

1 **The “two rules of speciation” in species with young sex chromosomes**

2

3 Dmitry A. Filatov

4

5 Email for correspondence: [Dmitry.Filatov@plants.ox.ac.uk](mailto:Dmitry.Filatov@plants.ox.ac.uk)

6

7 Address:

8 Department of Plant Sciences,

9 University of Oxford,

10 South Parks Road,

11 Oxford OX1 3RB

12 United Kingdom

13

14 Running title: Y-degeneration and faster-X in *Silene* speciation

15

16 Keywords: Haldane’s rule, large-X effect, faster-X effect, gene expression, Y-chromosome  
17 degeneration, speciation, plant sex chromosomes, gametophyte, haploid expression.

18

19

## 20 **Abstract**

21 The two “rules of speciation”, Haldane’s rule (HR) and the large-X effect (LXE), are thought to be  
22 caused by recessive species incompatibilities exposed in the phenotype due to the hemizyosity of X-  
23 linked genes in the heterogametic sex. Thus, the reports of HR and the LXE in species with recently  
24 evolved non- or partially-degenerate Y-chromosomes, such as *Silene latifolia* and its relatives, were  
25 surprising. Here I argue that rapid species-specific degeneration of Y-linked genes and associated  
26 adjustment of expression of X-linked gametologs (dosage compensation) may lead to rapid evolution  
27 of sex-linked species incompatibilities. This process is likely to be too slow in species with old  
28 degenerate Y-chromosomes (e.g. in mammals), but Y-degeneration in species with young gene-rich  
29 sex chromosomes may be fast enough to play a significant role in speciation. To illustrate this point I  
30 report the analysis of Y-degeneration and the associated evolution of gene expression on the X-  
31 chromosome of *Silene latifolia* and *Silene dioica*, a close relative that shares the same recently  
32 evolved sex chromosomes. Despite the recent ( $\leq 1$ MY) divergence of the two species, ~7% of Y-linked  
33 genes have undergone degeneration in one but not the other species. This species-specific  
34 degeneration appears to drive faster expression divergence of X-linked genes, which may account for  
35 HR and the LXE reported for these species. Furthermore, I suggest that “exposure” of autosomal or  
36 sex-linked recessive species incompatibilities in the haploid plant gametophyte may mimic the  
37 presence of HR in plants. Both haploid expression and species-specific Y-degeneration need to  
38 receive more attention if we are to understand the role of these processes in speciation.

39

## 40 **Introduction**

41 Many closely related species are known to form hybrids in nature, allowing them to exchange genes,  
42 which may slow down or prevent speciation and species divergence. The evolution of reproductive  
43 barriers is crucial for the speciation process to proceed and there is substantial evidence that sex  
44 chromosomes play a central role in the evolution of reproductive incompatibilities between incipient  
45 species (Larson *et al.* 2017; Laurie 1997; Masly & Presgraves 2007; Presgraves & Orr 1998; Turner  
46 *et al.* 2014). The importance of sex chromosomes in speciation is reflected in the “two rules of  
47 speciation”: Haldane’s rule and the large-X effect (Coyne & Orr 1989). Haldane’s rule (HR) states that  
48 the heterogametic sex is more likely to exhibit inviability or infertility in inter-specific hybrids, compared  
49 to the homogametic sex (Haldane 1922; Orr 1997). The large-X effect (LXE) posits that the X-

50 chromosome plays a disproportionately large role in hybrid dysfunction, compared to the autosomes  
51 (Coyne & Orr 2004; Jiggins *et al.* 2001; Turelli & Moyle 2007). Despite the striking variety in sex  
52 determination systems and sex chromosomes (Bachtrog *et al.* 2014) these rules of speciation are  
53 surprisingly universal. Both preferential hybrid inviability in the heterogametic sex and reduced  
54 interspecific gene flow of X-linked compared to autosomal genes – a convenient proxy for the LXE –  
55 have been described across plant and animal systems (Dufresnes *et al.* 2016; Ellegren *et al.* 2012;  
56 Hu & Filatov 2016; Payseur *et al.* 2004). This suggests the same, general mechanisms may be  
57 responsible for the special role of sex chromosomes in speciation across animal and plant groups.

58 The underlying causes of the two rules of speciation are not certain, but several hypotheses were  
59 proposed to account for HR and the LXE (Delph & Demuth 2016; Laurie 1997; Orr 1997). It is thought  
60 that, at least partly, both rules arise from recessive hybrid incompatibility alleles exhibiting full effects  
61 in the heterogametic sex (the dominance theory, (Coyne & Orr 2004; Turelli & Orr 1995)). The other  
62 possible causes include X-chromosome misregulation in hybrids (Larson *et al.* 2017; Masly &  
63 Presgraves 2007), greater density of male sterility loci on the X compared to autosomes (Masly &  
64 Presgraves 2007), meiotic drive on the sex chromosomes (Frank 1991; Hurst & Pomiankowski 1991;  
65 Tao & Hartl 2003) and faster evolution of X-linked genes (“faster-X” theory), which is predicted to  
66 arise if adaptive mutations are partially recessive (Charlesworth *et al.* 1987; Mank *et al.* 2010; Vicoso  
67 & Charlesworth 2006, 2009). It was also proposed that commonly observed hybrid male sterility but  
68 not female sterility, is due to faster evolution of genes involved in spermatogenesis compared to  
69 oogenesis-related genes and stronger sexual selection on males than females (“faster males” theory  
70 (Wu & Davis 1993)), which can help explain HR in species with male heterogamety (Laurie 1997; Orr  
71 1997; Wu *et al.* 1996). Extensive experimental work has provided support to the dominance and the  
72 faster males theories, which are now regarded as the most plausible explanations to the “two rules of  
73 speciation” (Laurie 1997; Masly & Presgraves 2007; Presgraves & Orr 1998).

74 Many of the aforementioned hypotheses assume that the Y-chromosome is already degenerate,  
75 resulting in the exposure of X-linked recessive mutations in hemizygous males. However, sex  
76 chromosomes are highly evolutionary labile, with many species (beyond well-studied mammals and  
77 *Drosophila*) exhibiting non-degenerate or partially degenerate Y-chromosomes (Bachtrog *et al.* 2014).  
78 The species with non-differentiated (homomorphic) sex chromosomes tend to show weaker  
79 reproductive isolation between closely related species, compared to species with heteromorphic sex

80 chromosomes (Lima 2014; Presgraves & Orr 1998). Nevertheless, the reports that the two “rules of  
81 speciation” also apply to species with non- or partly degenerate Y-chromosomes (Brothers & Delph  
82 2010; Dufresnes *et al.* 2016; Hu & Filatov 2016; Presgraves & Orr 1998) raise doubt that  
83 hemizygoty of the X-linked genes is the universal cause underpinning HR and the LXE.

84 Here I propose and investigate an additional potential cause for the special role that sex  
85 chromosomes play in speciation – species-specific Y-degeneration that drives divergent evolution of  
86 compensatory mechanisms, such as species-specific dosage compensation (ssDoC) where an X-  
87 linked gametolog is upregulated due to degeneration of the Y-linked gametolog. Following Y  
88 degeneration and evolution of dosage compensation, interspecific hybridisation with another species  
89 where the Y-copy of a gene is still functional, would result in a range of under- and over-  
90 compensation in male hybrids. For example, combining a degenerate Y-linked gene with non-  
91 compensating X-linked gametologs in hybrids would result in only half the normal gene dose, while a  
92 hybrid with functional Y-linked copy and dosage compensated (up-regulated) X-linked gametolog  
93 would have  $3/2$  the normal gene expression for the sex linked gene. As such, the modest differences  
94 in gene expression that are sufficient to drive the evolution of a dosage compensation system may  
95 also be sufficient to play a significant role in hybrid dysfunction. Furthermore, Y-degeneration may  
96 lead to compensatory evolution in non-homologous sex-linked or autosomal genes involved in the  
97 same biochemical pathway or gene regulatory network to adjust for lower dosage at the sex-linked  
98 gene with a degenerate Y-linked gametolog. Interspecific hybridisation may combine a functional Y-  
99 linked gene from one species with the genes that co-evolved with the degenerate Y of another  
100 species, leading to expression levels which are detrimental for hybrid viability or fertility.

101 This model is similar, but not identical to the model of Moyle *et al.* (Moyle *et al.* 2010) describing the  
102 contribution of gene movement between the chromosomes to the LXE and HR. In particular, both  
103 models predict transgressive gene expression in the heterogametic sex in hybrids. However, the  
104 ssDoC model described here implies the evolution of gene expression, while the gene movement  
105 model (Moyle *et al.* 2010) only requires that genes be moved between the X and the autosomes but  
106 does not involve any evolution of gene expression at the particular genes. This difference determines  
107 the type of effects contributing to hybrid dysfunction in the two models: unlike the gene movement  
108 model (Moyle *et al.* 2010), the ssDoC model implies that the fitness effects in hybrids can be

109 considered under-dominant, as combining a functional Y-linked copy and a dosage compensated X-  
110 linked copy is deleterious. The implications of this process for speciation have not been studied.

111 Potentially, rapid gene loss from the Y-chromosome and the associated evolution of dosage  
112 compensation on the X-chromosome could cause the rapid accumulation of hybrid incompatibilities  
113 between closely related species or sub-species with recently evolved gene-rich sex chromosomes.  
114 The rate of genetic degeneration in the non-recombining regions of sex-specific Y(or W)-  
115 chromosomes depends on the number of functional genes that are linked together, with Y-  
116 degeneration proceeding at a slower rate as the number of Y-linked genes decreases (Bachtrog 2008;  
117 Charlesworth 2008). This results in the rapid loss of Y-linked genes at the early stages of sex  
118 chromosome evolution, with the rate of gene loss slowing down on older Y(or W)-chromosomes. Thus,  
119 Y-degeneration is likely to be too slow to create species incompatibilities in species with old  
120 degenerate Y-chromosomes (e.g. in mammals), but rapid Y-degeneration of young gene-rich sex  
121 chromosomes may be fast enough to play a significant role in speciation. To illustrate this point I  
122 report an analysis of Y-degeneration and the associated evolution of gene expression on the X-  
123 chromosome of *Silene latifolia* and a close relative, *Silene dioica*, which share recently evolved sex  
124 chromosomes.

125 Both *S. latifolia* and *S. dioica* are common throughout Europe, with *S. latifolia* inhabiting open  
126 fields and road margins, and *S. dioica* tending to be found in more shady and moist habitats, such as  
127 forests. Despite the difference in preferred habitat, which likely plays the primary role in isolation of  
128 these species (Favre *et al.* 2017), *S. latifolia* and *S. dioica* often form hybrid swarms in places of co-  
129 occurrence. These species have identical karyotypes with clearly distinguishable X and Y-  
130 chromosomes (Armstrong & Filatov 2008; Ciupercescu *et al.* 1990). Due to the relatively recent (~11  
131 million years ago) origin of dioecy and sex chromosomes in the ancestor of these species (Krasovec  
132 *et al.* 2018), *S. latifolia* is being used to study the early stages of sex chromosome evolution  
133 (Bernasconi *et al.* 2009; Charlesworth 2015). Several recent studies demonstrated that most *S.*  
134 *latifolia* X-linked genes still appear to have functional Y-linked gametologs, though some Y-  
135 degeneration is apparent (Bergero *et al.* 2015; Chibalina & Filatov 2011; Krasovec *et al.* 2018;  
136 Papadopulos *et al.* 2015). The loss of at least some *S. latifolia* Y-linked genes appears to be  
137 compensated by a higher expression of their X-linked gametologs (dosage compensation, (Muyle *et*  
138 *al.* 2012; Papadopulos *et al.* 2015)). Below I report that rapid, on-going Y-chromosome degeneration

139 and an evolving dosage compensation system are making *Silene* sex chromosomes diverge faster  
140 than the autosomes. I propose that a higher rate of sex chromosome divergence may be an important  
141 contributor of HR and the LXE reported for these species (Brothers & Delph 2010; Hu & Filatov 2016).

142

## 143 **Materials and Methods**

### 144 *Plant material*

145 Six *S. latifolia* and six *S. dioica* plants (three males and three females of each species) were grown  
146 in the glasshouse (20°C and 15h lighting) from seed collected in the wild. The females of both species  
147 have already been used in a previous study assessing the LXE in *S. latifolia* and *S. dioica* (Hu &  
148 Filatov 2016). In addition, one male plant from another closely related dioecious species, *Silene*  
149 *diclinis* (Hu & Filatov 2016) was used as an outgroup in some of the analyses. While *S. latifolia* and *S.*  
150 *dioica* are very common all over Europe, *S. diclinis* is a rare endemic narrowly restricted to Xativa in  
151 Valencia (Spain) and material for this species was very limited, hence only one accession was  
152 available for analysis. It is worth noting that although *S. diclinis*, is closely related to *S. latifolia* and *S.*  
153 *dioica*, its sex chromosomes were rearranged, resulting in the evolution of neo-sex-chromosomes in  
154 that species (Howell *et al.* 2009), which may have affected expression at some genes.

155

### 156 *RNA extraction and sequencing*

157 RNA was extracted from actively growing shoots and flower buds from all of the plants, as  
158 described previously (Hu & Filatov 2016; Papadopulos *et al.* 2015). Total RNA from plant tissue was  
159 extracted using a Qiagen RNeasy Plant Mini Kit, including the optional on-column DNase digestion.  
160 Isolation of mRNA, cDNA synthesis and high-throughput sequencing were conducted according to the  
161 standard Illumina RNA-Seq procedure at the Oxford Genomics Centre of the Wellcome Trust Center  
162 for Human Genetics (WTCHG, Oxford). High-throughput sequencing for each individual was  
163 conducted at WTCHG using an Illumina HiSeq2000 instrument with 100 base, paired-end reads. All  
164 sequence reads were submitted to SRA database under project number PRJNA453413.

165

### 166 *Expression analyses*

167 The reference transcriptome used for RNA-seq read mapping was taken from a previous study  
168 (Papadopulos *et al.* 2015). A significant advantage of that reference transcriptome is that genomic

169 sequences (along with RNA-seq data) were used to reconstruct Y-linked genes (Papadopulos *et al.*  
170 2015), while the transcriptomes in other studies were based entirely on RNA-seq data (Bergero &  
171 Charlesworth 2011; Chibalina & Filatov 2011; Muyle *et al.* 2012; Zemp *et al.* 2016), resulting in under-  
172 representation of weakly expressed Y-linked genes. Furthermore, a large number of genes (2,114) in  
173 that transcriptome were genetically mapped to 12 chromosomes on *S. latifolia* (Papadopulos *et al.*  
174 2015), while other studies (Bergero *et al.* 2015; Chibalina & Filatov 2011; Muyle *et al.* 2012; Zemp *et*  
175 *al.* 2016) have only classified genes into 'X-linked', non-X-linked, and 'unknown' bins, with varying  
176 degrees of uncertainty.

177 RNA-seq reads were mapped to the reference transcriptome and gene expression was measured  
178 using RSEM (Li & Dewey 2011) with default parameters. "Fragments per kilobase per million reads"  
179 (FPKM (Mortazavi *et al.* 2008)) values from RSEM were used in all the analyses of expression. The  
180 accuracy of this approach in distinguishing homologous X- and Y-linked alleles in *S. latifolia* was  
181 demonstrated in the study by Papadopulos *et al.* (Papadopulos *et al.* 2015). In particular, the  
182 "expression" of Y-linked alleles in females was zero, as expected if RSEM accurately distinguishes  
183 between X- and Y-linked gametologs (see Suppl. methods page 5 in (Papadopulos *et al.* 2015)). The  
184 mapping of different *Silene* species to the same transcriptome is justified by the low sequence  
185 divergence between these species (silent site divergence ~1.5%, which is similar to intra-specific  
186 polymorphism in *S. latifolia* [ $\pi_s \sim 1.5\%$ ] (Hu & Filatov 2016)), resulting in similar proportions (~80%) of *S.*  
187 *latifolia*, *S. dioica* and *S. diclinis* RNA-seq reads successfully mapping to the reference transcriptome.  
188 Importantly, the three species had similar proportions of reads mapping to sex-linked and autosomal  
189 genes, illustrating that mapping to a heterospecific transcriptome did not differentially affect sex-linked  
190 genes.

191 Per-gene expression divergence between the species ( $D_e$ ) was calculated separately for males  
192 and females.  $D_e$  was calculated as the difference between medians for expression in the two species,  
193 normalised by the average of the two medians (Meisel *et al.* 2012). As the aim of the analyses in the  
194 current study was to quantify the overall divergence, rather than the direction of change in expression,  
195 the absolute value of  $D_e$  was used throughout this paper.

196 Statistical analyses ( $\chi^2$ , Wilcoxon tests and box plots) of gene expression were done in R, except  
197 figure 1, which was made in Excel (Microsoft).

198

**199 Results****200 *The extent of Y-degeneration in S. latifolia and S. dioica***

201 To test the extent of genetic degeneration in *S. latifolia* and *S. dioica*, Y-linked gene expression of  
202 homologous X- and Y-linked genes in males (mX and mY, respectively) were analysed relative to the  
203 expression of X-linked genes in females (fXX). The comparison with female expression is used to  
204 avoid confounding Y-degeneration and dosage compensation, both of which would affect the  
205 comparison between mY and mX. The expression analysis of sex-linked genes reveals extensive  
206 genetic degeneration on the Y-chromosome of both *S. latifolia* and *S. dioica*. In particular, out of 982  
207 genes for which both X- and Y-linked gametologs are available from the previous work (Chibalina &  
208 Filatov 2011; Papadopoulos *et al.* 2015), over 1/3<sup>rd</sup> of the Y-linked genes show more than 10-fold  
209 reduction in expression compared to X-linked gametologs (404, 419 and 360 genes in *S. latifolia*, *S.*  
210 *dioica* and both species, respectively). This result is consistent with the previous reports of Y-  
211 degeneration in *S. latifolia* (Bergero *et al.* 2015; Krasovec *et al.* 2018; Papadopoulos *et al.* 2015) and  
212 extends the analysis to *S. dioica* where Y-degeneration has not been analysed previously.

213 The analysis above was based on genes with intact X- and Y-linked gametologs (including genes  
214 where the Y-copy is not expressed, as long as it is detectable in the genomic sequence (Papadopoulos  
215 *et al.* 2015)) and does not take into account any genes without a detectable Y-linked copy. There are  
216 246 such X-only genes among the *S. latifolia* sex-linked genes detected previously (Chibalina &  
217 Filatov 2011; Papadopoulos *et al.* 2015). Many of these genes may have lost the Y-linked gametolog  
218 because of on-going Y-chromosome degeneration, though some may have been translocated to the  
219 X-chromosome from an autosome and never had a Y-linked gametolog. Taking these X-only genes  
220 into account brings the proportion of X-linked hemizygous genes to ~50%.

221 If *S. latifolia* and *S. dioica* Y-chromosomes have continued to degenerate, one would expect to see  
222 Y-linked genes that have been lost since these two species diverged. Such recently lost genes should  
223 be species-specific, that is, be actively expressed in one species and non-functional in the other  
224 dioecious species. Indeed, there is a considerable number of Y-linked genes with species-specific or  
225 nearly species-specific expression (Fig. 1). In particular, in *S. latifolia* and *S. dioica* 46 and 42 Y-linked  
226 genes, respectively, show >10-fold reduction in gene expression compared to expression of X-linked  
227 gametologs in females (fXX), while expression of the Y-linked copy in the other species is >30% fXX.  
228 As all these Y-linked genes are actively expressed in the closely related dioecious outgroup *S. diclinis*

229 (Fig. S1), they likely represent species-specific loss of Y-linked gene expression in *S. latifolia* or *S.*  
230 *dioica*. The loss of different sets of genes from the Y-chromosome may contribute to the evolution of  
231 reproductive isolation between closely related species with young actively degenerating sex  
232 chromosomes.

233

#### 234 *Y-degeneration accelerates the evolution of X-linked gene expression*

235 If the loss of a Y-linked gene is “compensated” by the upregulation of its X-linked gametolog  
236 (dosage compensation), the loss of different sets of Y-linked genes in closely related species is  
237 expected to accelerate divergence in gene expression for X-linked genes. Evidence of dosage  
238 compensation in *S. latifolia* sex chromosomes has been reported in two previous papers (Muyle *et al.*  
239 2012; Papadopulos *et al.* 2015), though two other analyses found no dosage compensation in that  
240 species (Bergero *et al.* 2015; Chibalina & Filatov 2011). To test whether the loss of different sets of  
241 genes from the Y-chromosomes of the two species accelerates divergence in X-linked gene  
242 expression, the analysis focused on 88 (=46+42, see previous section) sex-linked genes that are  
243 inferred to have lost Y-linked gametolog expression following the divergence of *S. latifolia* and *S.*  
244 *dioica*. Interestingly, the expression of X-linked gametologs of such genes in males is consistently  
245 higher in the species where the Y-copy is already degenerate (Fig. S2). In particular, the 46 sex-  
246 linked genes that lack expression of the Y-linked copy in *S. latifolia*, but are still actively expressed in  
247 *S. dioica*, show significantly higher expression of the X-linked gametolog in *S. latifolia* males,  
248 compared to *S. dioica* males (mean FPKM =  $14.7 \pm 33.04$  vs  $5.4 \pm 9.42$ ; paired t-test,  $P = 0.0069$ ).  
249 Conversely, for the 42 genes that lacked expression of the Y-linked copy in *S. dioica*, but not in *S.*  
250 *latifolia*, the former species shows significantly higher expression of X-linked gametologs, compared  
251 to the latter (mean FPKM =  $14.5 \pm 21.08$  vs  $8.9 \pm 14.82$ ; paired t-test,  $P = 0.009$ ). Thus, degeneration  
252 of different sets of Y-linked genes in *S. latifolia* and *S. dioica* is associated with the evolution of  
253 divergent gene expression on the X-chromosome, which may be partly responsible for the large-X  
254 effect in these species.

