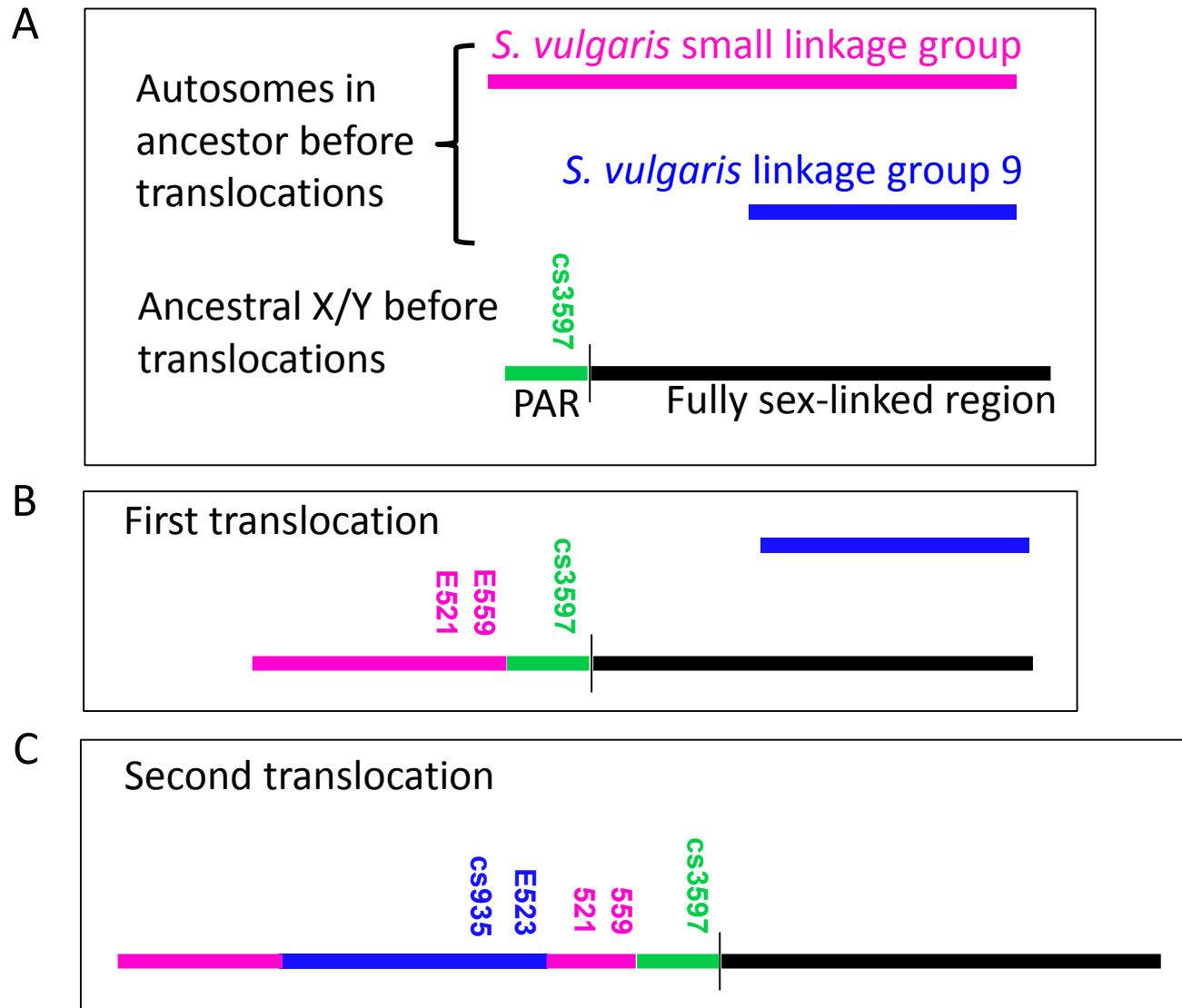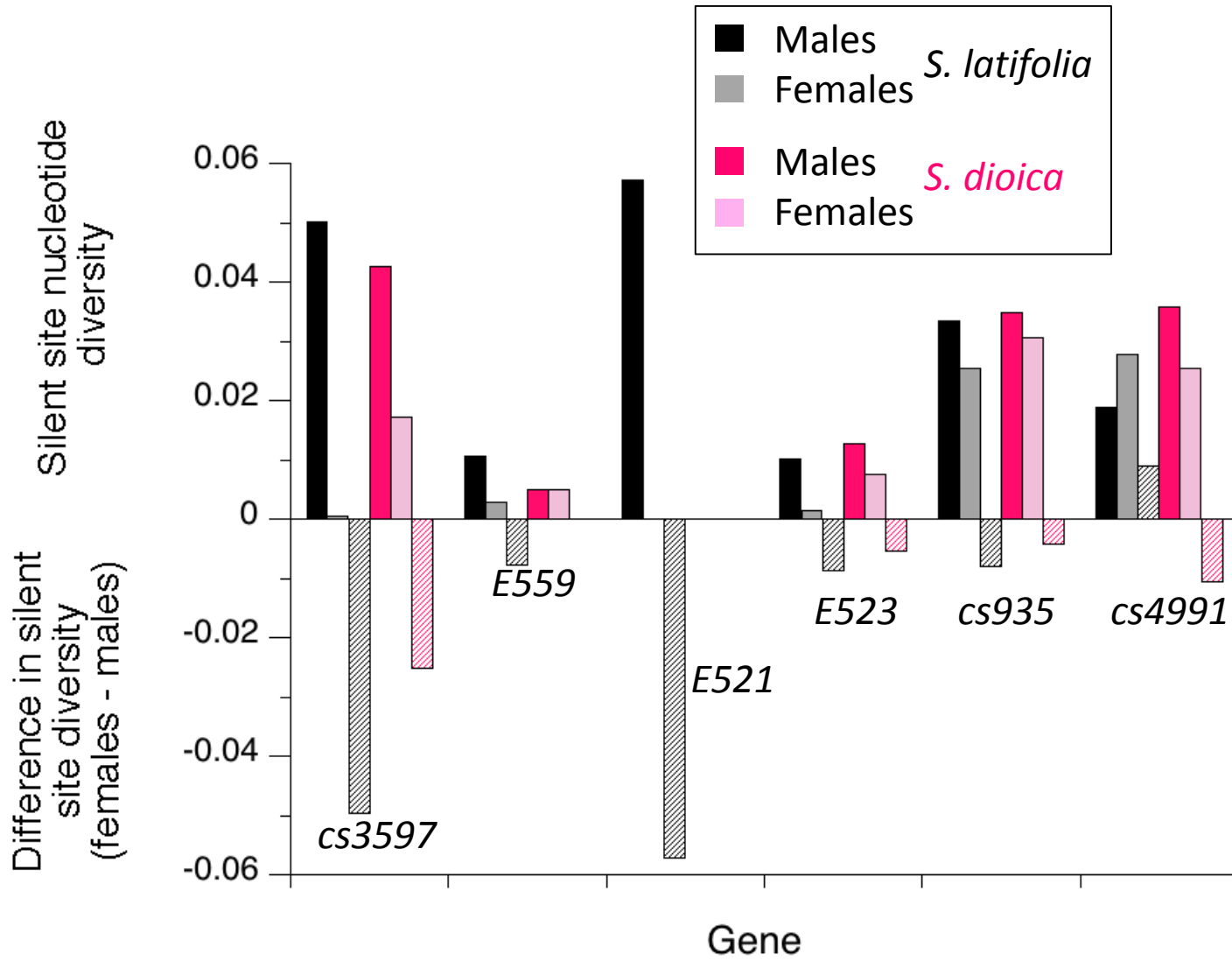