bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The effects of haploid selection on Y chromosome evolution in a dioecious plant

George Sandler*, Felix E.G. Beaudry, Spencer C.H. Barrett, and Stephen I. Wright

Department of Ecology and Evolutionary Biology, University of Toronto

25 Willcocks St., Toronto, ON M5S 3B2, Canada

Short running title: Haploid selection and sex chromosome evolution

Keyword: Sex chromosomes, recombination, ploidy

* Author for correspondence: george.sandler@mail.utoronto.ca

bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The evolution of sex chromosomes is classically considered to be driven by sexually antagonistic selection. However, selection during the haploid gametic phase of the lifecycle has recently received theoretical attention as possibly playing a central role in sex chromosome evolution, especially in plants. In particular, selection for reduced recombination on the sex chromosomes may occur as a result of intense haploid selection in males favouring the linkage of haploid beneficial alleles to an incipient Y chromosome. Here, we examine the evolution of gene expression in flower buds and pollen of two species of *Rumex* to test for signatures of sexual antagonism and haploid selection acting during sex chromosome evolution. We find that genes with high ancestral pollen expression bias occur more often on sex chromosomes than autosomes, and that genes on the Y chromosome are more likely to become enriched for pollen expression bias. We also find that genes with low expression in pollen are more likely to be lost from the Y chromosome. We found no comparable pattern for gene expression in male flower bud tissue suggesting that sexual antagonism among diploid parents may be a less important force in shaping Y chromosome evolution in *Rumex*. Our results suggest that selection during the haploid gametophytic stage of the lifecycle may be a major contributor to plant sex chromosome evolution.

Abstract word count: 220

Total word count: 4084

bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

IMPACT SUMMARY

Sex chromosome evolution is most commonly explained as resulting from the resolution of sexual antagonism in the genome. Evidence for this hypothesis, however, is fairly limited. Recently, an alternative hypothesis has been proposed which posits that sex chromosomes evolve to link haploid beneficial mutations to the chromosome which experiences the most selection during the haploid phase. In plants, linkage of such genes to the Y chromosome may allow alleles beneficial in the haploid competitive arena to consistently segregate into males where pollen competition occurs. Here, we analyse gene expression data from three tissues of two species of plants in *Rumex*. We demonstrate that the sex chromosomes in these species are enriched for pollen-expressed genes, that the genes have become more pollen biased in expression, and that Y-linked genes are overexpressed in pollen. We found no comparable pattern in flower bud tissue where we expect sexual antagonism to be acting in the diploid phase. Our results support previous findings in *Silene* that haploid selection contributes to the retention of genes on the Y chromosome, but also provides novel empirical evidence for adaptive specialization of Y linked genes to the haploid phase of the plant life cycle.

Dioecy, the occurrence of populations with separate male and female individuals, is relatively uncommon in flowering plants, occurring in only 5-6% of species (Renner 2014). Of the 175 families in which dioecy is reported, only 19 species from four families are known to possess heteromorphic sex chromosomes (Charlesworth 2002; Ming et al. 2011; Bachtrog et al. 2014). Since the development

bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

of modern visualization and genomic techniques, the number of species with sex chromosomes has increased, but most recently identified sex chromosome systems in flowering plants are homomorphic (Harkess et al. 2015; Pucholt et al. 2017). This rarity of sex chromosomes in angiosperms raises the question of whether they may evolve in fundamentally different ways in plants than in animals, and has spurred speculation about the fundamental differences between these groups that may cause this difference in abundance.

Theory suggests sex chromosomes evolve from a pair of autosomes that acquired a sex-determining region, and subsequently accumulate sexually antagonistic alleles (Charlesworth 1996). The loss of genetic recombination between the sex chromosomes is thought to be selected to assure the segregation of sexually antagonistic alleles into the sex in which they are beneficial (Rice 1984; Lenormand 2003). Though widely accepted, evidence supporting sex-specific selection is limited to a few systems (Foerster et al. 2007; Delph et al. 2010; Innocenti and Morrow 2010) and has rarely been conclusively related to sex chromosome evolution (but see Wright et al. 2017).

One key contrast between flowering plants and animals that could contribute to differences in the evolution of their sex chromosomes is the predominance of a haploid gametophytic phase in the life cycle of plants (Haldane 1933). Genes expressed in the haploid phase should be subject to a unique selection regime because any recessive alleles in the diploid phase are unmasked in the haploid phase. Haploid selection is thought to be pervasive in angiosperms; for example, in the hermaphroditic plant *Arabidopsis thaliana* 60-70% of all genes are expressed in pollen (Honys and Twell 2004; Borges et al. 2008). The haploid selective regime

bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

includes more efficient removal of deleterious mutations from a population as recessive deleterious phenotypic effects are expressed (Gerstein and Otto 2009). Similarly, recessive beneficial mutations are more likely to spread through populations. Indeed, in *Capsella grandiflora,* pollen-expressed genes experience stronger purifying and positive selection relative to non-pollen expressed genes (Arunkumar et al. 2013). Pollen competition is generally considered to be a common feature of angiosperms, increasing selective pressures imposed on plant genomes in the haploid phase (Moore and Pannell 2011).