255

#### 256 *Faster-X divergence for sex-linked and autosomal genes in Silene*

257 A faster evolution of X(or Z)-linked genes compared to autosomal genes has been reported in  
258 many animal species (Meisel & Connallon 2013) and is thought to contribute to the occurrence of HR

259 and the LXE. In most cases, the evidence for faster-X comes from sequence-based comparisons of  
260 evolutionary rates on X(or Z)-linked and autosomal genes, with the “rate” measured as  
261 nonsynonymous divergence ( $dN$ ) normalised by synonymous divergence ( $dS$ ) (Charlesworth *et al.*  
262 2018; Meisel & Connallon 2013). Such a comparison for divergence between *S. latifolia* and *S. dioica*  
263 also reveals significantly higher  $dN/dS$  ratio for X-linked compared to autosomal genes (Fig. 2A;  
264 Wilcoxon rank sum test,  $W = 84278$   $P = 0.00145$ ). However, there are far fewer studies comparing  
265 evolutionary rates at gene expression level. To test whether gene expression diverges faster for the  
266 X-linked compared to autosomal genes, I quantified gene expression divergence between *S. latifolia*  
267 and *S. dioica* as the absolute value of the difference between median expression in the two species  
268 normalised by their average ( $D_e$  (Meisel *et al.* 2012)). The X-linked genes showed significantly higher  
269 expression divergence (Wilcoxon two sided test  $P < 0.00001$ ), compared to 1787 autosomal genes  
270 genetically mapped in the previous study (Papadopulos *et al.* 2015). This was the case regardless of  
271 the sex analysed, though the signal in males was stronger (Fig. 2B). The Y-linked genes showed  
272 much higher expression divergence compared to the X-linked and autosomal genes in males.  
273 Provided the time of species divergence is the same for different chromosomes, I hereby refer to  
274 higher divergence as “faster” divergence. Faster expression divergence of the sex-linked genes was  
275 detectable regardless of the gene expression level (Fig. S3), thus it cannot be explained by the over-  
276 or under-representation of highly expressed genes on the X-chromosome. Furthermore, excluding  
277 weakly expressed genes (Fig. S4) or genes with significantly sex-biased expression does not change  
278 the conclusion (Fig. S5).

279 Faster expression divergence of X-linked genes in *Silene* may be caused by a range of factors that  
280 can be classified into selective and non-selective (e.g. demographic). Selective hypotheses for faster-  
281 X expression divergence include evolution of dosage compensation driven by selection to  
282 compensate for loss of rapidly degenerating Y-chromosome, as well as the classic arguments of the  
283 “faster-X” theory (Charlesworth *et al.* 1987). On the other hand, non-selective explanations include the  
284 difference in effective population size ( $N_e$ ) – the smaller  $N_e$  for the X-linked genes enables a larger  
285 fraction of weakly deleterious mutations to be fixed by drift (Mank *et al.* 2010; Vicoso & Charlesworth  
286 2009), result in faster expression divergence on X-chromosome compared to autosomes. Another  
287 non-selective factor – differential interspecific gene flow between the X and the autosomes reported  
288 for *S. latifolia* and *S. dioica* (Hu & Filatov 2016), may also result in faster expression divergence for X-

289 linked compared to autosomal genes. The non-selective and selective hypotheses for faster  
290 expression divergence on the X-chromosome make different predictions and can potentially be  
291 distinguished from each other. In particular, the non-selective explanations apply equally to all X-  
292 linked genes regardless of the presence of functional Y-linked gametologs for the particular X-linked  
293 genes. On the other hand, both evolving dosage compensation hypothesis and the classic “faster-X”  
294 theory predict that X-linked genes that lost Y-linked gametologs should diverge faster compared to  
295 genes actively expressed on both the X- and the Y-chromosomes.

296 To distinguish between the non-selective and selective hypotheses, I compared divergence  
297 between *S. latifolia* and *S. dioica* males for 334 hemizygous (no expression from the Y in either  
298 species; hereafter  $X_{noYexpr}$ ) and 399 non-hemizygous (Y-copy is actively expressed in both species;  
299 hereafter  $X_{Yexpr}$ ) X-linked genes. Consistent with the selective hypotheses, the  $X_{noYexp}$  genes show  
300 significantly faster expression divergence between *S. latifolia* and *S. dioica* males compared to  $X_{Yexpr}$   
301 genes (Fig 3A; Wilcoxon rank sum test  $W = 23019$ ,  $P = 0.0082$ ). The rate of sequence divergence  
302 (dN/dS) shows the same trend, though the difference between  $X_{noYexp}$  and  $X_{Yexpr}$  genes is not  
303 significant (Fig 3B; Wilcoxon rank sum test  $W = 11302$ ,  $P = 0.632$ ).

304

#### 305 *Does haploid expression affect species divergence?*

306 The X-chromosome is thought to be “special” because X-linked recessive mutations are expressed  
307 in the phenotype in hemizygous males (Charlesworth *et al.* 1987; Vicoso & Charlesworth 2006).  
308 However, in plants, a significant proportion of genes in the genome are expressed at the haploid  
309 stage of lifecycle – the male gametophyte (pollen in Angiosperms) (Honys & Twell 2004). Thus, if  
310 interspecific hybrid incompatibilities are primarily recessive, then the plant genes expressed in the  
311 gametophyte may be expected to play a similarly “special” role in plant speciation as the sex  
312 chromosomes in animals. The availability of pollen expression data for *S. latifolia* (Chibalina & Filatov  
313 2011) makes it possible to test whether exposure to haploid selection in the plant gametophyte affects  
314 the rate of gene expression divergence between the species.

315 The proportion of genes that evolved differential expression between *S. latifolia* and *S. dioica* is  
316 significantly lower for genes expressed in the gametophyte ( $G_{expr}$ ), compared to genes with a  
317 predominantly sporophytic expression ( $S_{expr}$ ; 9.9% and 12.3%, respectively; 2x2 contingency  $\chi^2=26.12$ ,  
318  $P < 0.000001$ ). Furthermore,  $D_e$  is significantly lower for  $G_{expr}$  compared to  $S_{expr}$  genes (Fig 4A;

319 Wilcoxon rank sum test  $P < 10^{-6}$ ), indicating that stronger purifying selection at the haploid stage of  
320 the lifecycle prevents expression divergence between species for most genes. In animals, the genes  
321 with functions related to male gametogenesis, such as accessory gland proteins in *Drosophila*, are  
322 often reported to show faster evolutionary rates, possibly due to positive selection fuelled by sexual  
323 conflict (e.g. (Ahmed-Braimah *et al.* 2017)). The *S. latifolia* and *S. dioica* genes that are over-  
324 expressed in pollen ( $P_{\text{expr}}$ ; the category includes genes with at least five-fold higher expression in  
325 pollen compared to male somatic tissues) also show accelerated expression divergence, compared to  
326 non-overexpressed genes; in particular, the distribution of  $D_e$  for  $P_{\text{expr}}$  genes is significantly shifted  
327 upwards compared to  $G_{\text{expr}}$  and  $S_{\text{expr}}$  genes (Fig 4A; Wilcoxon rank sum test  $P < 10^{-6}$  for both  
328  $P_{\text{expr}}:G_{\text{expr}}$  and  $P_{\text{expr}}:S_{\text{expr}}$  comparisons). Excluding weakly expressed genes does not change this  
329 result (Fig. S6A).

330 The distribution of  $P_{\text{expr}}$  genes across the X, Y and autosomes did not deviate from random. On the  
331 other hand,  $G_{\text{expr}}$  genes were significantly depleted on the X- and Y-chromosomes (83.6% and 62.1%  
332 of the expected, respectively; G-test,  $P < 0.001$ ), and  $S_{\text{expr}}$  genes were significantly depleted on the  
333 autosomes (79.8% of the expected; G-test,  $P < 0.001$ ) and over-represented on the X- and Y-  
334 chromosomes (132.5% and 176.4% of the expected; G-test,  $P < 0.001$ ). To test whether the under-  
335 representation of slow evolving  $G_{\text{expr}}$  genes on the sex chromosomes and under-representation of  
336 fast-evolving  $S_{\text{expr}}$  genes on the autosomes may be the cause of accelerated expression divergence  
337 on the X-chromosome (Fig. 1), the  $D_e$  was calculated separately for sex-linked and autosomal  $G_{\text{expr}}$   
338 and  $S_{\text{expr}}$  genes. As a faster evolution of sex-linked genes compared to those that are autosomal is  
339 still clearly detectable (Figures 4B and S6B) and significant (Wilcoxon rank sum test  $P < 0.0001$ ) for  
340 the  $G_{\text{expr}}$  and for  $S_{\text{expr}}$  genes separately, the uneven distribution of  $S_{\text{expr}}$  and  $G_{\text{expr}}$  genes across the  
341 chromosomes cannot explain faster expression divergence on the X- and Y-chromosomes compared  
342 to the autosomes.

343

## 344 Discussion

345 *Is the Silene Y-chromosome degenerate “enough”?*

346 Consistent with the dominance theory, species incompatibilities expressed in the heterogametic  
347 sex evolve faster in taxa with larger X-chromosomes (Turelli & Begun 1997). Although the X-  
348 chromosome in *S. latifolia* and *S. dioica* is relatively large (2<sup>nd</sup> largest chromosome in the genome

349 (Armstrong & Filatov 2008)), prior to this study, it had not been clear how many X-linked genes in *S.*  
350 *latifolia* and *S. dioica* comply with the assumption of the dominance theory that X-chromosome is  
351 hemizygous in the heterogametic sex. Due to the recent origin of sex chromosomes in the ancestor of  
352 *S. latifolia* and *S. dioica*, it was widely assumed that the Y-chromosome in these species is likely to be  
353 non-degenerate, which was supported by the isolation of apparently functional Y-linked genes in early  
354 low-throughput studies (reviewed by (Charlesworth 2008)). More recent analyses based on  
355 transcriptome (Bergero & Charlesworth 2011; Bergero *et al.* 2015; Chibalina & Filatov 2011) and  
356 genome (Papadopulos *et al.* 2015) sequence data reported various degrees of genetic degeneration  
357 (10 to 30%) of Y-linked genes in *S. latifolia*.

358 This study demonstrates that over 30% of Y-linked genes have effectively lost expression in either  
359 or both *S. latifolia* and *S. dioica*, leaving their X-linked gametologs hemizygous in males. The extent of  
360 Y-chromosome degeneration in *S. latifolia* and its close dioecious relatives may be considerably  
361 higher than 30% given that only genes with detectable X- and Y-linked gametologs were used in the  
362 analyses. The genes without detectable Y-linked gametologs (X-only genes) were excluded because  
363 an unknown proportion may represent translocations of autosomal genes to the X-chromosome rather  
364 than the loss from the Y-chromosome. Unfortunately, in the absence of chromosome-level assemblies  
365 of the genomes of *S. latifolia* and its non-dioecious relative, such as *S. vulgaris*, it is not possible to  
366 distinguish the X-ancestral genes that lost Y-linked gametologs from the genes that were translocated  
367 to the X and never had a Y-linked copy. A study that attempted to address this question using a  
368 comparative analysis of genetic maps of dioecious *S. latifolia* and non-dioecious *S. vulgaris* revealed  
369 that all 16 tested X-only genes of *S. latifolia* are X-ancestral rather than translocated to the X-  
370 chromosome secondarily (Bergero *et al.* 2015). Thus, the proportion of X-only genes translocated to  
371 the X-chromosome after the sex chromosomes evolved may be small, and most of the X-only genes  
372 are likely to have lost their Y-linked gametologs. When the X-only genes are included in the analysis,  
373 the proportion of Y-degenerate genes increases to about 50%.

374 The theory expressing the conditions for Haldane's rule as a function of  $p_x$ , the proportion of hybrid  
375 incompatibilities that are X-linked hemizygous ((Turelli & Orr 1995) equation B2), shows that with a  
376 smaller  $p_x$  the conditions for HR evolution become more restrictive (e.g. see fig. 1 in (Orr & Turelli  
377 1996)). The X-chromosome in *S. latifolia* and *S. dioica* contains about 10% of the *Silene* genome  
378 (Armstrong & Filatov 2008). If 50% of the Y-linked genes are degenerate, about 5% of the *Silene*

379 genome is hemizygous in males, which gives  $p_x \sim 0.1$  (from formula A3 in (Turelli & Begun 1997),  
380 assuming all incompatibilities involve two loci). To place this in the context of other species,  $p_x$  in  
381 *Silene* is larger than in mammals ( $p_x \sim 0.05$ ), but smaller than *Drosophila melanogaster*, where  $p_x \sim$   
382 0.36 (Orr & Turelli 1996). As both *Drosophila* and mammals comply with HR, it appears that the  
383 extent of hemizyosity in *S. latifolia* and *S. dioica* males is likely sufficient for the dominance theory to  
384 explain the presence of the Haldane's rule (Brothers & Delph 2010) in these species.

385

386 *Does on-going Y-degeneration contribute to the evolution of species divergence?*

387 Interestingly, 88 of the Y-linked genes analysed, have lost expression since the divergence of *S.*  
388 *latifolia* and *S. dioica*, demonstrating that genetic degeneration is rapidly progressing on Y-  
389 chromosomes of these species. For a comparison, only three human Y-linked genes have lost their  
390 function since they diverged from chimpanzees ~6 million years ago (Bellott *et al.* 2014). The  
391 comparison with human sex chromosomes is appropriate given the age of the youngest human  
392 stratum is similar to the age of *Silene* sex chromosomes after adjusting for the difference in  
393 generation time. Given the similar interspecific divergence at silent sites (~1.5%) in two species pairs  
394 (*S. latifolia/S. dioica* and *Homo sapiens/Pan troglodites*), genetic degeneration of Y-linked genes is  
395 proceeding at least 10 times faster in *Silene* compared to humans. On the other hand, the rate of Y-  
396 degeneration in *Silene* is comparable to that reported for recently evolved neo-Y chromosome of  
397 *Drosophila miranda* (Bachtrog *et al.* 2008).

398 Faster genetic degeneration is expected (e.g. (Bachtrog 2008)) for younger gene-rich Y-  
399 chromosomes (such as in *S. latifolia*), compared to the older stages when only a few functional genes  
400 remain on the Y-chromosome (as is the case in humans (Bellott *et al.* 2014)). This is the case  
401 because Y-chromosome degeneration is thought to be, at least partly, caused by interference of  
402 natural selection acting on multiple mutations linked together in the non-recombining region  
403 (Charlesworth 2008). With many functional genes linked together on the same Y-chromosome,  
404 natural selection is unable to eliminate deleterious mutations and fix advantageous mutations,  
405 resulting in gradual disfunctionalisation of Y-linked genes (Charlesworth & Charlesworth 2000). Thus,  
406 in species with young sex chromosomes, such as *S. latifolia* and *S. dioica*, the Y-chromosome  
407 contains hundreds to thousands of functional genes and Y-degeneration may be sufficiently rapid for  
408 closely related species to lose different sets of Y-linked genes. If the X-linked gametologs evolve

409 altered gene expression to compensate for the loss of Y-linked gametologs (i.e. gene-by-gene dosage  
410 compensation, such as reported for chicken (Mank & Ellegren 2009)), the introgressed X may not be  
411 compatible with the “local” Y-chromosome, resulting in the reduced fitness of hybrids. Such a  
412 reduction in fitness is expected to be present primarily in the heterogametic sex where the “foreign” Y  
413 meets the “local” X (or vice-versa). The expression analyses for *S. latifolia* and *S. dioica* sex-linked  
414 genes indicate that this “divergent Y degeneration” model is a plausible mechanism for the presence  
415 of HR and the LXE in species with young actively degenerating sex chromosomes.

416

417 *Y-degeneration, dosage compensation and faster-X evolution of gene expression in Silene*

418 Faster evolution of X-linked compared to autosomal genes may contribute to HR and the LXE  
419 because under the faster-X scenario recessive species incompatibilities would accumulate faster on  
420 the X-chromosome, resulting in a disproportionate contribution of the X-linked genes to reproductive  
421 barriers between the species (Coyne & Orr 1989). The analyses reported above provide compelling  
422 evidence for faster evolution of expression divergence in X-linked genes compared to autosomal  
423 genes. However, the causes of the accelerated evolution of gene expression on the *Silene* X-  
424 chromosome remain unclear.

425 An intriguing possibility explaining faster expression divergence in X-linked compared to autosomal  
426 genes is an on-going evolution of dosage compensation on the X-chromosome. The presence of  
427 dosage compensation on the *S. latifolia* sex chromosomes is a contested issue, with two studies  
428 finding no evidence supporting dosage compensation (Bergero *et al.* 2015; Chibalina & Filatov 2011)  
429 and two studies reporting the presence of at least partial dosage compensation in *S. latifolia* (Muyle *et*  
430 *al.* 2012; Papadopulos *et al.* 2015). If dosage compensation evolves in *S. latifolia* (and, by extension,  
431 *S. dioica*), this would accelerate gene expression divergence for X-linked genes, with selection to  
432 compensate for loss of Y-linked gametologs particularly affecting  $X_{\text{noYexpr}}$  genes. This is consistent  
433 with the observation of faster gene expression divergence at  $X_{\text{noYexpr}}$  compared to  $X_{\text{Yexpr}}$  genes. If  
434 dosage compensation in *Silene* evolves gene-by-gene (as opposed to chromosome-wide dosage  
435 compensation found in mammals and *Drosophila*), the X-linked genes with functional Y-linked  
436 gametologs would not be expected to be affected by this process, yet,  $X_{\text{Yexpr}}$  genes still show  
437 significantly faster evolution of gene expression compared to autosomal genes, suggesting some form  
438 of nascent chromosome-wide dosage compensation arising in these species.

439 If evolving dosage compensation is indeed the driver of faster-X evolution of gene expression in  
440 *Silene*, this could have broad implications on the speciation literature, in particular for species pairs  
441 with young sex chromosomes. Such species would be expected to show transgressive gene  
442 expression for sex-linked genes in male interspecific hybrids. Transgressive expression should be  
443 particularly pronounced for 'divergently degenerate' Y-linked genes – the genes that are lost in one  
444 and retained in the other hybridising species. These predictions need to be tested in future studies.  
445

#### 446 *The alternative explanations for faster-X evolution of gene expression in Silene*

447 The alternative explanations to faster-X gene expression driven by species-specific dosage  
448 compensation look less plausible. In particular, the non-selective explanations – the difference in  
449 effective population size ( $N_e$ ) or differential gene flow between the X-linked and autosomal genes are  
450 not compatible with the fact that  $X_{noYexpr}$  genes are diverging faster than  $X_{Yexpr}$  genes (Fig. 3).  
451 Furthermore, effective population size in *S. latifolia* and *S. dioica* is relatively large and likely  
452 comparable to that in *Drosophila* given the similar genetic diversity in these species (average  
453 heterozygosity at silent sites,  $\pi \sim 1.5\%$ ). For an  $N_e$  as large as in *Drosophila* purifying selection is  
454 expected to be highly efficient and the effect of a slightly smaller  $N_e$  for the X-chromosome should be  
455 marginal (Mank *et al.* 2010).

456 Faster-X evolution for gene expression in *Silene* is still detectable after exclusion of sex-biased  
457 genes (Fig. S5), or genes that are expressed or non-expressed in the gametophyte (Fig. 4B), or  
458 genes with high or with low expression only (Fig. S3). This indicates that none of these factors fully  
459 accounts for accelerated evolution of gene expression on the X-chromosome. However, the data  
460 presented above are compatible with the classic "faster-X" theory (Charlesworth *et al.* 1987) that  
461 predicts faster evolution of X-linked genes based on their female-biased transmission and exposure of  
462 X-linked recessive beneficial alleles in hemizygous males. Distinguishing between the classic faster-X  
463 theory and the species-specific dosage compensation hypothesis would require the analyses of gene  
464 expression in F1 interspecific hybrids, where transgressive segregation is predicted by the latter  
465 hypothesis. Unfortunately such data for crosses between *S. latifolia* and its close relatives is currently  
466 unavailable.

467 It is possible that the faster-X evolution in *Silene* has a composite nature, with accelerated positive  
468 selection, relaxed purifying selection, underrepresentation of slow evolving pollen-expressed genes

469 on the X-chromosome, differential gene flow on the X and autosomes and evolving dosage  
470 compensation all contributing to the observed faster evolution of gene expression on the X-  
471 chromosome. It remains unclear whether this “faster-X” for gene expression plays any role in  
472 speciation and future studies should test whether the *Silene* X-chromosome accumulates species  
473 incompatibilities faster than the autosomes.

474

475 *Does haploid expression of genes in the plant gametophyte contribute to speciation?*

476 The causes underlying HR and the LXE in *Silene* remain unclear, though the study that analysed  
477 the genetic basis of HR between *S. latifolia* and *S. diclinis* concluded that “the genetic architecture of  
478 Haldane’s rule in dioecious plants may differ from those commonly found in animals” (Demuth *et al.*  
479 2014). Widespread haploid expression of genes in the plant gametophyte may be one of the reasons  
480 for the difference in genetic architecture of HR between the two kingdoms. For example, in plants,  
481 hybrid male sterility may be caused by recessive species incompatibilities expressed in pollen, where  
482 over half of the genes in the genome are actively expressed (Honys & Twell 2004). These recessive  
483 species incompatibilities may not have anything to do with the sex chromosomes, but they can be  
484 interpreted as a manifestation of Haldane’s rule in species with heterogametic males. However, the  
485 first report of HR in plants – reduced pollen viability in hybrids between dioecious *Silene latifolia* and  
486 its close relatives (Brothers & Delph 2010), is unlikely to be the result of recessive hybrid  
487 incompatibilities in haploid pollen because the observation of hybrid male sterility in *Silene* was based  
488 on pollen stainability, a phenotype that is detectable before the haploid gene expression stage. Still,  
489 such mimicking of HR for hybrid sterility with autosomal recessive species incompatibilities expressed  
490 in the gametophyte may occur in other dioecious plants. More generally, the implications of haploid  
491 expression in plant gametophytes on hybrid incompatibilities remains unexplored. The evolutionary  
492 role of widespread haploid gene expression in plant gametophytes is poorly studied and its role in  
493 plant speciation remains unclear. Clearly, haploid expression in the plant gametophyte is potentially  
494 an important factor in plant evolution and it deserves more attention in the speciation literature.