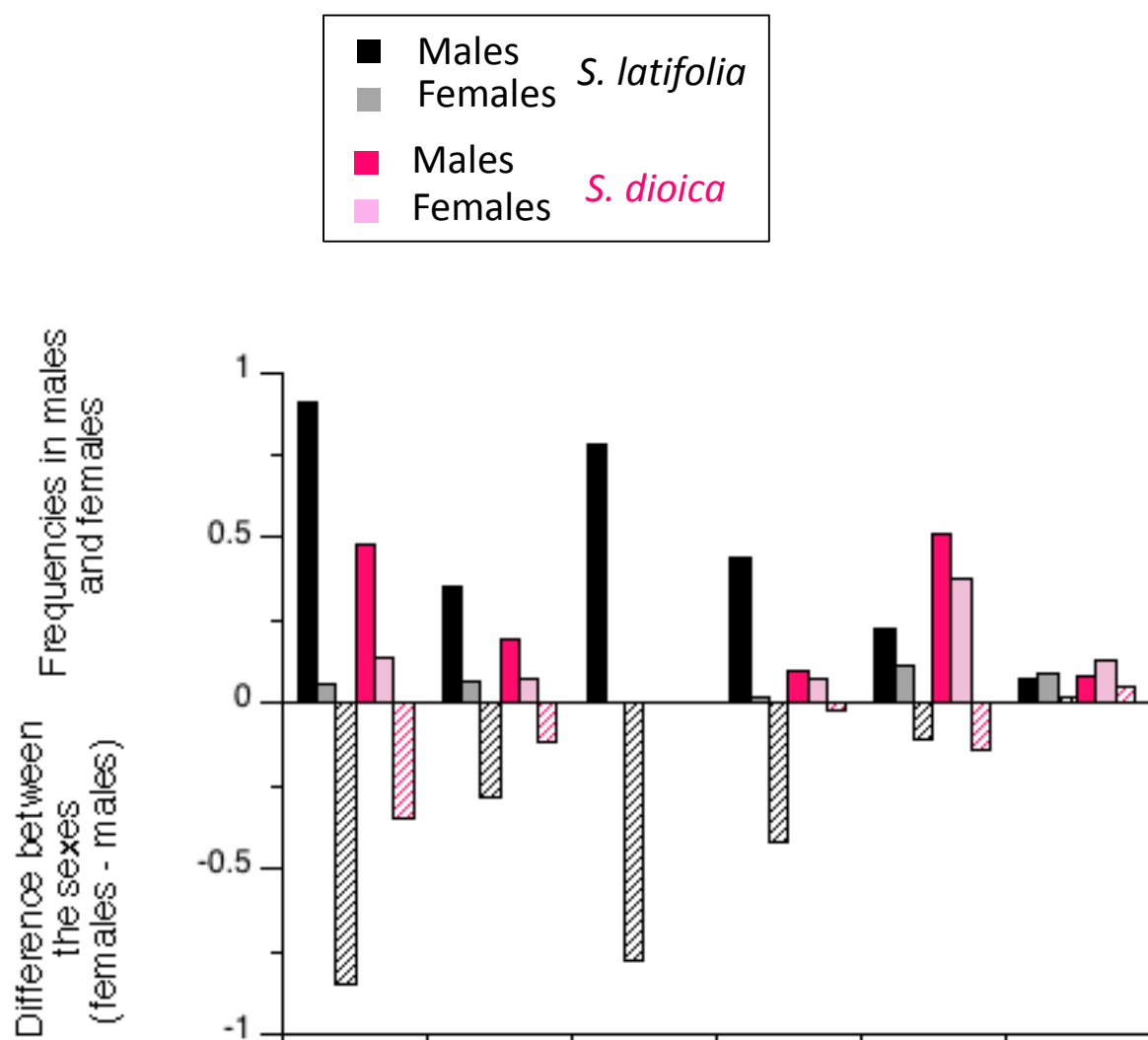[2] Number of sites excluding alignment gaps

Figure 1

Figure 2

Figure 3