Gene expression in pollen may contribute to the evolution of heteromorphic sex chromosomes in three complementary ways (Fig. 1). First, pollen-specific selection can favour the loss of recombination between the X and Y chromosomes, because linkage of alleles beneficial during pollen competition to the Y chromosome enables these alleles to spend more time in males where competition occurs (Scott and Otto 2017). Hereafter, we refer to this phenomenon as adaptive linkage (Fig. 1a). Second, once genes have become sex-linked, greater haploid selection in males may cause divergence and upregulation of Y-linked alleles specialized for pollen, akin to masculinization of the Y (Lahn and Page 1997; Zhou and Bachtrog 2012) and feminization of the X (Prince et al. 2010; Allen et al. 2013; Albritton et al. 2014) observed in some animal systems (hereafter “pollenization”) (Fig. 1b). Finally, haploid expression of genes on the Y can cause biased retention of pollen-expressed genes from the degenerating Y, as has been reported in *Silene latifolia* (Chibalina and Filatov 2011) (Fig. 2c).The study of pollen expression in young plant sex chromosome systems provides opportunities for the untangling of these three processes.

bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

124 *Rumex* (Polygonaceae) provides a valuable study system to investigate the
125 effects of haploid selection and sexual antagonism on sex chromosome evolution.
126 Members of this genus possess uniovulate flowers, and open-pollinated flowers
127 capture numerous pollen grains increasing the scope for pollen competition and
128 haploid selection (Stehlik et al. 2008). Here, we sequence RNA from flower buds and
129 mature pollen of annual, wind-pollinated *R. hastatulus* and *R. rothschildianus* to
130 investigate patterns of gene expression and test the predictions of models of sex
131 chromosome evolution driven by sexually antagonistic and haploid selection. Both
132 species possess heteromorphic sex chromosomes, with non-orthologous sex-linked
133 genes, consistent with independently-evolved sex chromosomes (Crowson et al.
134 2017). The species also exhibit different degrees of Y-chromosome degeneration
135 (Crowson et al. 2017). In *R. hastatulus* approximately three-quarters of genes have
136 retained an expressed Y copy, whereas in *R. rothschildianus* only ~10% have been
137 retained (Hough et al. 2014; Crowson et al. 2017). We find evidence for an increased
138 rate of loss of genes with very low or no pollen expression from the Y chromosome.
139 However, we also find evidence that genes on the sex chromosomes are ancestrally
140 more pollen biased in their expression, and that this bias has increased since the loss
141 of recombination. Finally, our results indicate that Y-linked alleles are overexpressed
142 relative to their X-linked counterparts in pollen but not in flower buds, a pattern
143 consistent with the hypothesis that haploid selection drives the cessation of
144 recombination on *Rumex* Y chromosomes.
145
146
147

bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

METHODS

Tissue collection

We used gene expression data from three tissues: pollen (a target for haploid selection), male flower bud (a target for sexual antagonism) and male leaf (control). We collected mature pollen and filtered it through a fine nylon mesh before RNA extraction. We pooled pollen from two *R. rothschildianus* males due to low tissue yields in this species but collected pollen from two male plants of *R. hastatulus* individually. We collected developing, unopened male flower buds from two *R. rothschildianus* individuals (sampled and sequenced independently) and one *R. hastatulus* individual. We performed all RNA extractions using Spectrum™ Plant Total RNA kits and stored RNA at -80ºC. Leaf expression data from three *R. hastatulus* and three *R. rothschildianus* males were obtained from previous work (see Hough et al. 2014; Crowson et al. 2017).

RNAseq and read analysis

We sequenced RNA samples using Illumina Hi-seq 2500 sequencing with 100bp paired end reads at the Centre for Applied Genomics, Toronto. We aligned samples to existing female leaf transcriptome assemblies from both species (Hough et al. 2014; Crowson et al. 2017). We performed alignments using STAR (Dobin et al. 2013) after which we removed duplicate reads using Picard (http://broadinstitute.github.io/picard). We used SAMtools to retrieve read counts for downstream differential expression analysis (Li et al. 2009). We performed differential expression analysis using the R package DESeq2 (Love et al. 2014) using read

bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

counts obtained from SAMtools. We used >0.3 FPKM as a cut-off for active transcription, as recommended in (Ramsköld et al. 2009).

Ortholog comparison

We compared expression of genes that were retained on both the X and Y (hereafter XY genes) and their orthologs in *R. hastatulus* and *R. rothschildianus* to infer changes in expression of genes that had become sex linked since the species diverged. This is possible because the sex-linked genes in the two species of *Rumex* arose independently (Crowson et al. 2017). We obtained lists of orthologs between the species from Crowson et al. (2017). Because we were interested in the relative expression bias of XY linked genes, and the normalization of expression may be influenced by the proportion and number of sex-linked genes in the genome, we corrected for this by normalizing expression bias of XY genes by the average expression bias of autosomes. The method for calculating the average autosomal expression bias is given by:

$$\bar{x} = 2^{\overline{\left\{log_2 \frac{\#normalised\ reads\ tissue\ 1}{\#normalised\ reads\ tissue\ 2}\right\}}}$$