495

## 496 **Conclusions**

497 The analyses reported above comprise the first demonstration of the “faster-X” effect in plants,  
498 though the underlying causes of the faster expression divergence on the *Silene* X-chromosome

499 appear different to the classic “faster-X” theory (Charlesworth *et al.* 1987). The most likely cause of  
500 the discovered “faster-X” in *Silene* is species-specific evolution of dosage compensation that is likely  
501 driven by on-going rapid degeneration of the Y-chromosome in *S. latifolia* and *S. dioica*. Although the  
502 connection between the observed “faster-X” evolution with HR and the LXE in *Silene* remains to be  
503 established, our results demonstrate that Y-degeneration and dosage compensations can be  
504 sufficiently rapid to proceed in species-specific manner, even between closely related species with  
505 young gene rich sex chromosomes. Potentially, this process may be a significant contributor of hybrid  
506 incompatibilities, though its role in speciation remains to be studied.

507

### 508 **Acknowledgements**

509 This work was supported by grants from the BBSRC (Grant BB/P009808/1 to DAF). I thank Bruno  
510 Nevado, Daven Presgraves and two anonymous reviewers for helpful comments, Michael Chester for  
511 proofreading the text and the staff at the Wellcome Trust Centre (Oxford) for sequencing and initial  
512 data processing.

513

### 514 **Data Accessibility**

515 The data used in this study are available from SRA database (project number PRJNA453413).

516

### 517 **Author Contributions**

518 DAF designed the study, generated and analysed the data and wrote the manuscript.

519

### 520 **References**

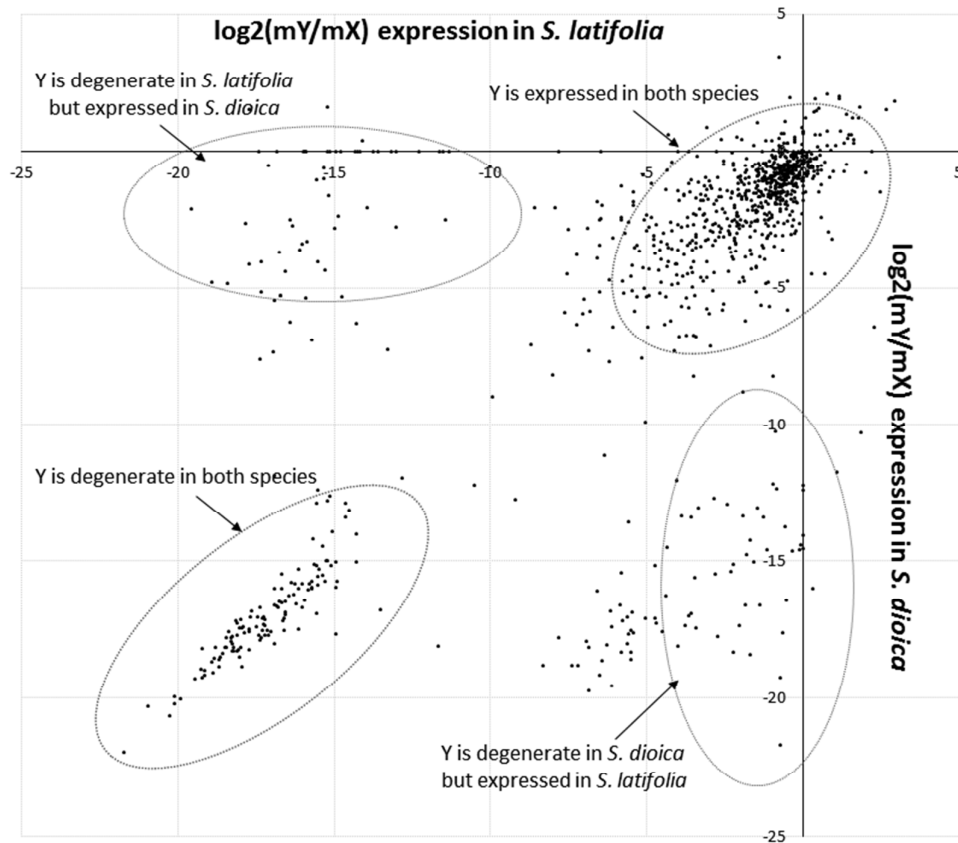
- 521 Ahmed-Braimah YH, Unckless RL, Clark AG (2017) Evolutionary dynamics of male reproductive  
522 genes in the *Drosophila virilis* subgroup. *G3 (Bethesda)* **7**, 3145-3155.
- 523 Armstrong SJ, Filatov DA (2008) A cytogenetic view of sex chromosome evolution in plants.  
524 *Cytogenet Genome Res* **120**, 241-246.
- 525 Bachtrog D (2008) The temporal dynamics of processes underlying Y chromosome degeneration.  
526 *Genetics* **179**, 1513-1525.
- 527 Bachtrog D, Hom E, Wong KM, Maside X, de Jong P (2008) Genomic degradation of a young Y  
528 chromosome in *Drosophila miranda*. *Genome Biol* **9**, R30.
- 529 Bachtrog D, Mank JE, Peichel CL, *et al.* (2014) Sex determination: why so many ways of doing it?  
530 *PLoS Biol* **12**, e1001899.
- 531 Bellott DW, Hughes JF, Skaletsky H, *et al.* (2014) Mammalian Y chromosomes retain widely  
532 expressed dosage-sensitive regulators. *Nature* **508**, 494-499.
- 533 Bergero R, Charlesworth D (2011) Preservation of the Y transcriptome in a 10-million-year-old plant  
534 sex chromosome system. *Curr Biol* **21**, 1470-1474.
- 535 Bergero R, Qiu S, Charlesworth D (2015) Gene loss from a plant sex chromosome system. *Curr Biol*  
536 **25**, 1234-1240.

- 537 Bernasconi G, Antonovics J, Biere A, *et al.* (2009) Silene as a model system in ecology and evolution.  
 538 *Heredity (Edinb)* **103**, 5-14.
- 539 Brothers AN, Delph LF (2010) Haldane's rule is extended to plants with sex chromosomes. *Evolution*  
 540 **64**, 3643-3648.
- 541 Charlesworth B, Campos JL, Jackson BC (2018) Faster-X evolution: Theory and evidence from  
 542 *Drosophila*. *Mol Ecol in press*.
- 543 Charlesworth B, Charlesworth D (2000) The degeneration of Y chromosomes. *Philos Trans R Soc*  
 544 *Lond B Biol Sci* **355**, 1563-1572.
- 545 Charlesworth B, Coyne JA, Barton NH (1987) The relative rates of evolution of sex-chromosomes and  
 546 autosomes. *Am Naturalist* **130**, 113-146.
- 547 Charlesworth D (2008) Sex chromosome origins and evolution. In: *Evolutionary Genomics and*  
 548 *Proteomics* (eds. Pagel M, Pomiankowski A), pp. 207-240. Sinauer Associates, Sunderland.
- 549 Charlesworth D (2015) Plant contributions to our understanding of sex chromosome evolution. *New*  
 550 *Phytol* **208**, 52-65.
- 551 Chibalina MV, Filatov DA (2011) Plant Y chromosome degeneration is retarded by haploid purifying  
 552 selection. *Curr Biol* **21**, 1475-1479.
- 553 Ciupercescu DD, Veuskens J, Mouras A, *et al.* (1990) Karyotyping *Melandrium album*, a dioecious  
 554 plant with heteromorphic sex chromosomes. *Genome* **33**, 556-562.
- 555 Coyne JA, Orr HA (1989) The two rules of speciation. In: *Speciation and its Consequences* (eds. Otte  
 556 D, Endler J), pp. 180-207. Sinauer Associates.
- 557 Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates, Sunderland, Mass.
- 558 Delph LF, Demuth JP (2016) Haldane's rule: genetic bases and their empirical support. *J Hered* **107**,  
 559 383-391.
- 560 Demuth JP, Flanagan RJ, Delph LF (2014) Genetic architecture of isolation between two species of  
 561 *Silene* with sex chromosomes and Haldane's rule. *Evolution* **68**, 332-342.
- 562 Dufresnes C, Majtyka T, Baird SJ, *et al.* (2016) Empirical evidence for large X-effects in animals with  
 563 undifferentiated sex chromosomes. *Sci Rep* **6**, 21029.
- 564 Ellegren H, Smeds L, Burri R, *et al.* (2012) The genomic landscape of species divergence in *Ficedula*  
 565 flycatchers. *Nature* **491**, 756-760.
- 566 Favre A, Widmer A, Karrenberg S (2017) Differential adaptation drives ecological speciation in  
 567 champions (*Silene*): evidence from a multi-site transplant experiment. *New Phytol* **213**, 1487-  
 568 1499.
- 569 Frank SA (1991) Divergence of meiotic drive-suppression systems as an explanation for sex-biased  
 570 hybrid sterility and inviability. *Evolution* **45**, 262-267.
- 571 Haldane JBS (1922) Sex-ratio and unisexual sterility in hybrid animals. *J. Genet.* **12**, 101-109.
- 572 Honys D, Twell D (2004) Transcriptome analysis of haploid male gametophyte development in  
 573 *Arabidopsis*. *Genome Biol* **5**, R85.
- 574 Howell EC, Armstrong SJ, Filatov DA (2009) Evolution of neo-sex chromosomes in *Silene diclinis*.  
 575 *Genetics* **182**, 1109-1115.
- 576 Hu XS, Filatov DA (2016) The large-X effect in plants: increased species divergence and reduced  
 577 gene flow on the *Silene* X-chromosome. *Molecular Ecology* **25**, 2609-2619.
- 578 Hurst LD, Pomiankowski A (1991) Causes of sex ratio bias may account for unisexual sterility in  
 579 hybrids: a new explanation of Haldane's rule and related phenomena. *Genetics* **128**, 841-858.
- 580 Jiggins CD, Linares M, Naisbit RE, *et al.* (2001) Sex-linked hybrid sterility in a butterfly. *Evolution* **55**,  
 581 1631-1638.
- 582 Krasovec M, Chester M, Ridout K, Filatov DA (2018) The mutation rate and the age of the sex  
 583 chromosomes in *Silene latifolia*. *Curr Biol in press*.
- 584 Larson EL, Keeble S, Vanderpool D, Dean MD, Good JM (2017) The composite regulatory basis of  
 585 the large X-effect in mouse speciation. *Mol Biol Evol* **34**, 282-295.
- 586 Laurie CC (1997) The weaker sex is heterogametic: 75 years of Haldane's rule. *Genetics* **147**, 937-  
 587 951.
- 588 Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a  
 589 reference genome. *BMC Bioinformatics* **12**, 323.
- 590 Lima TG (2014) Higher levels of sex chromosome heteromorphism are associated with markedly  
 591 stronger reproductive isolation. *Nat Commun* **5**, 4743.
- 592 Mank JE, Ellegren H (2009) All dosage compensation is local: Gene-by-gene regulation of sex-biased  
 593 expression on the chicken Z chromosome. *Heredity* **102**, 312-320.
- 594 Mank JE, Vicoso B, Berlin S, Charlesworth B (2010) Effective population size and the Faster-X effect:  
 595 empirical results and their interpretation. *Evolution* **64**, 663-674.

- 596 Masly JP, Presgraves DC (2007) High-resolution genome-wide dissection of the two rules of  
597 speciation in *Drosophila*. *PLoS Biol* **5**, e243.
- 598 Meisel RP, Connallon T (2013) The faster-X effect: integrating theory and data. *Trends Genet* **29**,  
599 537-544.
- 600 Meisel RP, Malone JH, Clark AG (2012) Faster-X evolution of gene expression in *Drosophila*. *PLoS*  
601 *Genet* **8**, e1003013.
- 602 Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian  
603 transcriptomes by RNA-Seq. *Nat Methods* **5**, 621-628.
- 604 Moyle LC, Muir CD, Han MV, Hahn MW (2010) The contribution of gene movement to the "two rules  
605 of speciation". *Evolution* **64**, 1541-1557.
- 606 Muyle A, Zemp N, Deschamps C, *et al.* (2012) Rapid de novo evolution of X chromosome dosage  
607 compensation in *Silene latifolia*, a plant with young sex chromosomes. *PLoS Biol* **10**,  
608 e1001308.
- 609 Orr HA (1997) Haldane's rule. *Annu. Rev. Ecol. Syst.* **28**, 195-218.
- 610 Orr HA, Turelli M (1996) Dominance and Haldane's rule. *Genetics* **143**, 613-616.
- 611 Papadopulos AS, Chester M, Ridout K, Filatov DA (2015) Rapid Y degeneration and dosage  
612 compensation in plant sex chromosomes. *Proc Natl Acad Sci U S A* **112**, 13021-13026.
- 613 Payseur BA, Krenz JG, Nachman MW (2004) Differential patterns of introgression across the X  
614 chromosome in a hybrid zone between two species of house mice. *Evolution* **58**, 2064-2078.
- 615 Presgraves DC, Orr HA (1998) Haldane's rule in taxa lacking a hemizygous X. *Science* **282**, 952-954.
- 616 Tao Y, Hartl DL (2003) Genetic dissection of hybrid incompatibilities between *Drosophila simulans*  
617 and *D. mauritiana*. III. Heterogeneous accumulation of hybrid incompatibilities, degree of  
618 dominance, and implications for Haldane's rule. *Evolution* **57**, 2580-2598.
- 619 Turelli M, Begun DJ (1997) Haldane's rule and X-chromosome size in *Drosophila*. *Genetics* **147**,  
620 1799-1815.
- 621 Turelli M, Moyle LC (2007) Asymmetric postmating isolation: Darwin's corollary to Haldane's rule.  
622 *Genetics* **176**, 1059-1088.
- 623 Turelli M, Orr HA (1995) The dominance theory of Haldane's rule. *Genetics* **140**, 389-402.
- 624 Turner LM, White MA, Tautz D, Payseur BA (2014) Genomic networks of hybrid sterility. *PLoS Genet*  
625 **10**, e1004162.
- 626 Vicoso B, Charlesworth B (2006) Evolution on the X chromosome: unusual patterns and processes.  
627 *Nat Rev Genet* **7**, 645-653.
- 628 Vicoso B, Charlesworth B (2009) Effective population size and the faster-X effect: an extended model.  
629 *Evolution* **63**, 2413-2426.
- 630 Wu C-I, Davis AW (1993) Evolution of postmating reproductive isolation: the composite nature of  
631 Haldane's rule and its genetic bases. *American Naturalist* **142**, 187-212.
- 632 Wu CI, Johnson NA, Palopoli MF (1996) Haldane's rule and its legacy: Why are there so many sterile  
633 males? *Trends in Ecology & Evolution* **11**, 281-284.
- 634 Zemp N, Tavares R, Muyle A, *et al.* (2016) Evolution of sex-biased gene expression in a dioecious  
635 plant. *Nat Plants* **2**, 16168.
- 636

637

638

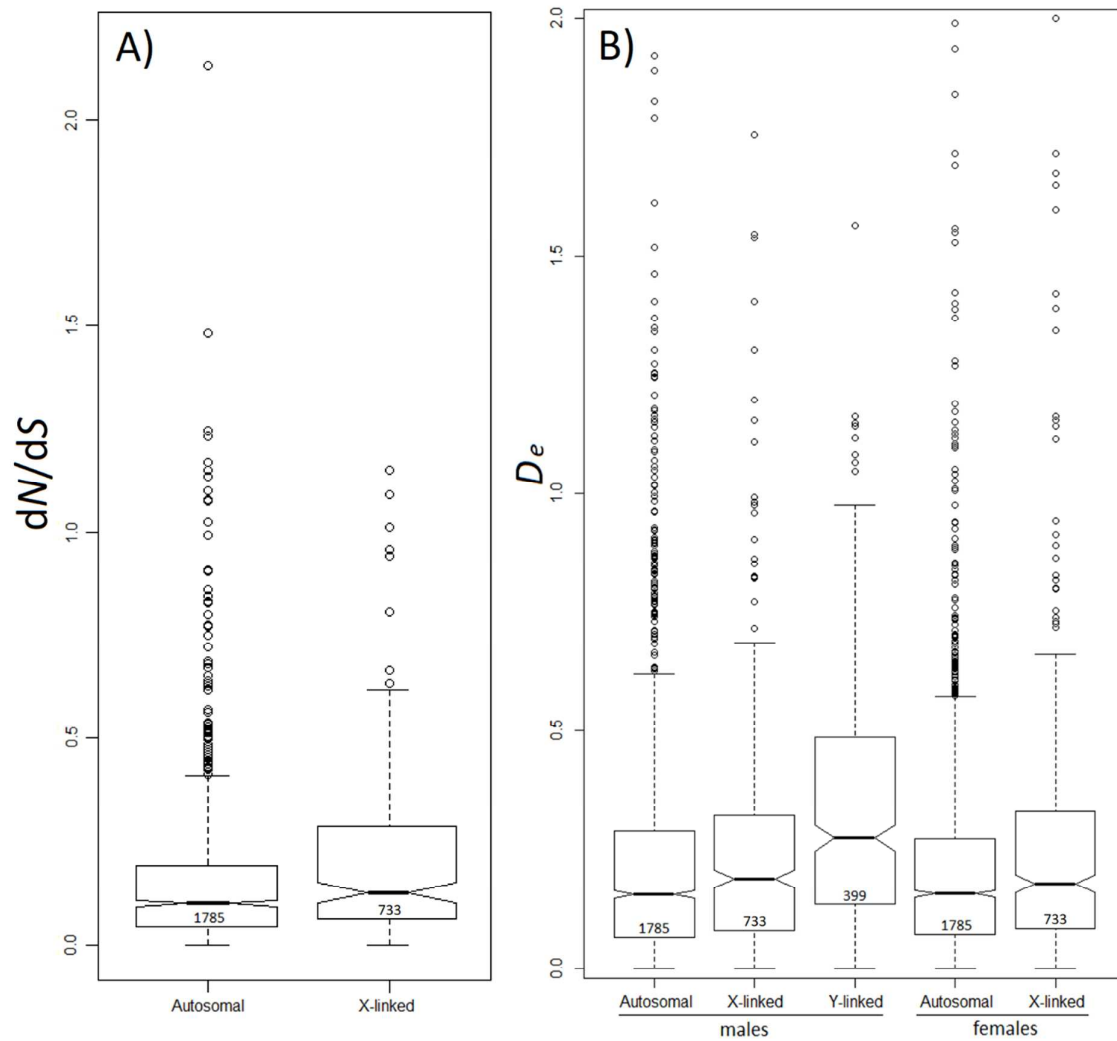


639

640 Figure 1. Loss of gene expression on the Y-chromosome. The plot shows  $\log_2$ -transformed ratios  
 641 of relative Y/X gene expression of individual sex-linked genes in *S. latifolia* (X-axis) and *S. dioica* (Y-  
 642 axis) males. To avoid infinity and division by zero errors for non-expressed genes, a small number  
 643 (0.0001) was added to all expression values. The ellipses were drawn by hand to highlight different  
 644 parts of the plot.

645

646

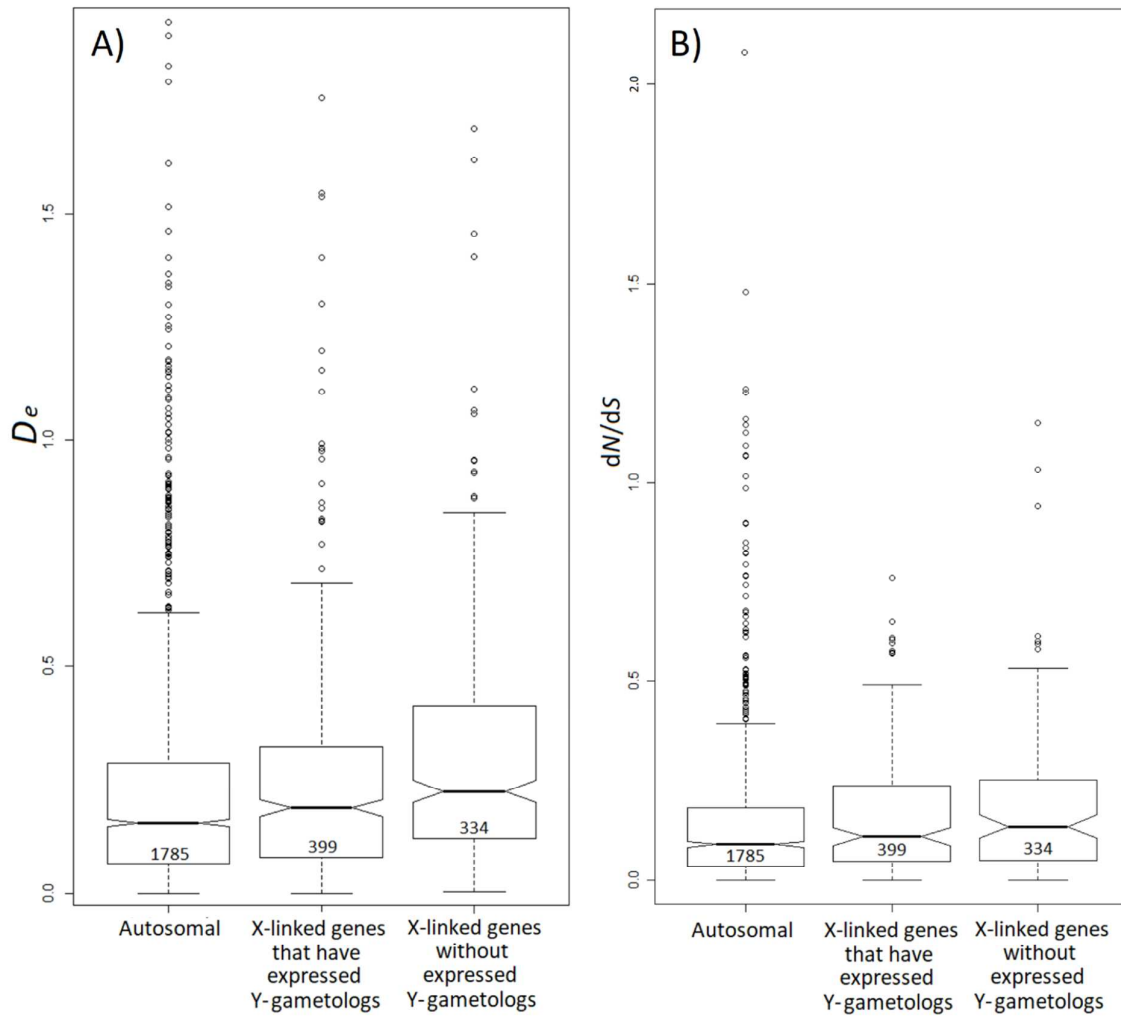


647

648 Figure 2. Non-synonymous ( $dN$ ) to synonymous ( $dS$ ) nucleotide substitution rate ratio (A) and gene  
 649 expression divergence (B) between *S. latifolia* and *S. dioica* for sex-linked and autosomal genes. The  
 650 numbers of genes analysed are shown inside the boxes.

651

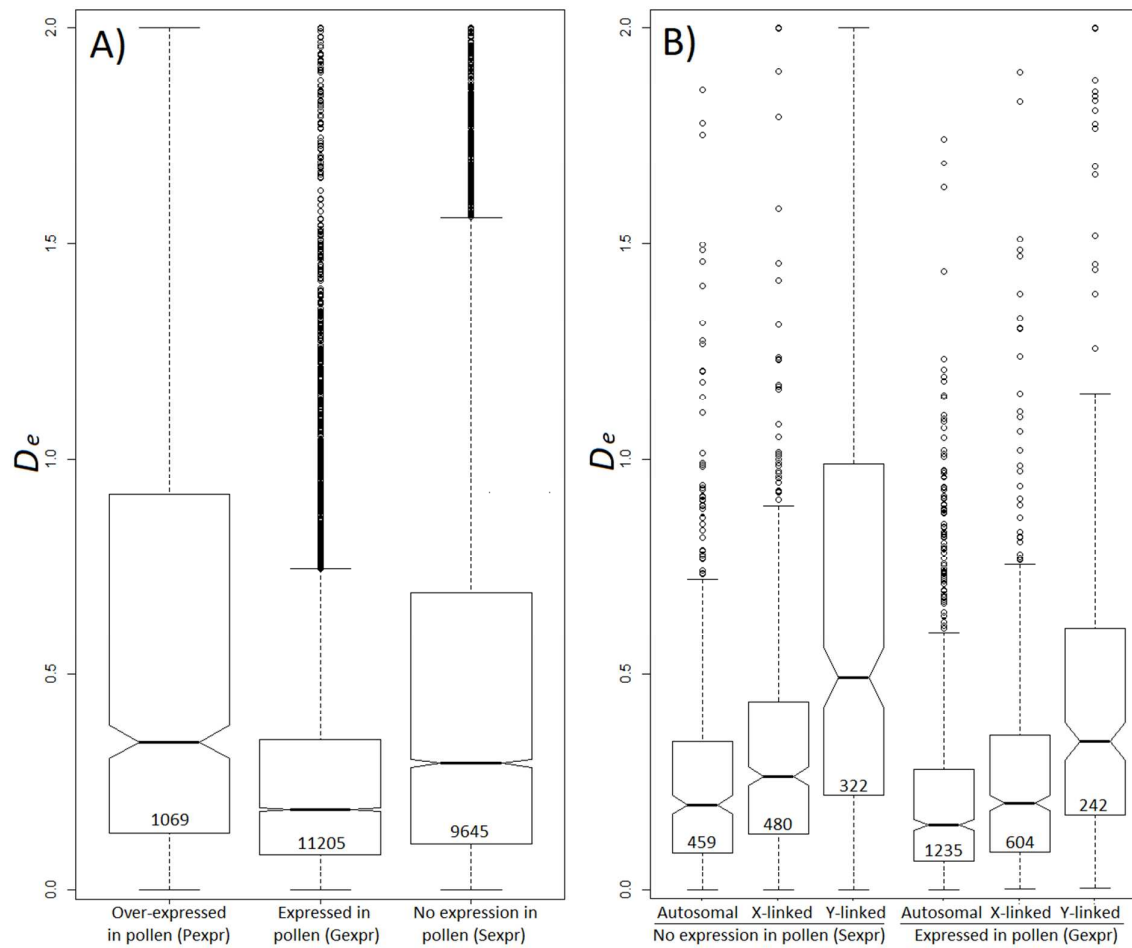
652



653

654 Figure 3. Faster-X evolution for gene expression divergence in males (A) and sequence divergence  
 655 ( $dN/dS$ , panel B) between *S. latifolia* and *S. dioica*. X-linked genes are split into two categories for X-  
 656 linked genes with expressed ( $X_{Y_{\text{expr}}}$ ) and non-expressed ( $X_{\text{no}Y_{\text{expr}}}$ ) Y-linked gametologs. The numbers  
 657 of genes analysed are shown inside the boxes.

658



659

660

661 Figure 4. Gene expression divergence between *S. latifolia* and *S. dioica* males for genes  
 662 predominantly expressed in the sporophyte ( $S_{expr}$ ) and the gametophyte ( $G_{expr}$  and  $P_{expr}$ ). A) All genes  
 663 regardless of linkage; B) A comparison of autosomal and sex-linked genes. The numbers of genes  
 664 analysed are shown inside the boxes.

664

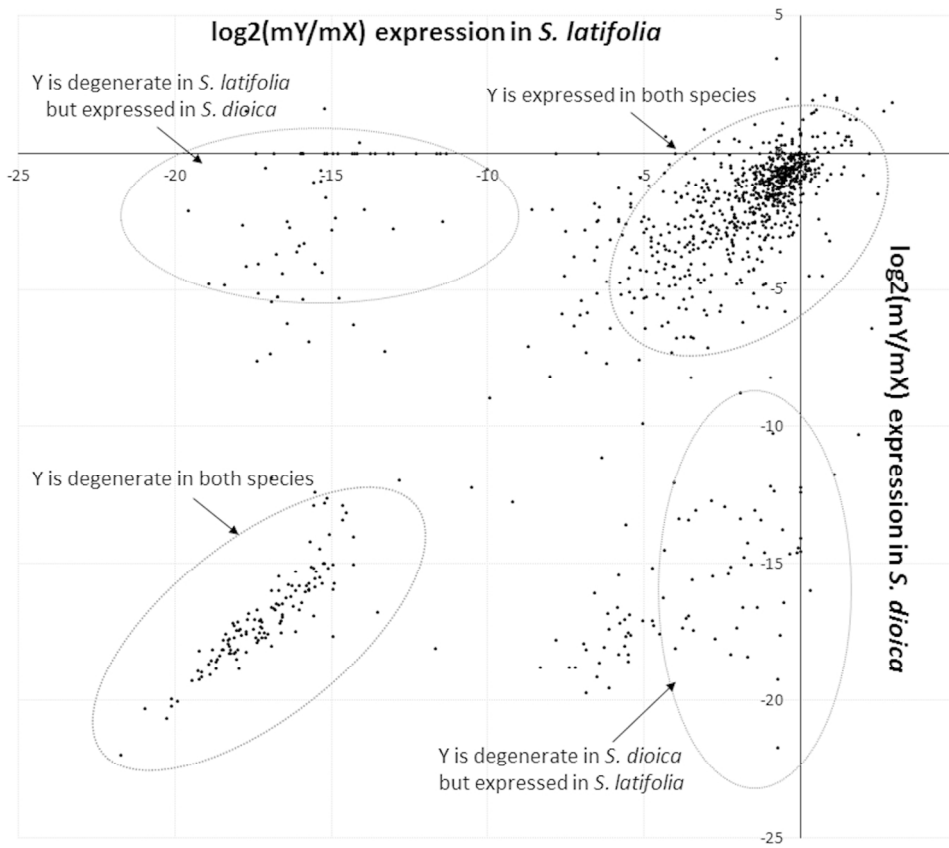


Figure 1

119x103mm (300 x 300 DPI)

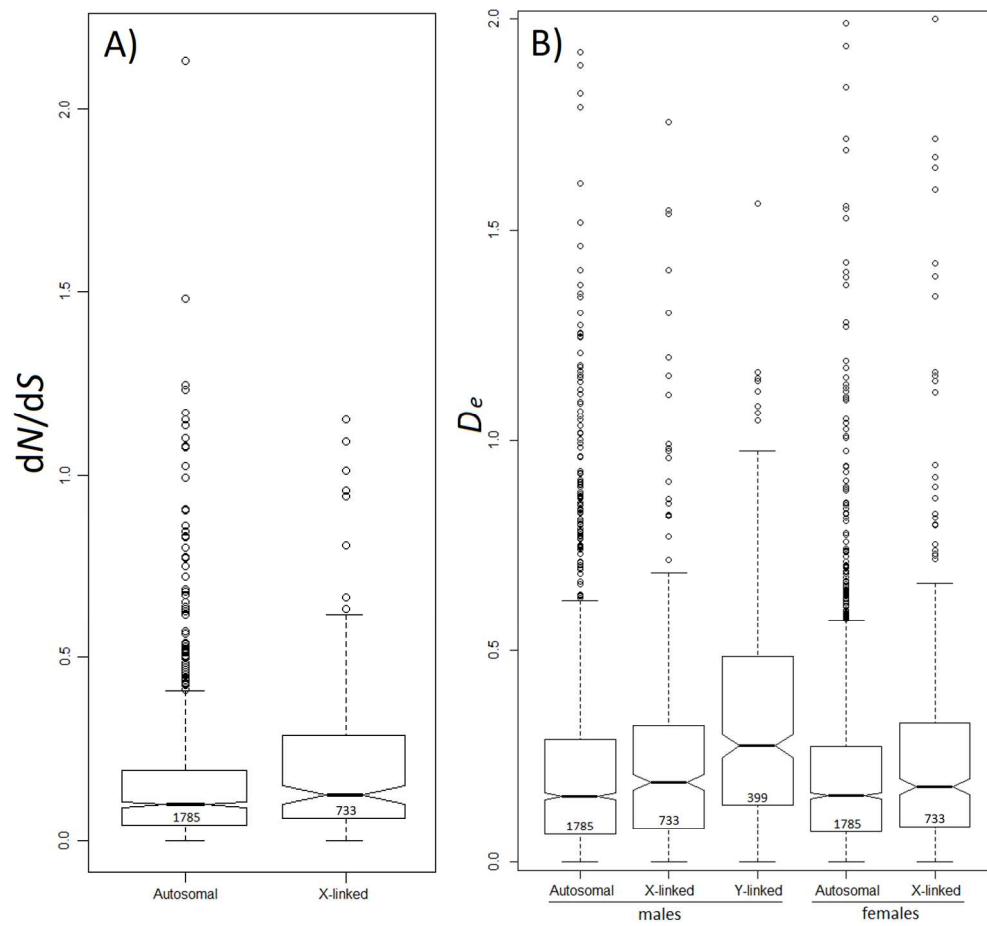


Figure 2

146x136mm (300 x 300 DPI)

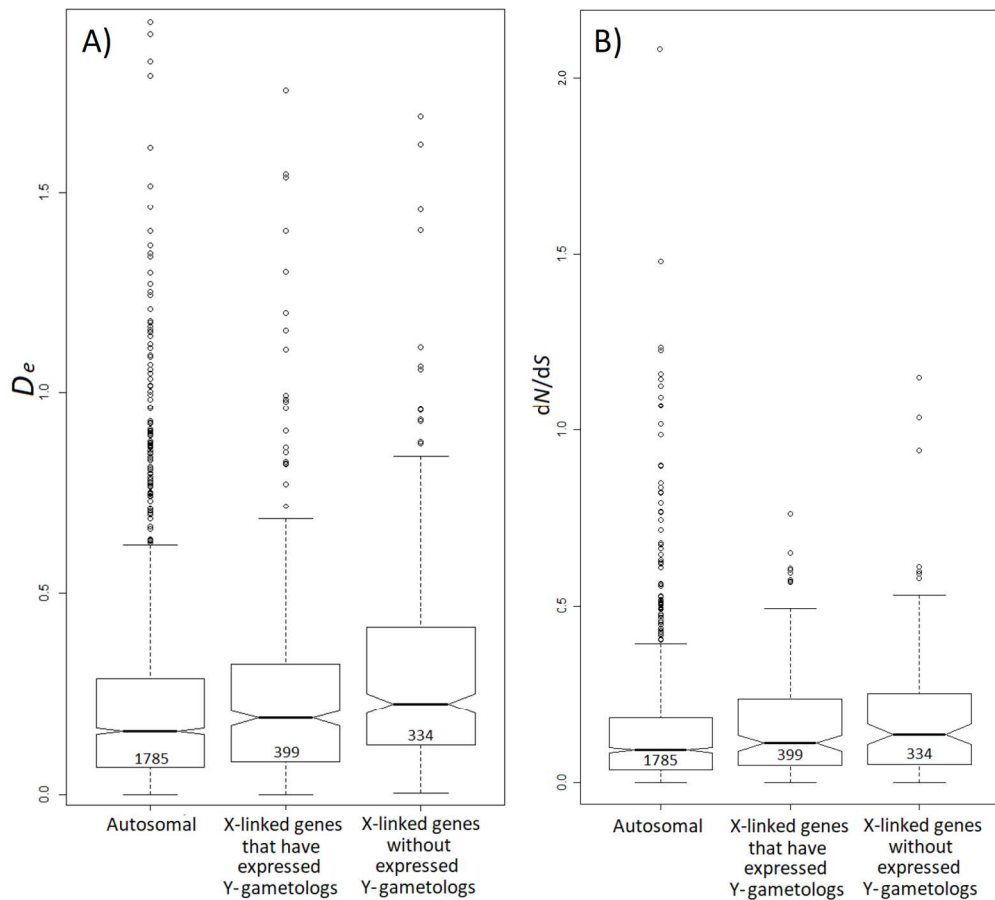


Figure 3

167x151mm (300 x 300 DPI)

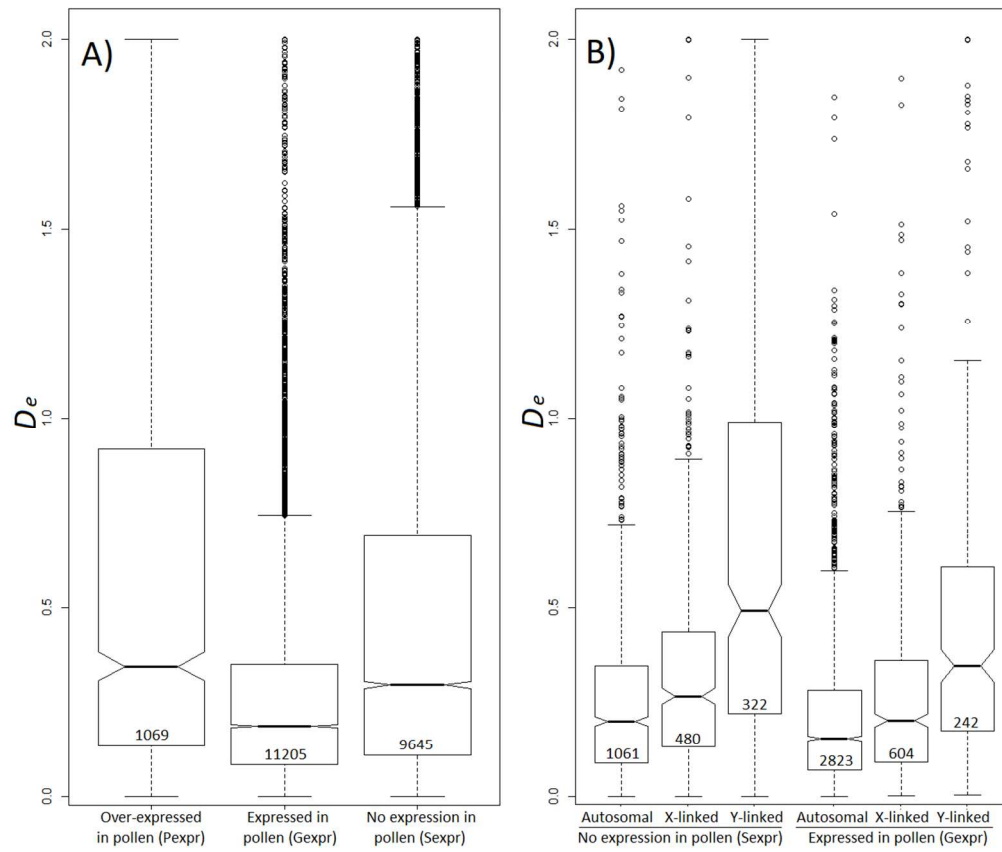


Figure 4

108x91mm (300 x 300 DPI)

1 **The “two rules of speciation” in species with young sex chromosomes**

2

3 Dmitry A. Filatov

4

5 Email for correspondence: [Dmitry.Filatov@plants.ox.ac.uk](mailto:Dmitry.Filatov@plants.ox.ac.uk)

6

7 Address:

8 Department of Plant Sciences,

9 University of Oxford,

10 South Parks Road,

11 Oxford OX1 3RB

12 United Kingdom

13

14 Running title: Y-degeneration and faster-X in *Silene* speciation

15

16 Keywords: Haldane’s rule, large-X effect, faster-X effect, gene expression, Y-chromosome  
17 degeneration, speciation, plant sex chromosomes, gametophyte, haploid expression.

18

19

## 20 **Abstract**

21 The two “rules of speciation”, Haldane’s rule (HR) and the large-X effect (LXE), are thought to be  
22 caused by recessive species incompatibilities exposed in the phenotype due to the hemizyosity of X-  
23 linked genes in the heterogametic sex. Thus, the reports of HR and the LXE in species with recently  
24 evolved non- or partially-degenerate Y-chromosomes, such as *Silene latifolia* and its relatives, were  
25 surprising. Here I argue that rapid species-specific degeneration of Y-linked genes and associated  
26 adjustment of expression of X-linked gametologs (dosage compensation) may lead to rapid evolution  
27 of sex-linked species incompatibilities. This process is likely to be too slow in species with old  
28 degenerate Y-chromosomes (e.g. in mammals), but Y-degeneration in species with young gene-rich  
29 sex chromosomes may be fast enough to play a significant role in speciation. To illustrate this point I  
30 report the analysis of Y-degeneration and the associated evolution of gene expression on the X-  
31 chromosome of *Silene latifolia* and *Silene dioica*, a close relative that shares the same recently  
32 evolved sex chromosomes. Despite the recent ( $\leq 1$ MY) divergence of the two species, ~7% of Y-linked  
33 genes have undergone degeneration in one but not the other species. This species-specific  
34 degeneration appears to drive faster expression divergence of X-linked genes, which may account for  
35 HR and the LXE reported for these species. Furthermore, I suggest that “exposure” of autosomal or  
36 sex-linked recessive species incompatibilities in the haploid plant gametophyte may mimic the  
37 presence of HR in plants. Both haploid expression and species-specific Y-degeneration need to  
38 receive more attention if we are to understand the role of these processes in speciation.

39

## 40 **Introduction**

41 Many closely related species are known to form hybrids in nature, allowing them to exchange genes,  
42 which may slow down or prevent speciation and species divergence. The evolution of reproductive  
43 barriers is crucial for the speciation process to proceed and there is substantial evidence that sex  
44 chromosomes play a central role in the evolution of reproductive incompatibilities between incipient  
45 species (Larson *et al.* 2017; Laurie 1997; Masly & Presgraves 2007; Presgraves & Orr 1998; Turner  
46 *et al.* 2014). The importance of sex chromosomes in speciation is reflected in the “two rules of  
47 speciation”: Haldane’s rule and the large-X effect (Coyne & Orr 1989). Haldane’s rule (HR) states that  
48 the heterogametic sex is more likely to exhibit inviability or infertility in inter-specific hybrids, compared  
49 to the homogametic sex (Haldane 1922; Orr 1997). The large-X effect (LXE) posits that the X-

50 chromosome plays a disproportionately large role in hybrid dysfunction, compared to the autosomes  
51 (Coyne & Orr 2004; Jiggins *et al.* 2001; Turelli & Moyle 2007). Despite the striking variety in sex  
52 determination systems and sex chromosomes (Bachtrog *et al.* 2014) these rules of speciation are  
53 surprisingly universal. Both preferential hybrid inviability in the heterogametic sex and reduced  
54 interspecific gene flow of X-linked compared to autosomal genes – a convenient proxy for the LXE –  
55 have been described across plant and animal systems (Dufresnes *et al.* 2016; Ellegren *et al.* 2012;  
56 Hu & Filatov 2016; Payseur *et al.* 2004). This suggests the same, general mechanisms may be  
57 responsible for the special role of sex chromosomes in speciation across animal and plant groups.