After we calculated the average autosomal expression bias, we corrected individual XY gene bias using equation 2:

$$log_2 \left\{ \frac{\#\ normalised\ reads\ tissue\ 1 \Big/ \#\ normalised\ reads\ tissue\ 2}{\bar{x}} \right\}$$

bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

193 **SNP calling and analysis of allele-specific expression**

194 We used the HaplotypeCaller tool in GATK (McKenna et al. 2010) to call SNP
195 variants in our transcriptomes followed by the SelectVariants tool in GATK to select
196 only biallelic, sex- linked SNPs. This list of SNPs was then run through the GATK
197 ASEReadcounter (McKenna et al. 2010) tool with default settings to allow for
198 estimation of allele-specific expression on the sex chromosomes. As all male tissue
199 transcriptomes were aligned to female references, most alternate SNPs in XY linked
200 genes represent the Y copy of a gene whereas most reference SNPs represent the X
201 copy. However, this assumption can be violated if a polymorphism exists on the X. To
202 account for this, we used population data from *R. hastatulus* to identify fixed SNP
203 differences between the X and Y chromosomes (Hough et al. 2017). Fixed
204 differences were determined if all females in a population were homozygous at an X
205 linked SNP whereas all males were heterozygous for the same site. This yielded a
206 list of high confidence sites for use in allele-specific expression analysis. No such
207 population data was available for *R. rothschildianus* so instead we used SNPs with
208 sex-linked segregation patterns within a family of sequenced plants (Crowson et al.
209 2017).

210 The lists of high confidence sites for allele-specific expression analysis were
211 then run through the tool GeneiASE (Edsgärd et al. 2016) to find statistically
212 significant cases of allelic bias in expression. Since GeneiASE does not require
213 phasing, the output identifies which genes have a bias in expression but not the
214 direction of this bias. We averaged the log 2-fold expression change of each high
215 confidence XY site to infer whether the reference (X) or alternate (Y) copy was
216 overexpressed.

bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Statistical analysis

All *P*-values reported in this study are two tailed. We used Fisher's exact tests to test for significant differences between counts of differentially expressed genes. Student's *t*-tests were used to compare mean FPKM pollen expression of hemizygous and XY linked genes. We conducted Wilcoxon signed-rank tests to compare tissue expression bias of XY and ortholog genes. Multiple test correction is applied by both DESeq2 and GeneiASE using the Benjamin-Hochberg method, only genes with an adjusted *P*-value of <0.05 were considered as differentially expressed (for DESeq2), or as having significant allele biased expression (for GeneiASE). We used Fisher's combined probability tests to combine the individual gene *P*-values using GeneiASE for each tissue sample to yield one *P*-value per gene. We then used Fisher's exact tests to test for differences in numbers of X vs. Y biased genes in our allele-specific expression analysis. We did not test for differences in numbers of genes without allele-specific expression, because the fraction of non-biased genes heavily depends on factors such as sequencing depth, which can vary from sample to sample. We used counts of overexpressed genes in leaf as the null expected counts in our statistical tests, as we did not expect either haploid selection or sexual antagonism to be driving Y chromosome evolution in leaf tissue.

Result and Discussion

Gene expression is widespread in pollen

In *R. hastatulus,* 39.1% of all predicted leaf transcripts had signatures of active pollen transcription, whereas 50.9% of leaf transcripts in *R. rothschildianus* showed

bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

241 evidence of pollen expression. These results suggest widespread gene expression in
242 the haploid phase. Nevertheless, the values we obtained were significantly lower
243 than in flower bud tissue, where 81.7% and 87.3% of predicted leaf transcripts were
244 actively transcribed in *R. hastatulus* and *R. rothschildianus,* respectively. Despite
245 overlap between tissues in genes with active expression, principle components
246 analysis (Supplementary Fig. 1) of individual samples indicated strong differentiation
247 between tissues in expression, particularly for pollen.
248
249 **Sex chromosomes are enriched for haploid expressed genes**
250 To investigate the relative importance of sexual antagonism and haploid selection
251 during sex chromosome evolution, we compared sex-linked and autosomal genes for
252 expression bias across three tissues. We focused on expression differences between
253 male leaf, male developing flower bud and mature pollen. The chromosomal location
254 of genes was previously evaluated using SNP segregation patterns for both *R.*
255 *hastatulus* (Hough et al. 2014) and *R. rothschildianus* (Crowson et al. 2017). We
256 compared counts of genes identified as having significantly different expression
257 (Benjamin-Hochberg FDR adjusted $P < 0.05$) between two pairs of tissues to quantify
258 expression bias. When comparing leaf and pollen expression patterns (Fig. 2a), we
259 found that sex-linked genes with retained Y copies (hereafter XY genes) were more
260 often significantly pollen biased than autosomal genes in both *R. hastatulus* (Fisher's
261 exact test, $P < 0.0001$) and *R. rothschildianus* (Fisher's exact test, $P< 0.0001$). The
262 same was true when comparing pollen to flower bud expression in *R. hastatulus*
263 (Fisher's exact test, $P = 0.0006$) and *R. rothschildianus* (Fisher's exact test, $P <$
264 0.0001) (Fig. 2b). Given this evidence for enrichment of pollen-biased genes on the

bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

sex chromosomes, we next sought to evaluate the possible role of different evolutionary mechanisms driving this enrichment (Fig. 1).