58 The underlying causes of the two rules of speciation are not certain, but several hypotheses were  
59 proposed to account for HR and the LXE (Delph & Demuth 2016; Laurie 1997; Orr 1997). It is thought  
60 that, at least partly, both rules arise from recessive hybrid incompatibility alleles exhibiting full effects  
61 in the heterogametic sex (the dominance theory, (Coyne & Orr 2004; Turelli & Orr 1995)). The other  
62 possible causes include X-chromosome misregulation in hybrids (Larson *et al.* 2017; Masly &  
63 Presgraves 2007), greater density of male sterility loci on the X compared to autosomes (Masly &  
64 Presgraves 2007), meiotic drive on the sex chromosomes (Frank 1991; Hurst & Pomiankowski 1991;  
65 Tao & Hartl 2003) and faster evolution of X-linked genes (“faster-X” theory), which is predicted to  
66 arise if adaptive mutations are partially recessive (Charlesworth *et al.* 1987; Mank *et al.* 2010; Vicoso  
67 & Charlesworth 2006, 2009). It was also proposed that commonly observed hybrid male sterility but  
68 not female sterility, is due to faster evolution of genes involved in spermatogenesis compared to  
69 oogenesis-related genes and stronger sexual selection on males than females (“faster males” theory  
70 (Wu & Davis 1993)), which can help explain HR in species with male heterogamety (Laurie 1997; Orr  
71 1997; Wu *et al.* 1996). Extensive experimental work has provided support to the dominance and the  
72 faster males theories, which are now regarded as the most plausible explanations to the “two rules of  
73 speciation” (Laurie 1997; Masly & Presgraves 2007; Presgraves & Orr 1998).

74 Many of the aforementioned hypotheses assume that the Y-chromosome is already degenerate,  
75 resulting in the exposure of X-linked recessive mutations in hemizygous males. However, sex  
76 chromosomes are highly evolutionary labile, with many species (beyond well-studied mammals and  
77 *Drosophila*) exhibiting non-degenerate or partially degenerate Y-chromosomes (Bachtrog *et al.* 2014).  
78 The species with non-differentiated (homomorphic) sex chromosomes tend to show weaker  
79 reproductive isolation between closely related species, compared to species with heteromorphic sex

80 chromosomes (Lima 2014; Presgraves & Orr 1998). Nevertheless, the reports that the two “rules of  
81 speciation” also apply to species with non- or partly degenerate Y-chromosomes (Brothers & Delph  
82 2010; Dufresnes *et al.* 2016; Hu & Filatov 2016; Presgraves & Orr 1998) raise doubt that  
83 hemizyosity of the X-linked genes is the universal cause underpinning HR and the LXE.

84 Here I propose and investigate an additional potential cause for the special role that sex  
85 chromosomes play in speciation – species-specific Y-degeneration that drives divergent evolution of  
86 compensatory mechanisms, such as species-specific dosage compensation (ssDoC) where an X-  
87 linked gametolog is upregulated due to degeneration of the Y-linked gametolog. Following Y  
88 degeneration and evolution of dosage compensation, interspecific hybridisation with another species  
89 where the Y-copy of a gene is still functional, would result in a range of under- and over-  
90 compensation in male hybrids. For example, combining a degenerate Y-linked gene with non-  
91 compensating X-linked gametologs in hybrids would result in only half the normal gene dose, while a  
92 hybrid with functional Y-linked copy and dosage compensated (up-regulated) X-linked gametolog  
93 would have  $3/2$  the normal gene expression for the sex linked gene. As such, the modest differences  
94 in gene expression that are sufficient to drive the evolution of a dosage compensation system may  
95 also be sufficient to play a significant role in hybrid dysfunction. Furthermore, Y-degeneration may  
96 lead to compensatory evolution in non-homologous sex-linked or autosomal genes involved in the  
97 same biochemical pathway or gene regulatory network to adjust for lower dosage at the sex-linked  
98 gene with a degenerate Y-linked gametolog. Interspecific hybridisation may combine a functional Y-  
99 linked gene from one species with the genes that co-evolved with the degenerate Y of another  
100 species, leading to expression levels which are detrimental for hybrid viability or fertility.

101 This model is similar, but not identical to the model of Moyle *et al.* (Moyle *et al.* 2010) describing the  
102 contribution of gene movement between the chromosomes to the LXE and HR. In particular, both  
103 models predict transgressive gene expression in the heterogametic sex in hybrids. However, the  
104 ssDoC model described here implies the evolution of gene expression, while the gene movement  
105 model (Moyle *et al.* 2010) only requires that genes be moved between the X and the autosomes but  
106 does not involve any evolution of gene expression at the particular genes. This difference determines  
107 the type of effects contributing to hybrid dysfunction in the two models: unlike the gene movement  
108 model (Moyle *et al.* 2010), the ssDoC model implies that the fitness effects in hybrids can be

109 considered under-dominant, as combining a functional Y-linked copy and a dosage compensated X-  
110 linked copy is deleterious. The implications of this process for speciation have not been studied.

111 Potentially, rapid gene loss from the Y-chromosome and the associated evolution of dosage  
112 compensation on the X-chromosome could cause the rapid accumulation of hybrid incompatibilities  
113 between closely related species or sub-species with recently evolved gene-rich sex chromosomes.  
114 The rate of genetic degeneration in the non-recombining regions of sex-specific Y(or W)-  
115 chromosomes depends on the number of functional genes that are linked together, with Y-  
116 degeneration proceeding at a slower rate as the number of Y-linked genes decreases (Bachtrog  
117 2008; Charlesworth 2008). This results in the rapid loss of Y-linked genes at the early stages of sex  
118 chromosome evolution, with the rate of gene loss slowing down on older Y(or W)-chromosomes.  
119 Thus, Y-degeneration is likely to be too slow to create species incompatibilities in species with old  
120 degenerate Y-chromosomes (e.g. in mammals), but rapid Y-degeneration of young gene-rich sex  
121 chromosomes may be fast enough to play a significant role in speciation. To illustrate this point I  
122 report an analysis of Y-degeneration and the associated evolution of gene expression on the X-  
123 chromosome of *Silene latifolia* and a close relative, *Silene dioica*, which share recently evolved sex  
124 chromosomes.

125 Both *S. latifolia* and *S. dioica* are common throughout Europe, with *S. latifolia* inhabiting open  
126 fields and road margins, and *S. dioica* tending to be found in more shady and moist habitats, such as  
127 forests. Despite the difference in preferred habitat, which likely plays the primary role in isolation of  
128 these species (Favre *et al.* 2017), *S. latifolia* and *S. dioica* often form hybrid swarms in places of co-  
129 occurrence. These species have identical karyotypes with clearly distinguishable X and Y-  
130 chromosomes (Armstrong & Filatov 2008; Ciupercescu *et al.* 1990). Due to the relatively recent  
131 (<10Myr) origin of dioecy and sex chromosomes in the ancestor of these species, *S. latifolia* is being  
132 used to study the early stages of sex chromosome evolution (Bernasconi *et al.* 2009; Charlesworth  
133 2015). Several recent studies demonstrated that most *S. latifolia* X-linked genes still appear to have  
134 functional Y-linked gametologs, though some Y-degeneration is apparent (Bergero *et al.* 2015;  
135 Chibalina & Filatov 2011; Papadopulos *et al.* 2015). The loss of at least some *S. latifolia* Y-linked  
136 genes appears to be compensated by a higher expression of their X-linked gametologs (dosage  
137 compensation, (Muyle *et al.* 2012; Papadopulos *et al.* 2015)). Below I report that rapid, on-going Y-  
138 chromosome degeneration and an evolving dosage compensation system are making *Silene* sex

139 chromosomes diverge faster than the autosomes. I propose that a higher rate of sex chromosome  
140 divergence may be an important contributor of HR and the LXE reported for these species (Brothers &  
141 Delph 2010; Hu & Filatov 2016).

142

## 143 **Materials and Methods**

### 144 *Plant material*

145 Six *S. latifolia* and six *S. dioica* plants (three males and three females of each species) were grown  
146 in the glasshouse (20°C and 15h lighting) from seed collected in the wild. The females of both species  
147 have already been used in a previous study assessing the LXE in *S. latifolia* and *S. dioica* (Hu &  
148 Filatov 2016). In addition, one male plant from another closely related dioecious species, *Silene*  
149 *diclinis* (Hu & Filatov 2016) was used as an outgroup in some of the analyses. While *S. latifolia* and *S.*  
150 *dioica* are very common all over Europe, *S. diclinis* is a rare endemic narrowly restricted to Xativa in  
151 Valencia (Spain) and material for this species was very limited, hence only one accession was  
152 available for analysis. It is worth noting that although *S. diclinis*, is closely related to *S. latifolia* and *S.*  
153 *dioica*, its sex chromosomes were rearranged, resulting in the evolution of neo-sex-chromosomes in  
154 that species (Howell *et al.* 2009), which may have affected expression at some genes.

155

### 156 *RNA extraction and sequencing*

157 RNA was extracted from actively growing shoots and flower buds from all of the plants, as  
158 described previously (Hu & Filatov 2016; Papadopoulos *et al.* 2015). Total RNA from plant tissue was  
159 extracted using a Qiagen RNeasy Plant Mini Kit, including the optional on-column DNase digestion.  
160 Isolation of mRNA, cDNA synthesis and high-throughput sequencing were conducted according to the  
161 standard Illumina RNA-Seq procedure at the Oxford Genomics Centre of the Wellcome Trust Center  
162 for Human Genetics (WTCHG, Oxford). High-throughput sequencing for each individual was  
163 conducted at WTCHG using an Illumina HiSeq2000 instrument with 100 base, paired-end reads. All  
164 sequence reads were submitted to SRA database under project number PRJNA453413.

165

### 166 *Expression analyses*

167 The reference transcriptome used for RNA-seq read mapping was taken from a previous study  
168 (Papadopoulos *et al.* 2015). A significant advantage of that reference transcriptome is that genomic

169 sequences (along with RNA-seq data) were used to reconstruct Y-linked genes (Papadopulos *et al.*  
170 2015), while the transcriptomes in other studies were based entirely on RNA-seq data (Bergero &  
171 Charlesworth 2011; Chibalina & Filatov 2011; Muyle *et al.* 2012; Zemp *et al.* 2016), resulting in under-  
172 representation of weakly expressed Y-linked genes. Furthermore, a large number of genes (2,114) in  
173 that transcriptome were genetically mapped to 12 chromosomes on *S. latifolia* (Papadopulos *et al.*  
174 2015), while other studies (Bergero *et al.* 2015; Chibalina & Filatov 2011; Muyle *et al.* 2012; Zemp *et*  
175 *al.* 2016) have only classified genes into 'X-linked', non-X-linked, and 'unknown' bins, with varying  
176 degrees of uncertainty.

177 RNA-seq reads were mapped to the reference transcriptome and gene expression was measured  
178 using RSEM (Li & Dewey 2011) with default parameters. "Fragments per kilobase per million reads"  
179 (FPKM (Mortazavi *et al.* 2008)) values from RSEM were used in all the analyses of expression. The  
180 accuracy of this approach in distinguishing homologous X- and Y-linked alleles in *S. latifolia* was  
181 demonstrated in the study by Papadopulos *et al.* (Papadopulos *et al.* 2015). In particular, the  
182 "expression" of Y-linked alleles in females was zero, as expected if RSEM accurately distinguishes  
183 between X- and Y-linked gametologs (see Suppl. methods page 5 in (Papadopulos *et al.* 2015)). The  
184 mapping of different *Silene* species to the same transcriptome is justified by the low sequence  
185 divergence between these species (silent site divergence ~1.5%, which is similar to intra-specific  
186 polymorphism in *S. latifolia* [ $\pi_s \sim 1.5\%$ ] (Hu & Filatov 2016)), resulting in similar proportions (~80%) of  
187 *S. latifolia*, *S. dioica* and *S. diclinis* RNA-seq reads successfully mapping to the reference  
188 transcriptome. Importantly, the three species had similar proportions of reads mapping to sex-linked  
189 and autosomal genes, illustrating that mapping to a heterospecific transcriptome did not differentially  
190 affect sex-linked genes.

191 Per-gene expression divergence between the species ( $D_e$ ) was calculated separately for males  
192 and females.  $D_e$  was calculated as the difference between medians for expression in the two species,  
193 normalised by the average of the two medians (Meisel *et al.* 2012). As the aim of the analyses in the  
194 current study was to quantify the overall divergence, rather than the direction of change in expression,  
195 the absolute value of  $D_e$  was used throughout this paper.

196 Statistical analyses ( $\chi^2$ , Wilcoxon tests and box plots) of gene expression were done in R, except  
197 figure 1, which was made in Excel (Microsoft).

198

## 199 **Results**

### 200 *The extent of Y-degeneration in S. latifolia and S. dioica*

201 To test the extent of genetic degeneration in *S. latifolia* and *S. dioica*, Y-linked gene expression of  
202 homologous X- and Y-linked genes in males (mX and mY, respectively) were analysed relative to the  
203 expression of X-linked genes in females (fXX). The comparison with female expression is used to  
204 avoid confounding Y-degeneration and dosage compensation, both of which would affect the  
205 comparison between mY and mX. The expression analysis of sex-linked genes reveals extensive  
206 genetic degeneration on the Y-chromosome of both *S. latifolia* and *S. dioica*. In particular, out of 982  
207 genes for which both X- and Y-linked gametologs are available from the previous work (Chibalina &  
208 Filatov 2011; Papadopulos *et al.* 2015), over 1/3<sup>rd</sup> of the Y-linked genes show more than 10-fold  
209 reduction in expression compared to X-linked gametologs (404, 419 and 360 genes in *S. latifolia*, *S.*  
210 *dioica* and both species, respectively). This result is consistent with the previous reports of Y-  
211 degeneration in *S. latifolia* (Bergero *et al.* 2015; Papadopulos *et al.* 2015) and extends the analysis to  
212 *S. dioica* where Y-degeneration has not been analysed previously.

213 The analysis above was based on genes with intact X- and Y-linked gametologs (including genes  
214 where the Y-copy is not expressed, as long as it is detectable in the genomic sequence (Papadopulos  
215 *et al.* 2015)) and does not take into account any genes without a detectable Y-linked copy. There are  
216 246 such X-only genes among the *S. latifolia* sex-linked genes detected previously (Chibalina &  
217 Filatov 2011; Papadopulos *et al.* 2015). Many of these genes may have lost the Y-linked gametolog  
218 because of on-going Y-chromosome degeneration, though some may have been translocated to the  
219 X-chromosome from an autosome and never had a Y-linked gametolog. Taking these X-only genes  
220 into account brings the proportion of X-linked hemizygous genes to ~50%.

221 If *S. latifolia* and *S. dioica* Y-chromosomes have continued to degenerate, one would expect to see  
222 Y-linked genes that have been lost since these two species diverged. Such recently lost genes should  
223 be species-specific, that is, be actively expressed in one species and non-functional in the other  
224 dioecious species. Indeed, there is a considerable number of Y-linked genes with species-specific or  
225 nearly species-specific expression (Fig. 1). In particular, in *S. latifolia* and *S. dioica* 46 and 42 Y-linked  
226 genes, respectively, show >10-fold reduction in gene expression compared to expression of X-linked  
227 gametologs in females (fXX), while expression of the Y-linked copy in the other species is >30% fXX.  
228 As all these Y-linked genes are actively expressed in the closely related dioecious outgroup *S. diclinis*

229 (Fig. S1), they likely represent species-specific loss of Y-linked gene expression in *S. latifolia* or *S.*  
230 *dioica*. The loss of different sets of genes from the Y-chromosome may contribute to the evolution of  
231 reproductive isolation between closely related species with young actively degenerating sex  
232 chromosomes.

233

#### 234 *Y-degeneration accelerates the evolution of X-linked gene expression*

235 If the loss of a Y-linked gene is “compensated” by the upregulation of its X-linked gametolog  
236 (dosage compensation), the loss of different sets of Y-linked genes in closely related species is  
237 expected to accelerate divergence in gene expression for X-linked genes. Evidence of dosage  
238 compensation in *S. latifolia* sex chromosomes has been reported in two previous papers (Muyle *et al.*  
239 2012; Papadopulos *et al.* 2015), though two other analyses found no dosage compensation in that  
240 species (Bergero *et al.* 2015; Chibalina & Filatov 2011). To test whether the loss of different sets of  
241 genes from the Y-chromosomes of the two species accelerates divergence in X-linked gene  
242 expression, the analysis focused on 88 (=46+42, see previous section) sex-linked genes that are  
243 inferred to have lost Y-linked gametolog expression following the divergence of *S. latifolia* and *S.*  
244 *dioica*. Interestingly, the expression of X-linked gametologs of such genes in males is consistently  
245 higher in the species where the Y-copy is already degenerate (Fig. S2). In particular, the 46 sex-  
246 linked genes that lack expression of the Y-linked copy in *S. latifolia*, but are still actively expressed in  
247 *S. dioica*, show significantly higher expression of the X-linked gametolog in *S. latifolia* males,  
248 compared to *S. dioica* males (mean FPKM =  $14.7 \pm 33.04$  vs  $5.4 \pm 9.42$ ; paired t-test,  $P = 0.0069$ ).  
249 Conversely, for the 42 genes that lacked expression of the Y-linked copy in *S. dioica*, but not in *S.*  
250 *latifolia*, the former species shows significantly higher expression of X-linked gametologs, compared  
251 to the latter (mean FPKM =  $14.5 \pm 21.08$  vs  $8.9 \pm 14.82$ ; paired t-test,  $P = 0.009$ ). Thus, degeneration  
252 of different sets of Y-linked genes in *S. latifolia* and *S. dioica* is associated with the evolution of  
253 divergent gene expression on the X-chromosome, which may be partly responsible for the large-X  
254 effect in these species.

255

#### 256 *Faster-X divergence for sex-linked and autosomal genes in Silene*

257 A faster evolution of X(or Z)-linked genes compared to autosomal genes has been reported in  
258 many animal species (Meisel & Connallon 2013) and is thought to contribute to the occurrence of HR

259 and the LXE. In most cases, the evidence for faster-X comes from sequence-based comparisons of  
260 evolutionary rates on X(or Z)-linked and autosomal genes, with the “rate” measured as  
261 nonsynonymous divergence ( $dN$ ) normalised by synonymous divergence ( $dS$ ) (Charlesworth *et al.*  
262 2018; Meisel & Connallon 2013). Such a comparison for divergence between *S. latifolia* and *S. dioica*  
263 also reveals significantly higher  $dN/dS$  ratio for X-linked compared to autosomal genes (Fig. 2A;  
264 Wilcoxon rank sum test,  $W = 84278$   $P = 0.00145$ ). However, there are far fewer studies comparing  
265 evolutionary rates at gene expression level. To test whether gene expression diverges faster for the  
266 X-linked compared to autosomal genes, I quantified gene expression divergence between *S. latifolia*  
267 and *S. dioica* as the absolute value of the difference between median expression in the two species  
268 normalised by their average ( $D_e$  (Meisel *et al.* 2012)). The X-linked genes showed significantly higher  
269 expression divergence (Wilcoxon two sided test  $P < 0.00001$ ), compared to 1787 autosomal genes  
270 genetically mapped in the previous study (Papadopulos *et al.* 2015). This was the case regardless of  
271 the sex analysed, though the signal in males was stronger (Fig. 2B). The Y-linked genes showed  
272 much higher expression divergence compared to the X-linked and autosomal genes in males.  
273 Provided the time of species divergence is the same for different chromosomes, I hereby refer to  
274 higher divergence as “faster” divergence. Faster expression divergence of the sex-linked genes was  
275 detectable regardless of the gene expression level (Fig. S3), thus it cannot be explained by the over-  
276 or under-representation of highly expressed genes on the X-chromosome. Furthermore, excluding  
277 weakly expressed genes (Fig. S4) or genes with significantly sex-biased expression does not change  
278 the conclusion (Fig. S5).

279 Faster expression divergence of X-linked genes in *Silene* may be caused by a range of factors that  
280 can be classified into selective and non-selective (e.g. demographic). Selective hypotheses for faster-  
281 X expression divergence include evolution of dosage compensation driven by selection to  
282 compensate for loss of rapidly degenerating Y-chromosome, as well as the classic arguments of the  
283 “faster-X” theory (Charlesworth *et al.* 1987). On the other hand, non-selective explanations include the  
284 difference in effective population size ( $N_e$ ) – the smaller  $N_e$  for the X-linked genes enables a larger  
285 fraction of weakly deleterious mutations to be fixed by drift (Mank *et al.* 2010; Vicoso & Charlesworth  
286 2009), result in faster expression divergence on X-chromosome compared to autosomes. Another  
287 non-selective factor – differential interspecific gene flow between the X and the autosomes reported  
288 for *S. latifolia* and *S. dioica* (Hu & Filatov 2016), may also result in faster expression divergence for X-

289 linked compared to autosomal genes. The non-selective and selective hypotheses for faster  
290 expression divergence on the X-chromosome make different predictions and can potentially be  
291 distinguished from each other. In particular, the non-selective explanations apply equally to all X-  
292 linked genes regardless of the presence of functional Y-linked gametologs for the particular X-linked  
293 genes. On the other hand, both evolving dosage compensation hypothesis and the classic “faster-X”  
294 theory predict that X-linked genes that lost Y-linked gametologs should diverge faster compared to  
295 genes actively expressed on both the X- and the Y-chromosomes.

296 To distinguish between the non-selective and selective hypotheses, I compared divergence  
297 between *S. latifolia* and *S. dioica* males for 334 hemizygous (no expression from the Y in either  
298 species; hereafter  $X_{noYexpr}$ ) and 399 non-hemizygous (Y-copy is actively expressed in both species;  
299 hereafter  $X_{Yexpr}$ ) X-linked genes. Consistent with the selective hypotheses, the  $X_{noYexp}$  genes show  
300 significantly faster expression divergence between *S. latifolia* and *S. dioica* males compared to  $X_{Yexpr}$   
301 genes (Fig 3A; Wilcoxon rank sum test  $W = 23019$ ,  $P = 0.0082$ ). The rate of sequence divergence  
302 ( $dN/dS$ ) shows the same trend, though the difference between  $X_{noYexp}$  and  $X_{Yexpr}$  genes is not  
303 significant (Fig 3B; Wilcoxon rank sum test  $W = 11302$ ,  $P = 0.632$ ).

304

305 *Does haploid expression affect species divergence?*

306 The X-chromosome is thought to be “special” because X-linked recessive mutations are expressed  
307 in the phenotype in hemizygous males (Charlesworth *et al.* 1987; Vicoso & Charlesworth 2006).  
308 However, in plants, a significant proportion of genes in the genome are expressed at the haploid  
309 stage of lifecycle – the male gametophyte (pollen in Angiosperms) (Honys & Twell 2004). Thus, if  
310 interspecific hybrid incompatibilities are primarily recessive, then the plant genes expressed in the  
311 gametophyte may be expected to play a similarly “special” role in plant speciation as the sex  
312 chromosomes in animals. The availability of pollen expression data for *S. latifolia* (Chibalina & Filatov  
313 2011) makes it possible to test whether exposure to haploid selection in the plant gametophyte affects  
314 the rate of gene expression divergence between the species.

315 The proportion of genes that evolved differential expression between *S. latifolia* and *S. dioica* is  
316 significantly lower for genes expressed in the gametophyte ( $G_{expr}$ ), compared to genes with a  
317 predominantly sporophytic expression ( $S_{expr}$ ; 9.9% and 12.3%, respectively; 2x2 contingency  
318  $\chi^2=26.12$ ,  $P < 0.000001$ ). Furthermore,  $D_e$  is significantly lower for  $G_{expr}$  compared to  $S_{expr}$  genes (Fig

319 4A; Wilcoxon rank sum test  $P < 10^{-6}$ ), indicating that stronger purifying selection at the haploid stage  
320 of the lifecycle prevents expression divergence between species for most genes. In animals, the  
321 genes with functions related to male gametogenesis, such as accessory gland proteins in *Drosophila*,  
322 are often reported to show faster evolutionary rates, possibly due to positive selection fuelled by  
323 sexual conflict (e.g. (Ahmed-Braimah *et al.* 2017)). The *S. latifolia* and *S. dioica* genes that are over-  
324 expressed in pollen ( $P_{\text{expr}}$ ; the category includes genes with at least five-fold higher expression in  
325 pollen compared to male somatic tissues) also show accelerated expression divergence, compared to  
326 non-overexpressed genes; in particular, the distribution of  $D_e$  for  $P_{\text{expr}}$  genes is significantly shifted  
327 upwards compared to  $G_{\text{expr}}$  and  $S_{\text{expr}}$  genes (Fig 4A; Wilcoxon rank sum test  $P < 10^{-6}$  for both  
328  $P_{\text{expr}}:G_{\text{expr}}$  and  $P_{\text{expr}}:S_{\text{expr}}$  comparisons). Excluding weakly expressed genes does not change this  
329 result (Fig. S6A).

330 The distribution of  $P_{\text{expr}}$  genes across the X, Y and autosomes did not deviate from random. On the  
331 other hand,  $G_{\text{expr}}$  genes were significantly depleted on the X- and Y-chromosomes (83.6% and 62.1%  
332 of the expected, respectively; G-test,  $P < 0.001$ ), and  $S_{\text{expr}}$  genes were significantly depleted on the  
333 autosomes (79.8% of the expected; G-test,  $P < 0.001$ ) and over-represented on the X- and Y-  
334 chromosomes (132.5% and 176.4% of the expected; G-test,  $P < 0.001$ ). To test whether the under-  
335 representation of slow evolving  $G_{\text{expr}}$  genes on the sex chromosomes and under-representation of  
336 fast-evolving  $S_{\text{expr}}$  genes on the autosomes may be the cause of accelerated expression divergence  
337 on the X-chromosome (Fig. 1), the  $D_e$  was calculated separately for sex-linked and autosomal  $G_{\text{expr}}$   
338 and  $S_{\text{expr}}$  genes. As a faster evolution of sex-linked genes compared to those that are autosomal is  
339 still clearly detectable (Figures 4B and S6B) and significant (Wilcoxon rank sum test  $P < 0.0001$ ) for  
340 the  $G_{\text{expr}}$  and for  $S_{\text{expr}}$  genes separately, the uneven distribution of  $S_{\text{expr}}$  and  $G_{\text{expr}}$  genes across the  
341 chromosomes cannot explain faster expression divergence on the X- and Y-chromosomes compared  
342 to the autosomes.

343

## 344 Discussion

345 *Is the Silene Y-chromosome degenerate “enough”?*

346 Consistent with the dominance theory, species incompatibilities expressed in the heterogametic  
347 sex evolve faster in taxa with larger X-chromosomes (Turelli & Begun 1997). Although the X-  
348 chromosome in *S. latifolia* and *S. dioica* is relatively large (2<sup>nd</sup> largest chromosome in the genome

349 (Armstrong & Filatov 2008)), prior to this study, it had not been clear how many X-linked genes in *S.*  
350 *latifolia* and *S. dioica* comply with the assumption of the dominance theory that X-chromosome is  
351 hemizygous in the heterogametic sex. Due to the recent origin of sex chromosomes in the ancestor of  
352 *S. latifolia* and *S. dioica*, it was widely assumed that the Y-chromosome in these species is likely to be  
353 non-degenerate, which was supported by the isolation of apparently functional Y-linked genes in early  
354 low-throughput studies (reviewed by (Charlesworth 2008)). More recent analyses based on  
355 transcriptome (Bergero & Charlesworth 2011; Bergero *et al.* 2015; Chibalina & Filatov 2011) and  
356 genome (Papadopoulos *et al.* 2015) sequence data reported various degrees of genetic degeneration  
357 (10 to 30%) of Y-linked genes in *S. latifolia*.

358 This study demonstrates that over 30% of Y-linked genes have effectively lost expression in either  
359 or both *S. latifolia* and *S. dioica*, leaving their X-linked gametologs hemizygous in males. The extent of  
360 Y-chromosome degeneration in *S. latifolia* and its close dioecious relatives may be considerably  
361 higher than 30% given that only genes with detectable X- and Y-linked gametologs were used in the  
362 analyses. The genes without detectable Y-linked gametologs (X-only genes) were excluded because  
363 an unknown proportion may represent translocations of autosomal genes to the X-chromosome rather  
364 than the loss from the Y-chromosome. Unfortunately, in the absence of chromosome-level assemblies  
365 of the genomes of *S. latifolia* and its non-dioecious relative, such as *S. vulgaris*, it is not possible to  
366 distinguish the X-ancestral genes that lost Y-linked gametologs from the genes that were translocated  
367 to the X and never had a Y-linked copy. A study that attempted to address this question using a  
368 comparative analysis of genetic maps of dioecious *S. latifolia* and non-dioecious *S. vulgaris* revealed  
369 that all 16 tested X-only genes of *S. latifolia* are X-ancestral rather than translocated to the X-  
370 chromosome secondarily (Bergero *et al.* 2015). Thus, the proportion of X-only genes translocated to  
371 the X-chromosome after the sex chromosomes evolved may be small, and most of the X-only genes  
372 are likely to have lost their Y-linked gametologs. When the X-only genes are included in the analysis,  
373 the proportion of Y-degenerate genes increases to about 50%.

374 The theory expressing the conditions for Haldane's rule as a function of  $p_x$ , the proportion of hybrid  
375 incompatibilities that are X-linked hemizygous ((Turelli & Orr 1995) equation B2), shows that with a  
376 smaller  $p_x$  the conditions for HR evolution become more restrictive (e.g. see fig. 1 in (Orr & Turelli  
377 1996)). The X-chromosome in *S. latifolia* and *S. dioica* contains about 10% of the *Silene* genome  
378 (Armstrong & Filatov 2008). If 50% of the Y-linked genes are degenerate, about 5% of the *Silene*

379 genome is hemizygous in males, which gives  $p_x \sim 0.1$  (from formula A3 in (Turelli & Begun 1997),  
380 assuming all incompatibilities involve two loci). To place this in the context of other species,  $p_x$  in  
381 *Silene* is larger than in mammals ( $p_x \sim 0.05$ ), but smaller than *Drosophila melanogaster*, where  $p_x \sim$   
382 0.36 (Orr & Turelli 1996). As both *Drosophila* and mammals comply with HR, it appears that the  
383 extent of hemizygoty in *S. latifolia* and *S. dioica* males is likely sufficient for the dominance theory to  
384 explain the presence of the Haldane's rule (Brothers & Delph 2010) in these species.

385

386 *Does on-going Y-degeneration contribute to the evolution of species divergence?*

387 Interestingly, 88 of the Y-linked genes analysed, have lost expression since the divergence of *S.*  
388 *latifolia* and *S. dioica*, demonstrating that genetic degeneration is rapidly progressing on Y-  
389 chromosomes of these species. For a comparison, only three human Y-linked genes have lost their  
390 function since they diverged from chimpanzees ~6 million years ago (Bellott *et al.* 2014). The  
391 comparison with human sex chromosomes is appropriate given the age of the youngest human  
392 stratum is similar to the age of *Silene* sex chromosomes after adjusting for the difference in  
393 generation time. Given the similar interspecific divergence at silent sites (~1.5%) in two species pairs  
394 (*S. latifolia/S. dioica* and *Homo sapiens/Pan troglodites*), genetic degeneration of Y-linked genes is  
395 proceeding at least 10 times faster in *Silene* compared to humans. On the other hand, the rate of Y-  
396 degeneration in *Silene* is comparable to that reported for recently evolved neo-Y chromosome of  
397 *Drosophila miranda* (Bachtrog *et al.* 2008).

398 Faster genetic degeneration is expected (e.g. (Bachtrog 2008)) for younger gene-rich Y-  
399 chromosomes (such as in *S. latifolia*), compared to the older stages when only a few functional genes  
400 remain on the Y-chromosome (as is the case in humans (Bellott *et al.* 2014)). This is the case  
401 because Y-chromosome degeneration is thought to be, at least partly, caused by interference of  
402 natural selection acting on multiple mutations linked together in the non-recombining region  
403 (Charlesworth 2008). With many functional genes linked together on the same Y-chromosome,  
404 natural selection is unable to eliminate deleterious mutations and fix advantageous mutations,  
405 resulting in gradual disfunctionalisation of Y-linked genes (Charlesworth & Charlesworth 2000). Thus,  
406 in species with young sex chromosomes, such as *S. latifolia* and *S. dioica*, the Y-chromosome  
407 contains hundreds to thousands of functional genes and Y-degeneration may be sufficiently rapid for  
408 closely related species to lose different sets of Y-linked genes. If the X-linked gametologs evolve

409 altered gene expression to compensate for the loss of Y-linked gametologs (i.e. gene-by-gene dosage  
410 compensation, such as reported for chicken (Mank & Ellegren 2009)), the introgressed X may not be  
411 compatible with the “local” Y-chromosome, resulting in the reduced fitness of hybrids. Such a  
412 reduction in fitness is expected to be present primarily in the heterogametic sex where the “foreign” Y  
413 meets the “local” X (or vice-versa). The expression analyses for *S. latifolia* and *S. dioica* sex-linked  
414 genes indicate that this “divergent Y degeneration” model is a plausible mechanism for the presence  
415 of HR and the LXE in species with young actively degenerating sex chromosomes.

416

417 *Y-degeneration, dosage compensation and faster-X evolution of gene expression in Silene*

418 Faster evolution of X-linked compared to autosomal genes may contribute to HR and the LXE  
419 because under the faster-X scenario recessive species incompatibilities would accumulate faster on  
420 the X-chromosome, resulting in a disproportionate contribution of the X-linked genes to reproductive  
421 barriers between the species (Coyne & Orr 1989). The analyses reported above provide compelling  
422 evidence for faster evolution of expression divergence in X-linked genes compared to autosomal  
423 genes. However, the causes of the accelerated evolution of gene expression on the *Silene* X-  
424 chromosome remain unclear.

425 An intriguing possibility explaining faster expression divergence in X-linked compared to autosomal  
426 genes is an on-going evolution of dosage compensation on the X-chromosome. The presence of  
427 dosage compensation on the *S. latifolia* sex chromosomes is a contested issue, with two studies  
428 finding no evidence supporting dosage compensation (Bergero *et al.* 2015; Chibalina & Filatov 2011)  
429 and two studies reporting the presence of at least partial dosage compensation in *S. latifolia* (Muyle *et al.*  
430 2012; Papadopulos *et al.* 2015). If dosage compensation evolves in *S. latifolia* (and, by extension,  
431 *S. dioica*), this would accelerate gene expression divergence for X-linked genes, with selection to  
432 compensate for loss of Y-linked gametologs particularly affecting  $X_{noY_{expr}}$  genes. This is consistent  
433 with the observation of faster gene expression divergence at  $X_{noY_{expr}}$  compared to  $X_{Y_{expr}}$  genes. If  
434 dosage compensation in *Silene* evolves gene-by-gene (as opposed to chromosome-wide dosage  
435 compensation found in mammals and *Drosophila*), the X-linked genes with functional Y-linked  
436 gametologs would not be expected to be affected by this process, yet,  $X_{Y_{expr}}$  genes still show  
437 significantly faster evolution of gene expression compared to autosomal genes, suggesting some form  
438 of nascent chromosome-wide dosage compensation arising in these species.

439 If evolving dosage compensation is indeed the driver of faster-X evolution of gene expression in  
440 *Silene*, this could have broad implications on the speciation literature, in particular for species pairs  
441 with young sex chromosomes. Such species would be expected to show transgressive gene  
442 expression for sex-linked genes in male interspecific hybrids. Transgressive expression should be  
443 particularly pronounced for ‘divergently degenerate’ Y-linked genes – the genes that are lost in one  
444 and retained in the other hybridising species. These predictions need to be tested in future studies.

445

#### 446 *The alternative explanations for faster-X evolution of gene expression in Silene*

447 The alternative explanations to faster-X gene expression driven by species-specific dosage  
448 compensation look less plausible. In particular, the non-selective explanations – the difference in  
449 effective population size ( $N_e$ ) or differential gene flow between the X-linked and autosomal genes are  
450 not compatible with the fact that  $X_{noYexpr}$  genes are diverging faster than  $X_{Yexpr}$  genes (Fig. 3).  
451 Furthermore, effective population size in *S. latifolia* and *S. dioica* is relatively large and likely  
452 comparable to that in *Drosophila* given the similar genetic diversity in these species (average  
453 heterozygosity at silent sites,  $\pi \sim 1.5\%$ ). For an  $N_e$  as large as in *Drosophila* purifying selection is  
454 expected to be highly efficient and the effect of a slightly smaller  $N_e$  for the X-chromosome should be  
455 marginal (Mank *et al.* 2010).

456 Faster-X evolution for gene expression in *Silene* is still detectable after exclusion of sex-biased  
457 genes (Fig. S5), or genes that are expressed or non-expressed in the gametophyte (Fig. 4B), or  
458 genes with high or with low expression only (Fig. S3). This indicates that none of these factors fully  
459 accounts for accelerated evolution of gene expression on the X-chromosome. However, the data  
460 presented above are compatible with the classic “faster-X” theory (Charlesworth *et al.* 1987) that  
461 predicts faster evolution of X-linked genes based on their female-biased transmission and exposure of  
462 X-linked recessive beneficial alleles in hemizygous males. Distinguishing between the classic faster-X  
463 theory and the species-specific dosage compensation hypothesis would require the analyses of gene  
464 expression in F1 interspecific hybrids, where transgressive segregation is predicted by the latter  
465 hypothesis. Unfortunately such data for crosses between *S. latifolia* and its close relatives is currently  
466 unavailable.

467 It is possible that the faster-X evolution in *Silene* has a composite nature, with accelerated positive  
468 selection, relaxed purifying selection, underrepresentation of slow evolving pollen-expressed genes

469 on the X-chromosome, differential gene flow on the X and autosomes and evolving dosage  
470 compensation all contributing to the observed faster evolution of gene expression on the X-  
471 chromosome. It remains unclear whether this “faster-X” for gene expression plays any role in  
472 speciation and future studies should test whether the *Silene* X-chromosome accumulates species  
473 incompatibilities faster than the autosomes.

474

475 *Does haploid expression of genes in the plant gametophyte contribute to speciation?*

476 The causes underlying HR and the LXE in *Silene* remain unclear, though the study that analysed  
477 the genetic basis of HR between *S. latifolia* and *S. diclinis* concluded that “the genetic architecture of  
478 Haldane’s rule in dioecious plants may differ from those commonly found in animals” (Demuth *et al.*  
479 2014). Widespread haploid expression of genes in the plant gametophyte may be one of the reasons  
480 for the difference in genetic architecture of HR between the two kingdoms. For example, in plants,  
481 hybrid male sterility may be caused by recessive species incompatibilities expressed in pollen, where  
482 over half of the genes in the genome are actively expressed (Honys & Twell 2004). These recessive  
483 species incompatibilities may not have anything to do with the sex chromosomes, but they can be  
484 interpreted as a manifestation of Haldane’s rule in species with heterogametic males. However, the  
485 first report of HR in plants – reduced pollen viability in hybrids between dioecious *Silene latifolia* and  
486 its close relatives (Brothers & Delph 2010), is unlikely to be the result of recessive hybrid  
487 incompatibilities in haploid pollen because the observation of hybrid male sterility in *Silene* was based  
488 on pollen stainability, a phenotype that is detectable before the haploid gene expression stage. Still,  
489 such mimicking of HR for hybrid sterility with autosomal recessive species incompatibilities expressed  
490 in the gametophyte may occur in other dioecious plants. More generally, the implications of haploid  
491 expression in plant gametophytes on hybrid incompatibilities remains unexplored. The evolutionary  
492 role of widespread haploid gene expression in plant gametophytes is poorly studied and its role in  
493 plant speciation remains unclear. Clearly, haploid expression in the plant gametophyte is potentially  
494 an important factor in plant evolution and it deserves more attention in the speciation literature.

495

## 496 **Conclusions**

497 The analyses reported above comprise the first demonstration of the “faster-X” effect in plants,  
498 though the underlying causes of the faster expression divergence on the *Silene* X-chromosome

499 appear different to the classic “faster-X” theory (Charlesworth *et al.* 1987). The most likely cause of  
 500 the discovered “faster-X” in *Silene* is species-specific evolution of dosage compensation that is likely  
 501 driven by on-going rapid degeneration of the Y-chromosome in *S. latifolia* and *S. dioica*. Although the  
 502 connection between the observed “faster-X” evolution with HR and the LXE in *Silene* remains to be  
 503 established, our results demonstrate that Y-degeneration and dosage compensations can be  
 504 sufficiently rapid to proceed in species-specific manner, even between closely related species with  
 505 young gene rich sex chromosomes. Potentially, this process may be a significant contributor of hybrid  
 506 incompatibilities, though its role in speciation remains to be studied.

507

### 508 **Acknowledgements**

509 This work was supported by grants from the BBSRC (Grant BB/P009808/1 to DAF). I thank Bruno  
 510 Nevado, Daven Presgraves and two anonymous reviewers for helpful comments, Michael Chester for  
 511 proofreading the text and the staff at the Wellcome Trust Centre (Oxford) for sequencing and initial  
 512 data processing.

513

### 514 **Data Accessibility**

515 The data used in this study are available from SRA database (project number PRJNA453413).

516

### 517 **Author Contributions**

518 DAF designed the study, generated and analysed the data and wrote the manuscript.

519

### 520 **References**

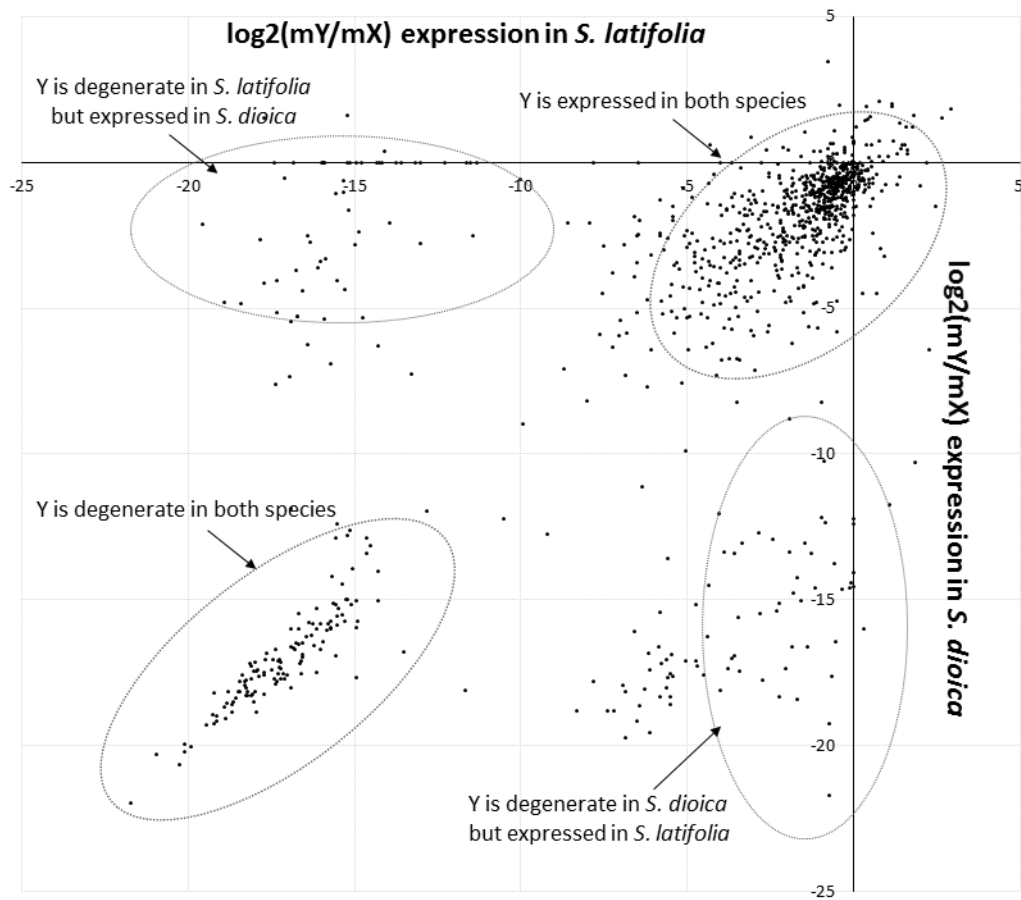
- 521 Ahmed-Braimah YH, Unckless RL, Clark AG (2017) Evolutionary dynamics of male reproductive  
 522 genes in the *Drosophila virilis* subgroup. *G3 (Bethesda)* **7**, 3145-3155.  
 523 Armstrong SJ, Filatov DA (2008) A cytogenetic view of sex chromosome evolution in plants.  
 524 *Cytogenet Genome Res* **120**, 241-246.  
 525 Bachtrog D (2008) The temporal dynamics of processes underlying Y chromosome degeneration.  
 526 *Genetics* **179**, 1513-1525.  
 527 Bachtrog D, Hom E, Wong KM, Maside X, de Jong P (2008) Genomic degradation of a young Y  
 528 chromosome in *Drosophila miranda*. *Genome Biol* **9**, R30.  
 529 Bachtrog D, Mank JE, Peichel CL, *et al.* (2014) Sex determination: why so many ways of doing it?  
 530 *PLoS Biol* **12**, e1001899.  
 531 Bellott DW, Hughes JF, Skaletsky H, *et al.* (2014) Mammalian Y chromosomes retain widely  
 532 expressed dosage-sensitive regulators. *Nature* **508**, 494-499.  
 533 Bergero R, Charlesworth D (2011) Preservation of the Y transcriptome in a 10-million-year-old plant  
 534 sex chromosome system. *Curr Biol* **21**, 1470-1474.  
 535 Bergero R, Qiu S, Charlesworth D (2015) Gene loss from a plant sex chromosome system. *Curr Biol*  
 536 **25**, 1234-1240.