Haploid selection maintains pollen-expressed genes on the Y chromosome

To test whether haploid expression of Y-linked genes slows down Y chromosome degeneration in *Rumex* (Fig. 1c), we compared the pollen expression of hemizygous genes (which lack a Y-expressed copy) and XY genes (Chibalina and Filatov 2011; Crowson et al. 2017). We found that hemizygous genes showed significantly reduced pollen expression compared with XY genes in *R. hastatulus* (Welch Two Sample *t*-test, t = -6.7295, df = 154.31, P = 3.145e-10) and *R. rothschildianus* (Welch Two Sample *t*-test, t = -14.019, df = 552.18, P = 2.2e-16) (Supplementary Fig. 2). This effect is particularly prominent in the more degenerated *R. rothschildianus*: the effect size (Cohen's D) in *R. hastatulus* is 0.74, whereas it is 1.16 in *R. rothschildianus*. This difference was not simply due to hemizygous genes being generally less expressed (Crowson et al. 2017); differential expression analyses indicated that hemizygous genes had a deficiency of pollen-biased relative to either leaf- or flower-biased genes compared with XY genes in both *R. hastatulus* (Fisher's exact test, P < 0.0001 leaf/pollen; P < 0.0252 flower bud/pollen) and *R. rothschildianus* (Fisher's exact test, P < 0.0001 leaf/pollen; P < 0.0001 flower bud/pollen) (Fig. 2). Again, this effect was more pronounced in *R. rothschildianus.* Our results suggest that haploid selection retains pollen expressed genes, as also reported in *Silene latifolia*, another XY plant sex chromosome system (Chibalina and Filatov 2011). It is interesting to note that the difference in pollen expression between XY and hemizygous genes has diverged to a greater extent in *R. rothschildianus* suggesting that haploid selection

bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

289 does indeed slow down Y chromosome degeneration even in highly heteromorphic
290 plant sex chromosome systems.
291
292 **Sex chromosome linked genes show signals of ancestral and derived pollen**
293 **bias**
294 We next investigated whether the footprint of differential Y chromosome degeneration
295 could fully account for the patterns of pollen bias at XY genes, without needing to
296 invoke adaptive evolution of sex linkage or Y chromosome pollenization. To do this,
297 we examined the extent of pollen bias on all sex chromosome-linked genes,
298 combining both hemizygous and XY genes. By combining these gene sets, our
299 analysis should more closely resemble the ancestral set of genes that evolved to
300 become linked to the sex-determining region. The combined data still showed an
301 enrichment of pollen-biased genes across all sex-linked genes analysed together for
302 both *R. hastatulus* (Fisher's exact test, $P < 0.0001$, $P = 0.0015$; pollen/leaf and
303 pollen/flower bud respectively) and *R. rothschildianus* (Fisher's exact test, $P =$
304 0.0005, $P = 0.0314$; pollen/leaf and pollen/flower bud, respectively) (Supplementary
305 Fig. 3). However, there exists an ascertainment bias in this reconstructed gene set
306 as more SNP segregation patterns can be used to identify XY genes compared to
307 hemizygous genes (in particular, divergent SNPs between X and Y chromosomes)
308 resulting in overrepresentation of XY genes on reconstructed XY chromosomes
309 (Hough et al. 2014; Crowson et al. 2017). To account for this bias, we used existing
310 published XY gene lists (Hough et al. 2014; Crowson et al. 2017) which only
311 contained XY genes identified with the same set of SNP segregation patterns
312 (polymorphisms on the X chromosome) as hemizygous genes. This procedure

bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

313 therefore removed any ascertainment bias. We still found evidence for a significant
314 enrichment of pollen-biased genes in *R. hastatulus* (Fisher's exact test, $P < 0.0001$
315 for both tissue comparisons), but not in *R. rothschildianus*, where sex-linked genes
316 as a whole were significantly depleted for pollen enrichment (Fisher's exact test, $P =$
317 0.0224, $P = 0.0074$; pollen/leaf and pollen/flower bud, respectively). Thus, overall the
318 enrichment of pollen-expressed genes does not appear to be a simple function of the
319 Y degeneration of genes not expressed in pollen. However, any signal of early
320 enrichment may be eroded by extensive Y degeneration in the more degenerated *R.*
321 *rothschildianus* and possibly secondary movement of genes on and off the X
322 chromosome.

323 Because Y-chromosome degeneration alone is not sufficient to explain the
324 enrichment of pollen expression on XY linked genes, we further searched for
325 evidence of Y chromosome pollenization (Fig. 1b), or adaptive sex linkage (Fig. 1a).
326 The sex-linked genes in *R. rothschildianus* and *R. hastatulus* have arisen
327 independently (Crowson et al. 2017), thus ancestral expression of XY linked genes in
328 one species should be represented by the autosomal orthologs of these genes in the
329 other species. Therefore, we next attempted to disentangle the ancestral and
330 subsequent evolution of expression bias in sex-linked genes.

331 We first investigated whether orthologs of XY-linked genes were ancestrally
332 more pollen biased than other autosomal genes to determine whether pollen bias is
333 present before linkage to the sex chromosomes, indicative of adaptive sex linkage.
334 We found that *R. hastatulus* XY-linked genes were ancestrally more pollen biased
335 than other autosomal genes in a comparison between leaf and pollen (Fisher's exact
336 test, $P = 0.0344$) (Supplementary Fig. 4); although a similar trend was evident

bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

337 comparing pollen and flower buds, the difference was not significant (Fisher's exact
338 test, $P = 0.1197$). Similarly, *R. rothschildianus* XY-linked genes were ancestrally
339 more pollen biased than other autosomal genes in both tissue comparisons (Fisher's
340 exact test $P < 0.0001$, $P = 0.0002$; pollen/leaf and pollen/flower bud respectively).
341 These findings suggest that *R. hastatulus* XY ancestors were mildly pollen biased
342 and *R. rothschildianus* XY ancestors were highly pollen biased, a difference that may
343 be explained by the large difference in levels of Y chromosome degeneration
344 between the species. In particular, because the relatively intact Y chromosomes of *R.*
345 *hastatulus* still contain a considerable number of genes not expressed in the haploid
346 phase, the bias towards pollen expression should be less severe. Overall, our results
347 suggest that pollen-biased genes may have been involved early in the evolution of
348 sex chromosomes, suggesting adaptive linkage of haploid beneficial alleles.

349 To determine whether XY-linked gene ancestors became further pollenized
350 after linkage to the sex chromosomes, we performed reciprocal pairwise comparisons
351 of expression patterns of XY genes and their non-XY linked orthologs. Direct
352 comparisons between these gene sets are difficult to interpret due to apparent
353 genome-wide divergence in expression patterns between *R. hastatulus and R.*
354 *rothschildianus* (Fig. 2). To account for this divergence, we normalized the
355 expression bias of XY and orthologous genes by the average expression bias of their
356 respective autosomal genes to uncover relative differences in expression between
357 species (for details see methods).

358 We found that XY-linked genes in *R. hastatulus* were significantly more pollen
359 biased than their autosomal orthologs in *R. rothschildianus* in both tissue
360 comparisons (Wilcoxon's sign rank test $P < 0.01$ both tissue comparisons, $Z = -13.58$

bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

pollen/leaf, Z = -3.941 pollen/flower bud) (Fig. 3). The same was true for *R. rothschildianus* when comparing pollen and flower buds (Wilcoxon's sign rank test $P = 0.04$, Z = -2.022) but not when comparing pollen and leaf where XY genes appeared to be more pollen biased ancestrally (Wilcoxon's sign rank test $P = 0.01$, Z = -6.418). Our results support the hypothesis that XY-linked genes play an important role in the haploid gametophytic phase, particularly for *R. hastatulus,* by becoming enriched for pollen expression following sex linkage. We posit that the lack of pollenization observed when comparing leaf and pollen in *R. rothschildianus* may be related to the highly degenerate nature of the Y chromosomes in this species, where long periods of inefficient selection may have eroded signatures of pollenization on this chromosome.

Widespread Y overexpression is present specifically in pollen

If pollen overexpression on sex chromosomes is due to adaptive linkage on the Y chromosome (Fig. 1a), we would predict that it is driven by upregulation of Y-linked genes expressed in pollen. To test this hypothesis, we examined allele-specific gene expression on the sex chromosomes across several tissues. We predicted Y-bias in pollen if adaptive linkage contributes to the formation of plant sex chromosomes, and Y-bias in male flower buds if sexual antagonism is the dominant force driving sex chromosome evolution. In contrast, we predicted minimal Y-bias in leaf, or reduced expression on the Y due to degeneration (Hough et al. 2014; Crowson et al. 2017), and therefore this tissue was used as a control.

We found no consistent chromosomal bias for allelic overexpression in leaf (14.6% X overexpressed, 16.4% Y overexpressed) and flower bud (7.7% X

bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

overexpressed, 7.0% Y overexpressed) of *R. hastatulus*. There was also no significant difference in the pattern of allelic overexpression between leaf and flower bud (Fisher's exact test $P = 0.7083$) (Fig. 4). In pollen, however, 44.9% of XY genes exhibited Y overexpression whereas only 16.4% had X overexpression, indicating that XY genes had significantly more Y overexpression in pollen than leaf (Fisher's exact test $P = 0.0021$).

In *R. rothschildianus* we found overall more X allele-biased genes in flower bud (67.5% X overexpression, 18.2% Y overexpression) and leaf (64.4% X overexpression, 24.3% Y overexpression), with no significant difference in the direction of allele-specific expression between these tissues (Fisher's exact test, $P = 0.1167$). Once again, we found significantly more Y-biased expression in pollen (47.8% Y overexpressed, 35.5% X overexpressed) relative to leaf (Fisher's exact test, $P < 0.0001$).

The occurrence of widespread Y-overexpression in pollen of both species is consistent with the hypothesis that haploid gametophytic selection plays a significant role in the evolution of sex chromosomes (adaptive linkage) and suggests that Y-linked alleles are preferentially upregulated in the haploid phase, for which they have been optimised. The overall prevalence of X-overexpression in *R. rothschildianus* confirms previous findings that X-linked alleles appear to be more highly expressed when their Y-linked orthologs accumulate deleterious mutations due to inefficient selection (Crowson et al. 2017). Given that the sex chromosomes of *R. rothschildianus* are far more degenerate than those of *R. hastatulus* it is expected that this pattern of widespread X overexpression is more prominent in *R. rothschildianus*. We interpret the lack of widespread Y-overexpression in male flower

bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

409 buds as an indicator that sexual antagonism in flower buds contributes less to the
410 evolution of sex chromosomes in the two *Rumex* species than haploid selection.
411 Alternatively, it could be that sexual antagonism is resolved through mechanisms that
412 do not leave a signature in allelic expression patterns. However, this may be unlikely
413 given that sexual antagonism has previously been linked with allele-specific
414 expression changes in plants (Zemp et al. 2016).