- 537 Bernasconi G, Antonovics J, Biere A, *et al.* (2009) *Silene* as a model system in ecology and evolution.  
538 *Heredity (Edinb)* **103**, 5-14.
- 539 Brothers AN, Delph LF (2010) Haldane's rule is extended to plants with sex chromosomes. *Evolution*  
540 **64**, 3643-3648.
- 541 Charlesworth B, Campos JL, Jackson BC (2018) Faster-X evolution: Theory and evidence from  
542 *Drosophila*. *Mol Ecol* **in press**.
- 543 Charlesworth B, Charlesworth D (2000) The degeneration of Y chromosomes. *Philos Trans R Soc*  
544 *Lond B Biol Sci* **355**, 1563-1572.
- 545 Charlesworth B, Coyne JA, Barton NH (1987) The relative rates of evolution of sex-chromosomes and  
546 autosomes. *American Naturalist* **130**, 113-146.
- 547 Charlesworth D (2008) Sex chromosome origins and evolution. In: *Evolutionary Genomics and*  
548 *Proteomics* (eds. Pagel M, Pomiankowski A), pp. 207-240. Sinauer Associates, Sunderland.
- 549 Charlesworth D (2015) Plant contributions to our understanding of sex chromosome evolution. *New*  
550 *Phytol* **208**, 52-65.
- 551 Chibalina MV, Filatov DA (2011) Plant Y chromosome degeneration is retarded by haploid purifying  
552 selection. *Curr Biol* **21**, 1475-1479.
- 553 Ciupercescu DD, Veuskens J, Mouras A, *et al.* (1990) Karyotyping *Melandrium album*, a dioecious  
554 plant with heteromorphic sex chromosomes. *Genome* **33**, 556-562.
- 555 Coyne JA, Orr HA (1989) The two rules of speciation. In: *Speciation and its Consequences* (eds. Otte  
556 D, Endler J), pp. 180-207. Sinauer Associates.
- 557 Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates, Sunderland, Mass.
- 558 Delph LF, Demuth JP (2016) Haldane's rule: genetic bases and their empirical support. *J Hered* **107**,  
559 383-391.
- 560 Demuth JP, Flanagan RJ, Delph LF (2014) Genetic architecture of isolation between two species of  
561 *Silene* with sex chromosomes and Haldane's rule. *Evolution* **68**, 332-342.
- 562 Dufresnes C, Majtyka T, Baird SJ, *et al.* (2016) Empirical evidence for large X-effects in animals with  
563 undifferentiated sex chromosomes. *Sci Rep* **6**, 21029.
- 564 Ellegren H, Smeds L, Burri R, *et al.* (2012) The genomic landscape of species divergence in *Ficedula*  
565 flycatchers. *Nature* **491**, 756-760.
- 566 Favre A, Widmer A, Karrenberg S (2017) Differential adaptation drives ecological speciation in  
567 champions (*Silene*): evidence from a multi-site transplant experiment. *New Phytol* **213**, 1487-  
568 1499.
- 569 Frank SA (1991) Divergence of meiotic drive-suppression systems as an explanation for sex-biased  
570 hybrid sterility and inviability. *Evolution* **45**, 262-267.
- 571 Haldane JBS (1922) Sex-ratio and unisexual sterility in hybrid animals. *J. Genet.* **12**, 101-109.
- 572 Honys D, Twell D (2004) Transcriptome analysis of haploid male gametophyte development in  
573 *Arabidopsis*. *Genome Biol* **5**, R85.
- 574 Howell EC, Armstrong SJ, Filatov DA (2009) Evolution of neo-sex chromosomes in *Silene diclinis*.  
575 *Genetics* **182**, 1109-1115.
- 576 Hu XS, Filatov DA (2016) The large-X effect in plants: increased species divergence and reduced  
577 gene flow on the *Silene* X-chromosome. *Molecular Ecology* **25**, 2609-2619.
- 578 Hurst LD, Pomiankowski A (1991) Causes of sex ratio bias may account for unisexual sterility in  
579 hybrids: a new explanation of Haldane's rule and related phenomena. *Genetics* **128**, 841-858.
- 580 Jiggins CD, Linares M, Naisbit RE, *et al.* (2001) Sex-linked hybrid sterility in a butterfly. *Evolution* **55**,  
581 1631-1638.
- 582 Larson EL, Keeble S, Vanderpool D, Dean MD, Good JM (2017) The composite regulatory basis of  
583 the large X-effect in mouse speciation. *Mol Biol Evol* **34**, 282-295.
- 584 Laurie CC (1997) The weaker sex is heterogametic: 75 years of Haldane's rule. *Genetics* **147**, 937-  
585 951.
- 586 Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a  
587 reference genome. *BMC Bioinformatics* **12**, 323.
- 588 Lima TG (2014) Higher levels of sex chromosome heteromorphism are associated with markedly  
589 stronger reproductive isolation. *Nat Commun* **5**, 4743.
- 590 Mank JE, Ellegren H (2009) All dosage compensation is local: Gene-by-gene regulation of sex-biased  
591 expression on the chicken Z chromosome. *Heredity* **102**, 312-320.
- 592 Mank JE, Vicoso B, Berlin S, Charlesworth B (2010) Effective population size and the Faster-X effect:  
593 empirical results and their interpretation. *Evolution* **64**, 663-674.
- 594 Masly JP, Presgraves DC (2007) High-resolution genome-wide dissection of the two rules of  
595 speciation in *Drosophila*. *PLoS Biol* **5**, e243.

- 596 Meisel RP, Connallon T (2013) The faster-X effect: integrating theory and data. *Trends Genet* **29**,  
597 537-544.
- 598 Meisel RP, Malone JH, Clark AG (2012) Faster-X evolution of gene expression in *Drosophila*. *PLoS*  
599 *Genet* **8**, e1003013.
- 600 Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian  
601 transcriptomes by RNA-Seq. *Nat Methods* **5**, 621-628.
- 602 Moyle LC, Muir CD, Han MV, Hahn MW (2010) The contribution of gene movement to the "two rules  
603 of speciation". *Evolution* **64**, 1541-1557.
- 604 Muyle A, Zemp N, Deschamps C, *et al.* (2012) Rapid de novo evolution of X chromosome dosage  
605 compensation in *Silene latifolia*, a plant with young sex chromosomes. *PLoS Biol* **10**,  
606 e1001308.
- 607 Orr HA (1997) Haldane's rule. *Annu. Rev. Ecol. Svst.* **28**, 195-218.
- 608 Orr HA, Turelli M (1996) Dominance and Haldane's rule. *Genetics* **143**, 613-616.
- 609 Papadopulos AS, Chester M, Ridout K, Filatov DA (2015) Rapid Y degeneration and dosage  
610 compensation in plant sex chromosomes. *Proc Natl Acad Sci U S A* **112**, 13021-13026.
- 611 Payseur BA, Krenz JG, Nachman MW (2004) Differential patterns of introgression across the X  
612 chromosome in a hybrid zone between two species of house mice. *Evolution* **58**, 2064-2078.
- 613 Presgraves DC, Orr HA (1998) Haldane's rule in taxa lacking a hemizygous X. *Science* **282**, 952-954.
- 614 Tao Y, Hartl DL (2003) Genetic dissection of hybrid incompatibilities between *Drosophila simulans*  
615 and *D. mauritiana*. III. Heterogeneous accumulation of hybrid incompatibilities, degree of  
616 dominance, and implications for Haldane's rule. *Evolution* **57**, 2580-2598.
- 617 Turelli M, Begun DJ (1997) Haldane's rule and X-chromosome size in *Drosophila*. *Genetics* **147**,  
618 1799-1815.
- 619 Turelli M, Moyle LC (2007) Asymmetric postmating isolation: Darwin's corollary to Haldane's rule.  
620 *Genetics* **176**, 1059-1088.
- 621 Turelli M, Orr HA (1995) The dominance theory of Haldane's rule. *Genetics* **140**, 389-402.
- 622 Turner LM, White MA, Tautz D, Payseur BA (2014) Genomic networks of hybrid sterility. *PLoS Genet*  
623 **10**, e1004162.
- 624 Vicoso B, Charlesworth B (2006) Evolution on the X chromosome: unusual patterns and processes.  
625 *Nat Rev Genet* **7**, 645-653.
- 626 Vicoso B, Charlesworth B (2009) Effective population size and the faster-X effect: an extended model.  
627 *Evolution* **63**, 2413-2426.
- 628 Wu C-I, Davis AW (1993) Evolution of postmating reproductive isolation: the composite nature of  
629 Haldane's rule and its genetic bases. *American Naturalist* **142**, 187-212.
- 630 Wu CI, Johnson NA, Palopoli MF (1996) Haldane's rule and its legacy: Why are there so many sterile  
631 males? *Trends in Ecology & Evolution* **11**, 281-284.
- 632 Zemp N, Tavares R, Muyle A, *et al.* (2016) Evolution of sex-biased gene expression in a dioecious  
633 plant. *Nat Plants* **2**, 16168.
- 634

635

636



637

638 Figure 1. Loss of gene expression on the Y-chromosome. The plot shows  $\log_2$ -transformed ratios639 of relative Y/X gene expression of individual sex-linked genes in *S. latifolia* (X-axis) and *S. dioica* (Y-

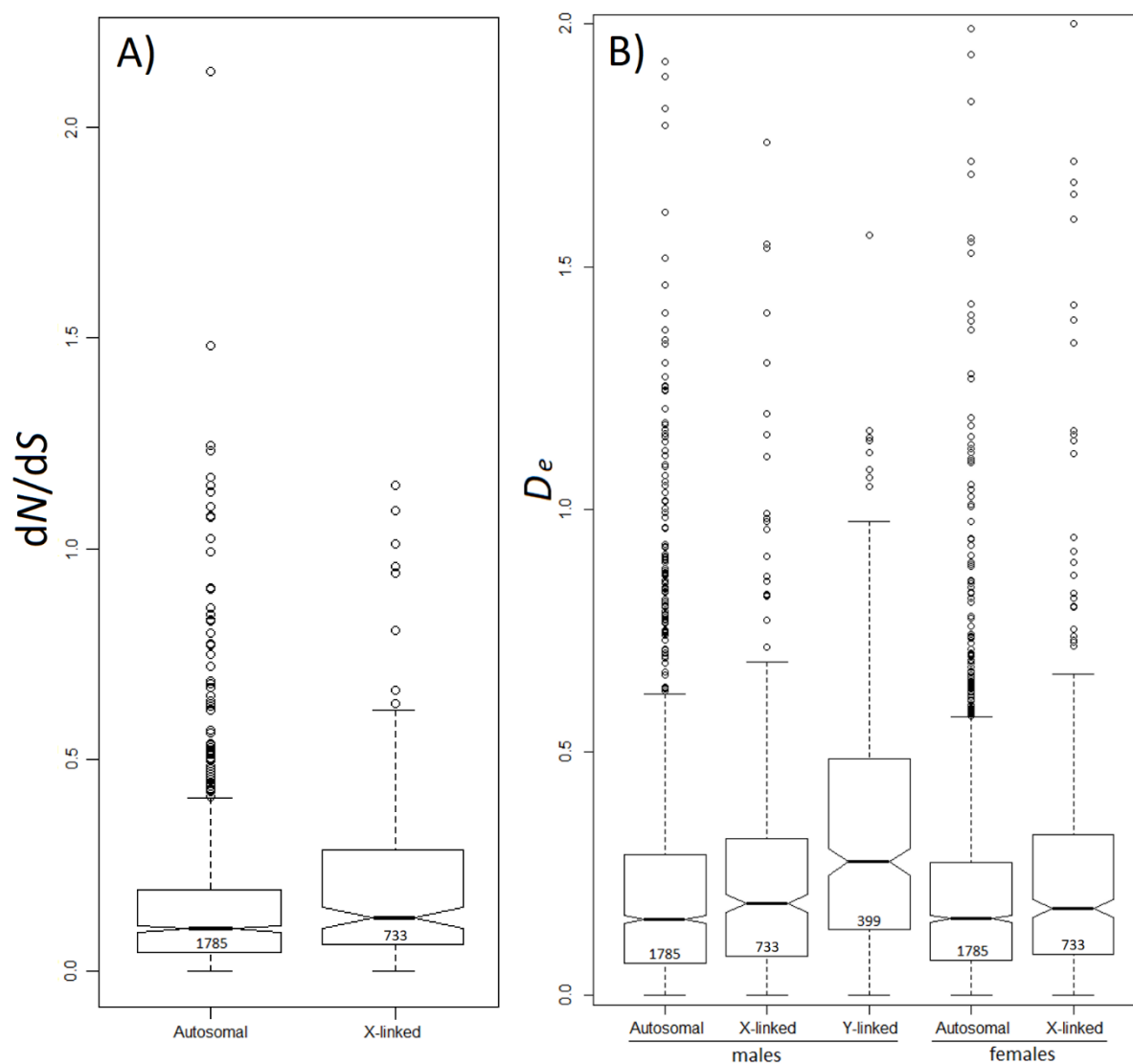
640 axis) males. To avoid infinity and division by zero errors for non-expressed genes, a small number

641 (0.0001) was added to all expression values. The ellipses were drawn by hand to highlight different

642 parts of the plot.

643

644

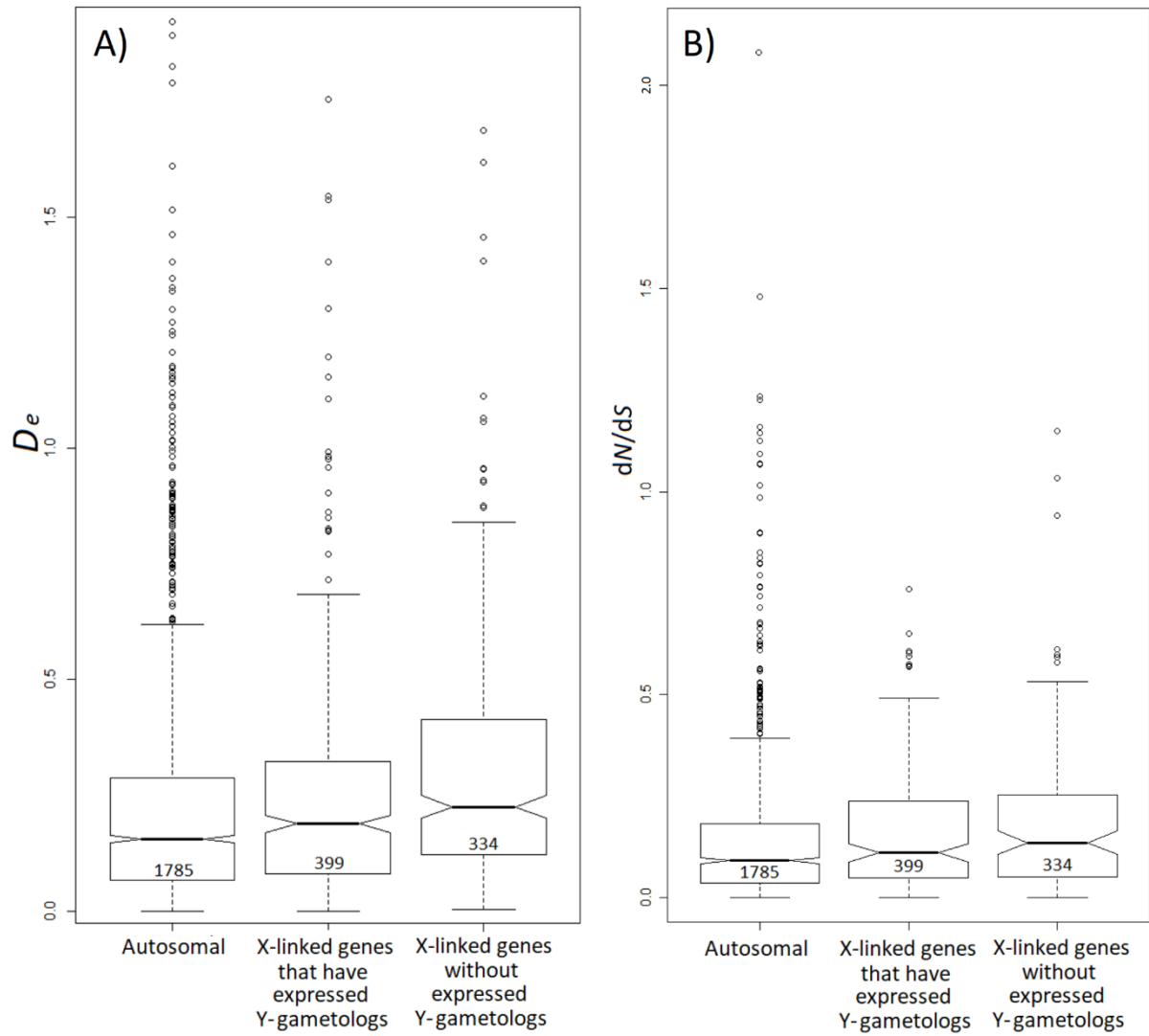


645

646 Figure 2. Non-synonymous ( $dN$ ) to synonymous ( $dS$ ) nucleotide substitution rate ratio (A) and gene  
 647 expression divergence (B) between *S. latifolia* and *S. dioica* for sex-linked and autosomal genes. The  
 648 numbers of genes analysed are shown inside the boxes.

649

650



651

652

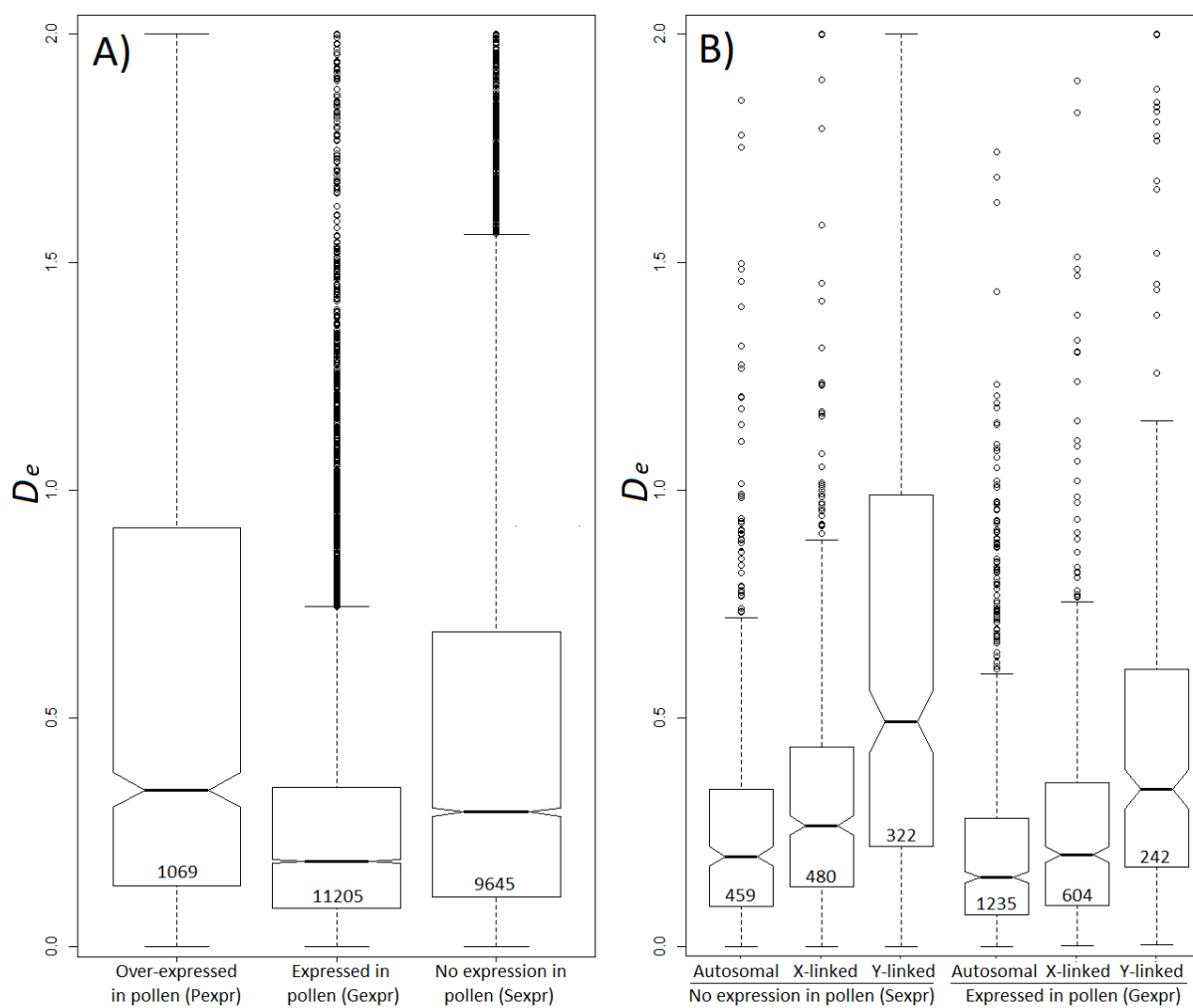
653

654

655

656

Figure 3. Faster-X evolution for gene expression divergence in males (A) and sequence divergence ( $dN/dS$ , panel B) between *S. latifolia* and *S. dioica*. X-linked genes are split into two categories for X-linked genes with expressed ( $X_{Y_{\text{expr}}}$ ) and non-expressed ( $X_{\text{no}Y_{\text{expr}}}$ ) Y-linked gametologs. The numbers of genes analysed are shown inside the boxes.



657

658

659

660

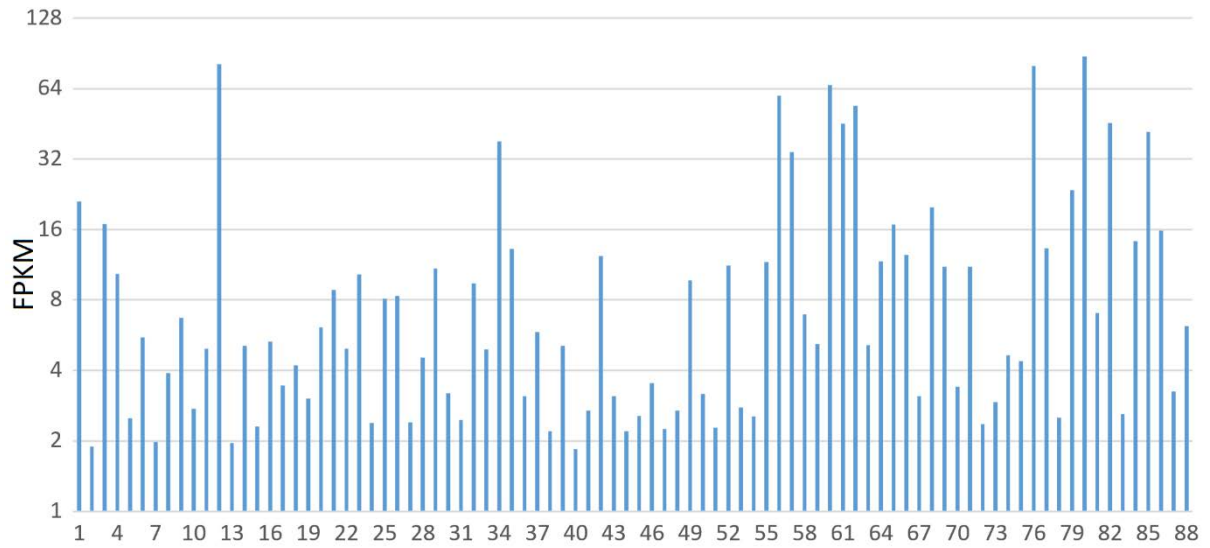
661

662

Figure 4. Gene expression divergence between *S. latifolia* and *S. dioica* males for genes predominantly expressed in the sporophyte ( $S_{expr}$ ) and the gametophyte ( $G_{expr}$  and  $P_{expr}$ ). A) All genes regardless of linkage; B) A comparison of autosomal and sex-linked genes. The numbers of genes analysed are shown inside the boxes.

663

664

**Suppl. Figures**

665

666

667

Figure S1 Expression of 88 *S. diclinis* Y-linked genes that lost expression in *S. latifolia* or *S. dioica*.

668

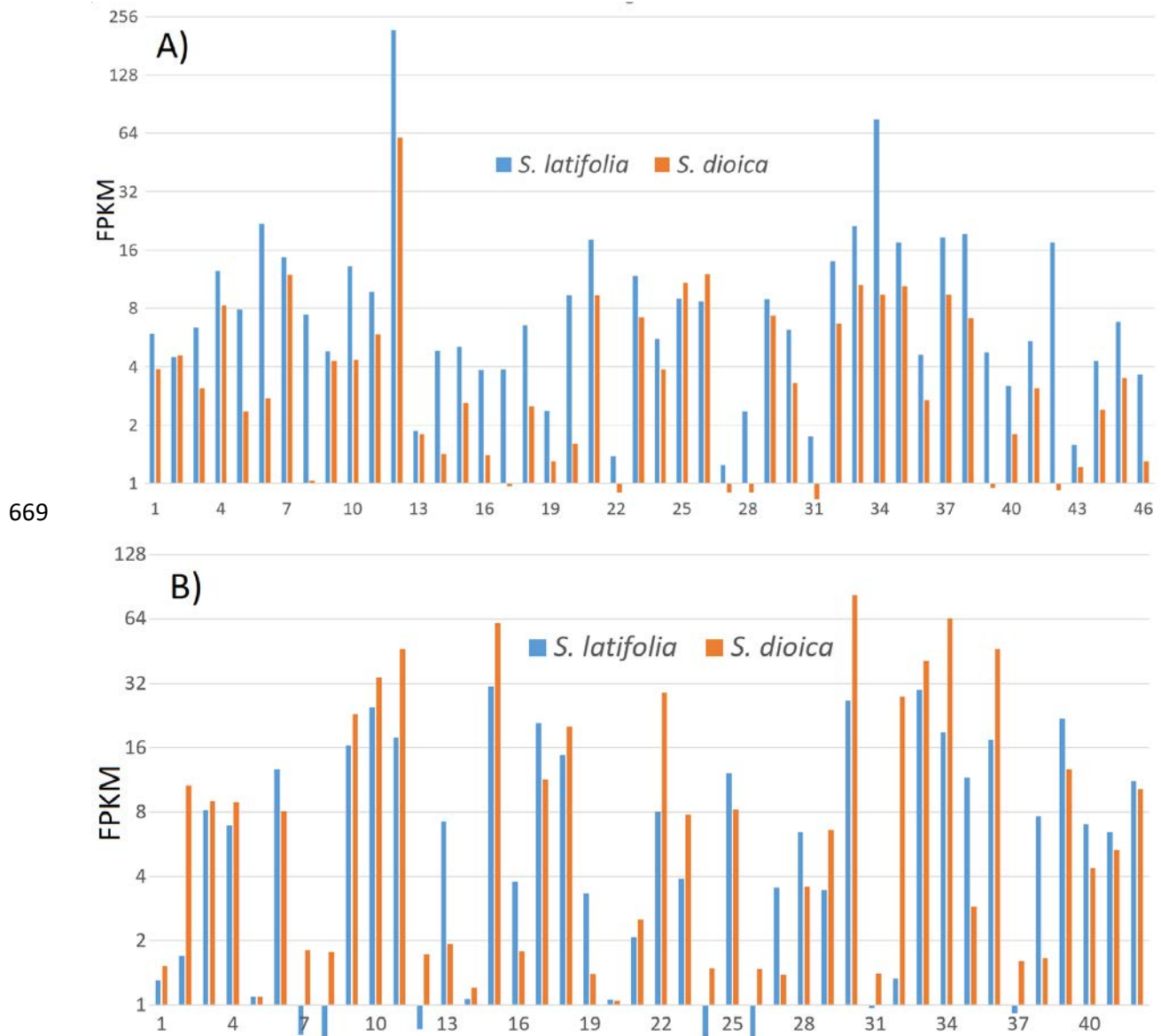
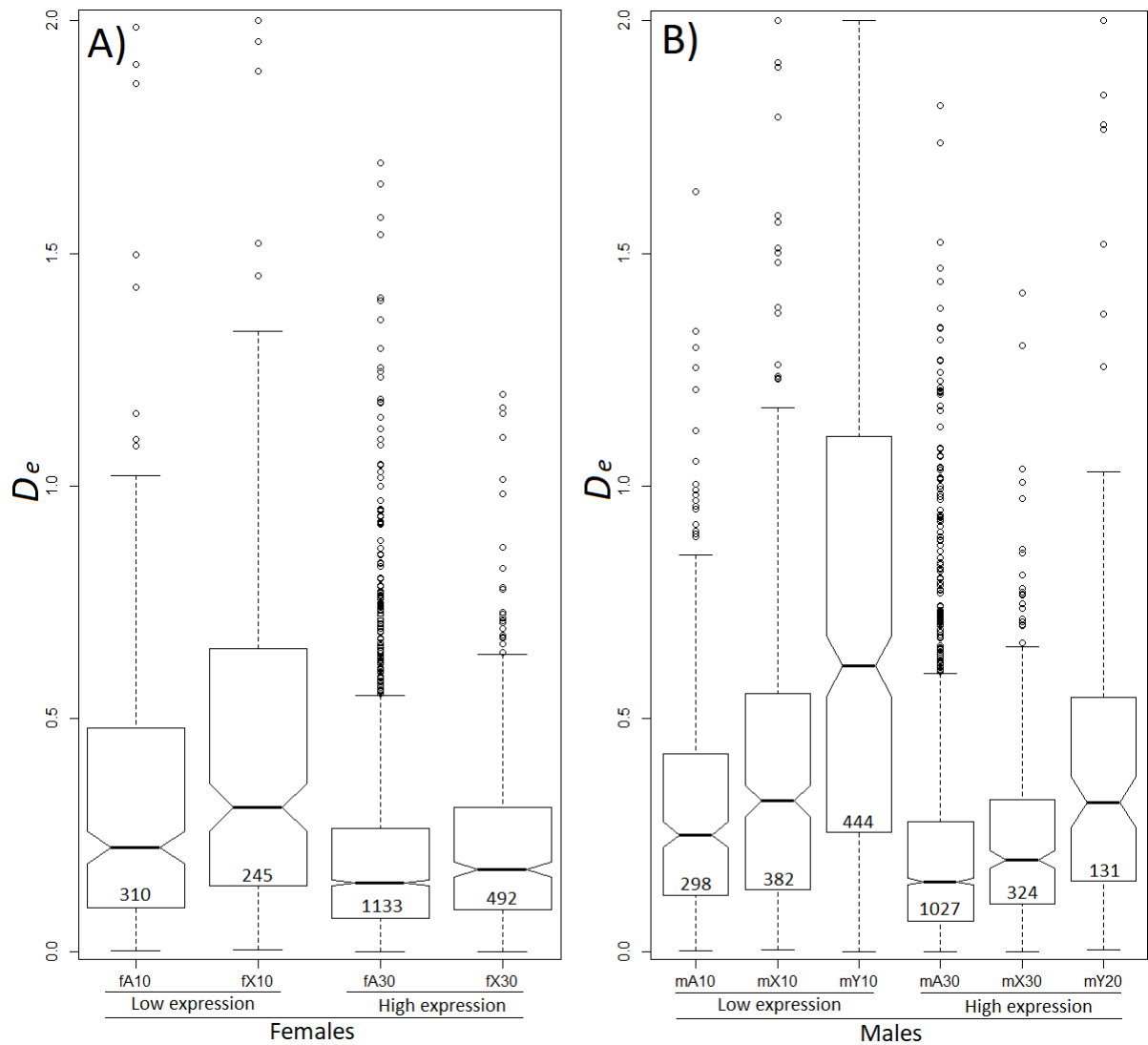


Figure S2. Expression of 46 and 42 X-linked genes that lost expression of the Y-linked gametologs ( $X_{noYexpr}$ ) in *S. latifolia* (A) or *S. dioica* (B) males, respectively. Expression of  $X_{noYexpr}$  genes in males is consistently higher in the species where the Y-copy is already degenerate.

675

676



677

678

679

680

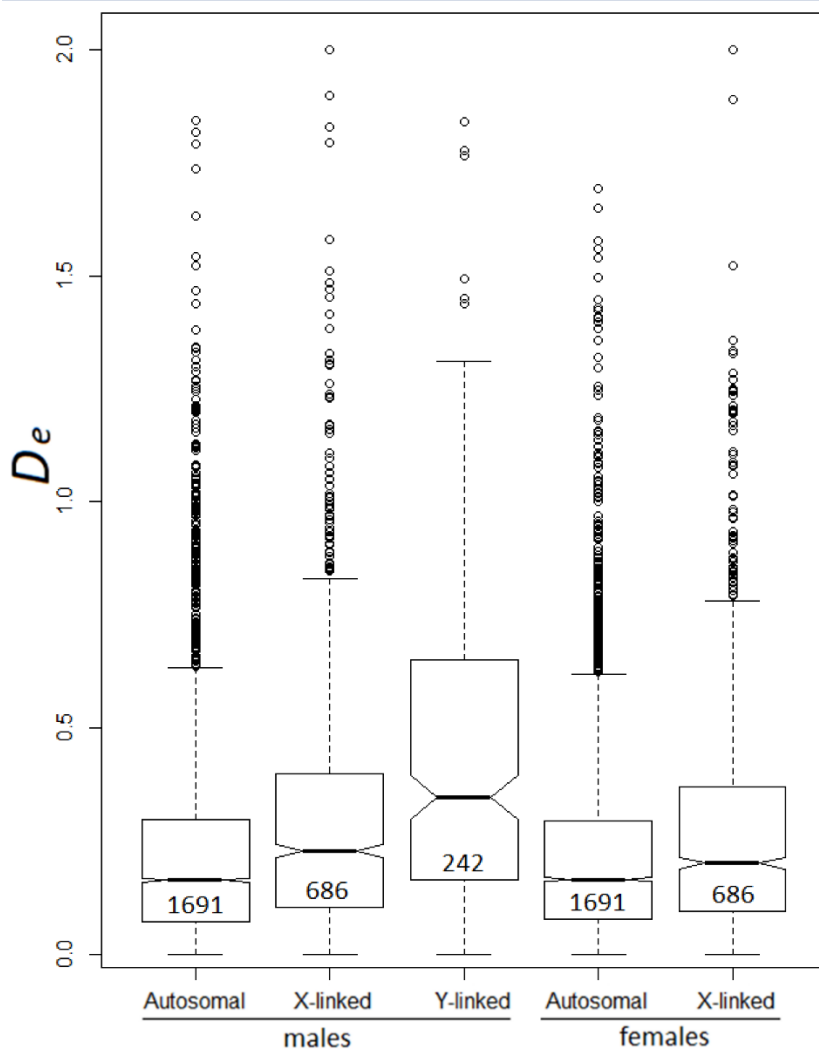
681

682

683

684

Figure S3. A faster expression divergence ( $D_e$ ) of the sex-linked genes was detectable regardless of the gene expression level. “Low expression” bars show expression divergence between *S. latifolia* and *S. dioica* for genes with  $1 < \text{FPKM} < 10$ , while the “high expression” bars show expression divergence for highly expressed genes ( $\text{FPKM} > 30$  for autosomal and X-linked genes;  $\text{FPKM} > 20$  for Y-linked genes). A lower FPKM threshold for highly expressed Y-linked genes was chosen to avoid sampling too few genes in this class. The numbers of genes analysed are shown inside the boxes.



685

686

687

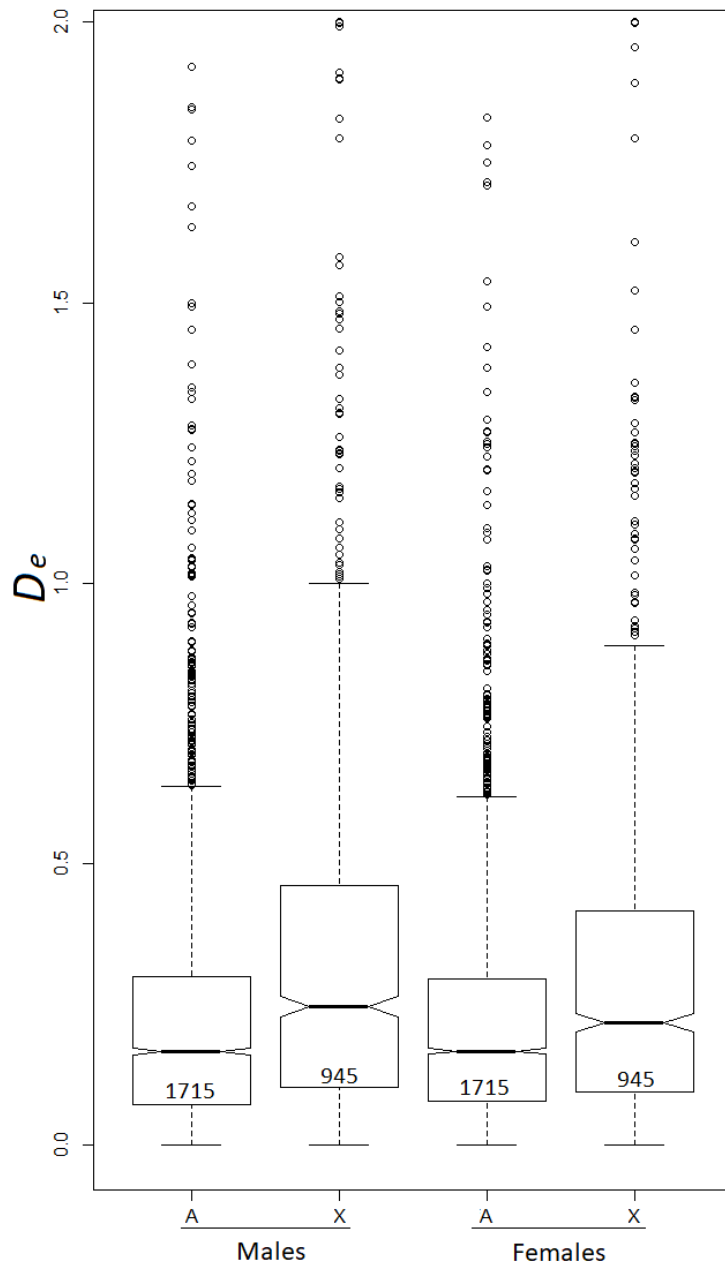
688

689

690

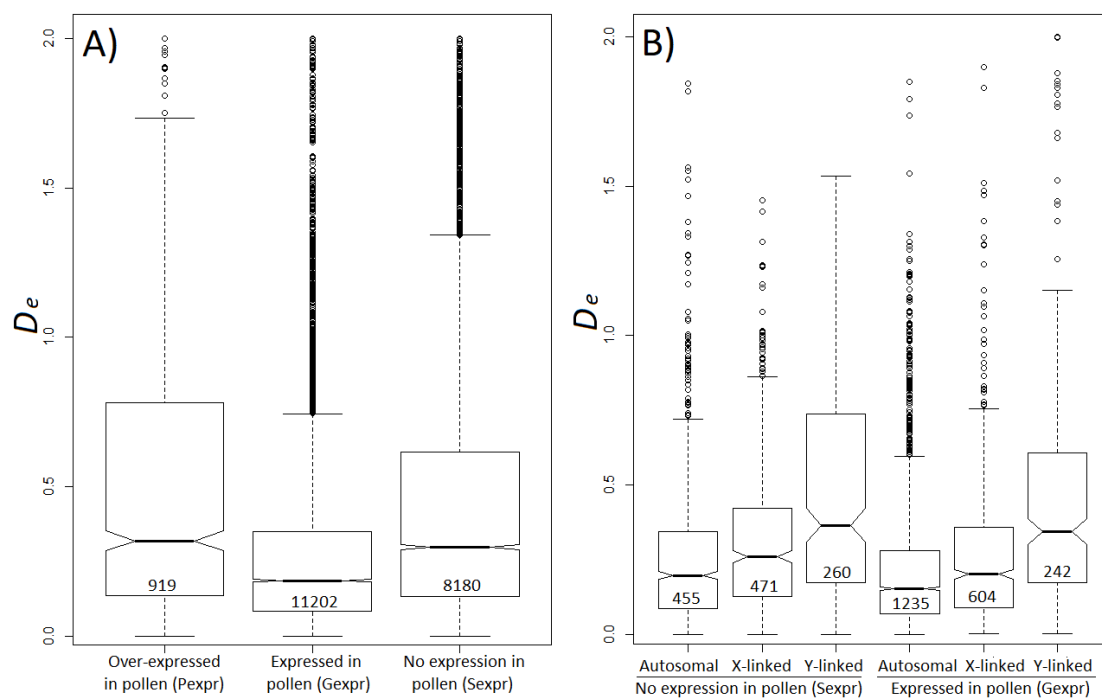
Figure S4. A faster expression divergence ( $D_e$ ) of the sex-linked genes was detectable after exclusion of weakly expressed genes (f<sub>pkm</sub><1). The numbers of genes analysed are shown inside the boxes.

691  
692



693  
694  
695  
696  
697  
698  
699

Figure S5. Gene expression divergence ( $D_e$ ) between *S. latifolia* and *S. dioica* for X-linked (X) and autosomal (A) genes after exclusion of sex-biased genes. Y-linked genes are not shown because by definition all Y-linked genes are male-specific. The numbers of genes analysed are shown inside the boxes.



700

701

702

703

704

705

706

707

Figure S6. Gene expression divergence between *S. latifolia* and *S. dioica* males for genes predominantly expressed in the sporophyte ( $S_{expr}$ ) and the gametophyte ( $G_{expr}$  and  $P_{expr}$ ). A) All genes regardless of linkage; B) A comparison of autosomal and sex-linked genes. The numbers of genes analysed are shown inside the boxes. Weakly expressed genes ( $f_{pkm} < 1$ ) were excluded from analysis.