415 Given the evidence for pollen specialization on the Y chromosome, the early
416 stages of sex chromosome evolution in *Rumex* may have been driven by haploid
417 selection. It has been proposed that this should lead to male-biased sex ratios in
418 populations (Scott and Otto 2017); however, contemporary populations of several
419 *Rumex* species, including both species studied here typically show female biased sex
420 ratios (Putwain and Harper 1972; Zarzycki and Rychlewski 1972; Klimes 1993;
421 Rottenberg 1998; Stehlik and Barrett 2005; Pickup and Barrett 2013), suggesting
422 superior pollen competitive ability of the X chromosome. This discrepancy might be
423 explained by subsequent Y-chromosome degeneration driven by inefficient linked
424 selection, despite the residual signal of pollen specialization on the Y (Scott and Otto
425 2017). If plant Y chromosomes do indeed contain many genes important in pollen
426 competition, Y degeneration should then be particularly harmful for the competitive
427 ability of Y-bearing pollen leading to the observed female biased sex ratios.
428

429 **Conclusion**

430 We report evidence that differential retention of pollen-expressed genes during
431 degeneration, pollenization upon divergence, and adaptive linkage of pollen-
432 expressed genes each contribute to the enrichment of Y chromosomes for pollen

bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

433 expressed genes. As previously reported in *Silene latifolia* (Chibalina and Filatov
434 2011), haploid selection can slow the degeneration of Y chromosomes despite the
435 reduced efficacy of selection predicted to be associated with the loss of
436 recombination (Charlesworth 1996). The slow pace of degeneration of Y
437 chromosomes as a result of haploid selection may help to explain the overall rarity of
438 heteromorphic sex chromosomes in plants. Furthermore, our results provide
439 evidence for the influence of haploid selection during the early stages of sex
440 chromosome evolution. Similar to the increased sex-specific expression on animal
441 sex chromosome, we find sex-specific transmission produces unique conditions
442 whereby the Y chromosome can specialize for pollen competition. These results
443 support the theoretical prediction of Scott and Otto (2017) that haploid selection can
444 contribute to the arrest in recombination between nascent sex chromosomes.
445

446 **Acknowledgements**

447 The authors would like to thank Daisy Crowson and Tia Harrison for help with RNA
448 extraction from pollen and flower buds. This research was funded by Natural
449 Sciences and Engineering Research Council of Canada Discovery grants (to
450 S.C.H.B. and S.I.W.) and an EWR Steacie fellowship (to S.I.W.).
451

452 **Author Contributions**

453 SIW, SCHB and GS conceived of and designed the study, GS collected the data, GS
454 and FEGB analysed the data, all authors contributed to the writing of the paper.
455
456

bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Conflict of interest

The authors declare no conflict of interest.

Literature Cited

Albritton, S. E., A.-L. Kranz, P. Rao, M. Kramer, C. Dieterich, and S. Ercan. 2014. Sex-biased gene expression and evolution of the X chromosome in nematodes. *Genetics* **197**:865–883.

Allen, S. L., R. Bonduriansky, and S. F. Chenoweth. 2013. The genomic distribution of sex-biased genes in *Drosophila serrata*: X chromosome demasculinization, feminization, and hyperexpression in both sexes. *Genome Biol. Evol.* **5**:1986–1994.

Arunkumar, R., E. B. Josephs, R. J. Williamson, and S. I. Wright. 2013. Pollen-specific, but not sperm-specific, genes show stronger purifying selection and higher rates of positive selection than sporophytic genes in *Capsella grandiflora*. *Mol. Biol. Evol*. **30**:2475–2486.

Borges, F., G. Gomes, R. Gardner, N. Moreno, S. McCormick, J. A. Feijó, and J. D. Becker. 2008. Comparative transcriptomics of *Arabidopsis* sperm cells. *Plant Physiol*. **148**:1168–1181.

Charlesworth, B. 1996. The evolution of chromosomal sex determination and dosage compensation. *Curr. Biol*. **6**:149–162.

Charlesworth, D. 2002. Plant sex determination and sex chromosomes. *Heredity* **88**:94–101.

Chibalina, M. V., and D. A. Filatov. 2011. Plant Y chromosome degeneration is retarded by haploid purifying selection. *Curr. Biol.* **21**:1475–1479.

bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

481 Crowson, D., S. C. H. Barrett, and S. I. Wright. 2017. Purifying and positive selection
482 influence patterns of gene loss and gene expression in the evolution of a plant
483 sex chromosome system. *Mol. Biol. Evol*. **34**:1140–1154.
484 Delph, L. F., A. M. Arntz, C. Scotti-Saintagne, and I. Scotti. 2010. The genomic
485 architecture of sexual dimorphism in the dioecious plant *Silene latifolia*.
486 *Evolution* **64**:2873–2886.
487 Dobin, A., C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M.
488 Chaisson, and T. R. Gingeras. 2013. STAR: ultrafast universal RNA-seq
489 aligner. *Bioinformatics* **29**:15–21.
490 Edsgärd, D., M. J. Iglesias, S.-J. Reilly, A. Hamsten, P. Tornvall, J. Odeberg, and O.
491 Emanuelsson. 2016. GeneiASE: Detection of condition-dependent and static
492 allele-specific expression from RNA-seq data without haplotype information.
493 *Sci. Rep.* **6**:21134.
494 Foerster, K., T. Coulson, B. C. Sheldon, J. M. Pemberton, T. H. Clutton-Brock, and L.
495 E. B. Kruuk. 2007. Sexually antagonistic genetic variation for fitness in red
496 deer. *Nature* **447**:1107–1110.
497 Gerstein, A. C., and S. P. Otto. 2009. Ploidy and the causes of genomic evolution. *J.*
498 *Hered.* **100**:571–581.
499 Haldane, J. B. S. 1933. The part played by recurrent mutation in evolution. *Am. Nat.*
500 **67**:5–19.
501 Harkess, A., F. Mercati, H.-Y. Shan, F. Sunseri, A. Falavigna, and J. Leebens-Mack.
502 2015. Sex-biased gene expression in dioecious garden asparagus (*Asparagus*
503 *officinalis*). *New Phytol*. **207**:883–892.

bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

504 Honys, D., and D. Twell. 2004. Transcriptome analysis of haploid male gametophyte
505 development in *Arabidopsis*. *Genome Biol*. **5**:R85.
506 Hough, J., J. D. Hollister, W. Wang, S. C. H. Barrett, and S. I. Wright. 2014. Genetic
507 degeneration of old and young Y chromosomes in the flowering plant *Rumex*
508 *hastatulus*. *Proc. Natl. Acad. Sci.* **111**:7713–7718.
509 Hough, J., W. Wang, S. C. H. Barrett, and S. I. Wright. 2017. Hill-Robertson
510 interference reduces genetic diversity on a young plant Y-chromosome.
511 *Genetics* **207**:685-695.
512 Innocenti, P., and E. H. Morrow. 2010. The sexually antagonistic genes of *Drosophila*
513 *melanogaster*. *PLOS Biol*. **8:**e1000335.
514 Klimes, L. 1993. Sex ratios in dioecious *Rumex tuberosus* populations along a
515 successional gradient. *Preslia* **65**:335–340.
516 Lahn, B. T., and D. C. Page. 1997. Functional coherence of the human Y
517 chromosome. *Science* **278**:675–680.
518 Lenormand, T. 2003. The evolution of sex dimorphism in recombination. *Genetics*
519 **163**:811–822.
520 Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G.
521 Abecasis, R. Durbin, and 1000 Genome Project Data Processing Subgroup.
522 2009. The sequence alignment/map format and SAMtools. *Bioinforma. Oxf.*
523 *Engl*. **25**:2078–2079.
524 Love, M. I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and
525 dispersion for RNA-seq data with DESeq2. *Genome Biol*. **15**:550.
526 McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K.
527 Garimella, D. Altshuler, S. Gabriel, M. Daly, and M. A. DePristo. 2010. The

bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

genome analysis toolkit: a mapreduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. **20**:1297–1303.

Moore, J. C., and J. R. Pannell. 2011. Sexual selection in plants. *Curr. Biol*. **21**:R176–R182.

Pickup, M., and S. C. H. Barrett. 2013. The influence of demography and local mating environment on sex ratios in a wind-pollinated dioecious plant. *Ecol. Evol*. **3**:629–639.

Prince, E. G., D. Kirkland, and J. P. Demuth. 2010. Hyperexpression of the X chromosome in both sexes results in extensive female bias of X-linked genes in the flour beetle. *Genome Biol. Evol*. **2**:336–346.

Pucholt, P., A. E. Wright, L. L. Conze, J. E. Mank, and S. Berlin. 2017. Recent sex chromosome divergence despite ancient dioecy in the willow *Salix viminalis*. *Mol. Biol. Evol*. **34**:1991–2001.

Putwain, P. D., and J. L. Harper. 1972. Studies in the dynamics of plant populations: V. Mechanisms governing the sex ratio in *Rumex acetosa* and *R. acetosella*. *J. Ecol*. **60**:113–129.

Ramsköld, D., E. T. Wang, C. B. Burge, and R. Sandberg. 2009. An abundance of ubiquitously expressed genes revealed by tissue transcriptome sequence data. *PLOS Comput. Biol*. **5**:e1000598.

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

Rice, W. R. 1984. Sex chromosomes and the evolution of sexual dimorphism. *Evol. Int. J. Org. Evol*. **38**:735–742.

bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

552 Rottenberg, A. 1998. Sex ratio and gender stability in the dioecious plants of Israel.
553 *Bot. J. Linn. Soc*. **128**:137–148.
554 Scott, M. F., and S. P. Otto. 2017. Haploid selection favors suppressed
555 recombination between sex chromosomes despite causing biased sex ratios.
556 *Genetics* **207**:1631–1649.
557 Stehlik, I., and S. C. H. Barrett. 2005. Mechanisms governing sex-ratio variation in
558 dioecious *Rumex nivalis*. *Evol. Int. J. Org. Evol*. **59**:814–825.
559 Stehlik, I., J. Friedman, and S. C. H. Barrett. 2008. Environmental influence on
560 primary sex ratio in a dioecious plant. *Proc. Natl. Acad. Sci*. **105**:10847–
561 10852.
562 Wright, A. E., I. Darolti, N. I. Bloch, V. Oostra, B. Sandkam, S. D. Buechel, N. Kolm,
563 F. Breden, B. Vicoso, and J. E. Mank. 2017. Convergent recombination
564 suppression suggests role of sexual selection in guppy sex chromosome
565 formation. *Nat. Commun*. **8**:14251.
566 Zarzycki, K., and J. Rychlewski. 1972. Sex ratios in polish natural populations and in
567 seedling samples of *Rumex acetosa L*. and *Rumex thyrsiflorus Fing*. *Acta Biol.*
568 *Cracoviensia Ser. Bot*. **15**:335–340.
569 Zemp, N., R. Tavares, A. Muyle, D. Charlesworth, G. A. B. Marais, and A. Widmer.
570 2016. Evolution of sex-biased gene expression in a dioecious plant. *Nat.*
571 *Plants* **2**:16168.
572 Zhou, Q., and D. Bachtrog. 2012. Sex-specific adaptation drives early sex
573 chromosome evolution in *Drosophila*. *Science* **337**:341–345.
574

bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Tables and Figures

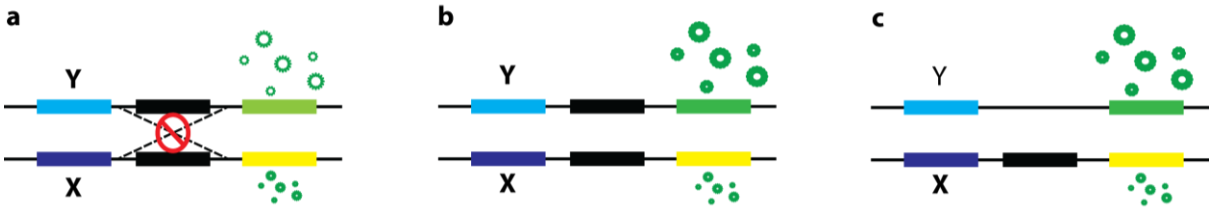
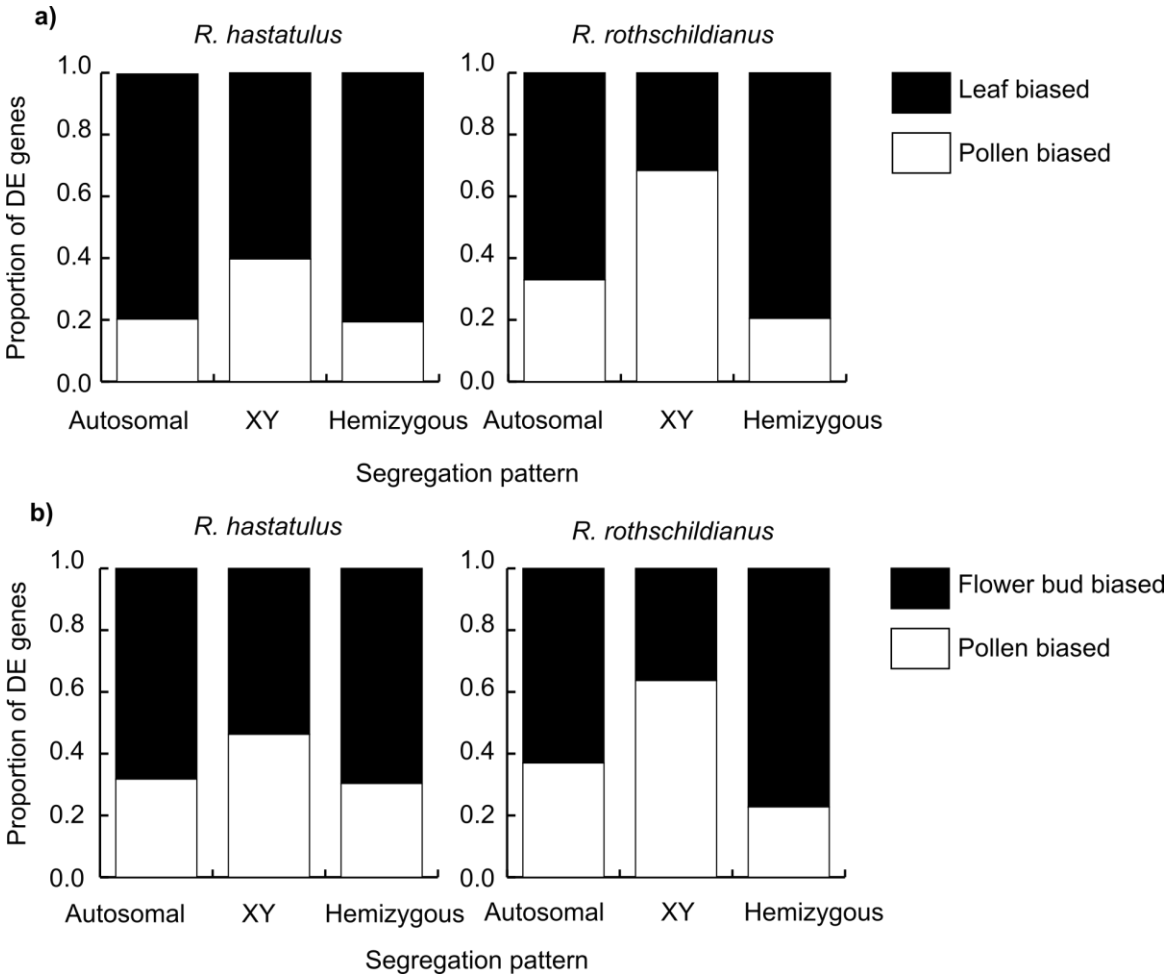


Fig. 1: Depiction of the effects of haploid gametophytic selection on sex chromosomes. Three distinct processes can potentially contribute to biased expression and overrepresentation of pollen genes on the Y chromosome. **a)** Recombination can be lost between the male determining region (Y) and any allele that increases pollen fitness (the green allele). **b)** Without recombination, alleles with pollen-specific fitness can diverge, or their expression can increase relative to the X-linked allele. **c)** As inefficient selection due to linkage causes degeneration of Y-linked alleles, haploid selection during pollen competition may cause biased retention of genes with pollen-specific fitness effects

bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

589



590

591

592 Fig. 2: Tissue expression bias of different gene groups in two *Rumex* species. Bar
593 segments represent the fraction of genes with significant differential-expression (DE)
594 in two pairwise tissue comparisons.

595

596

597

598

599

600

bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

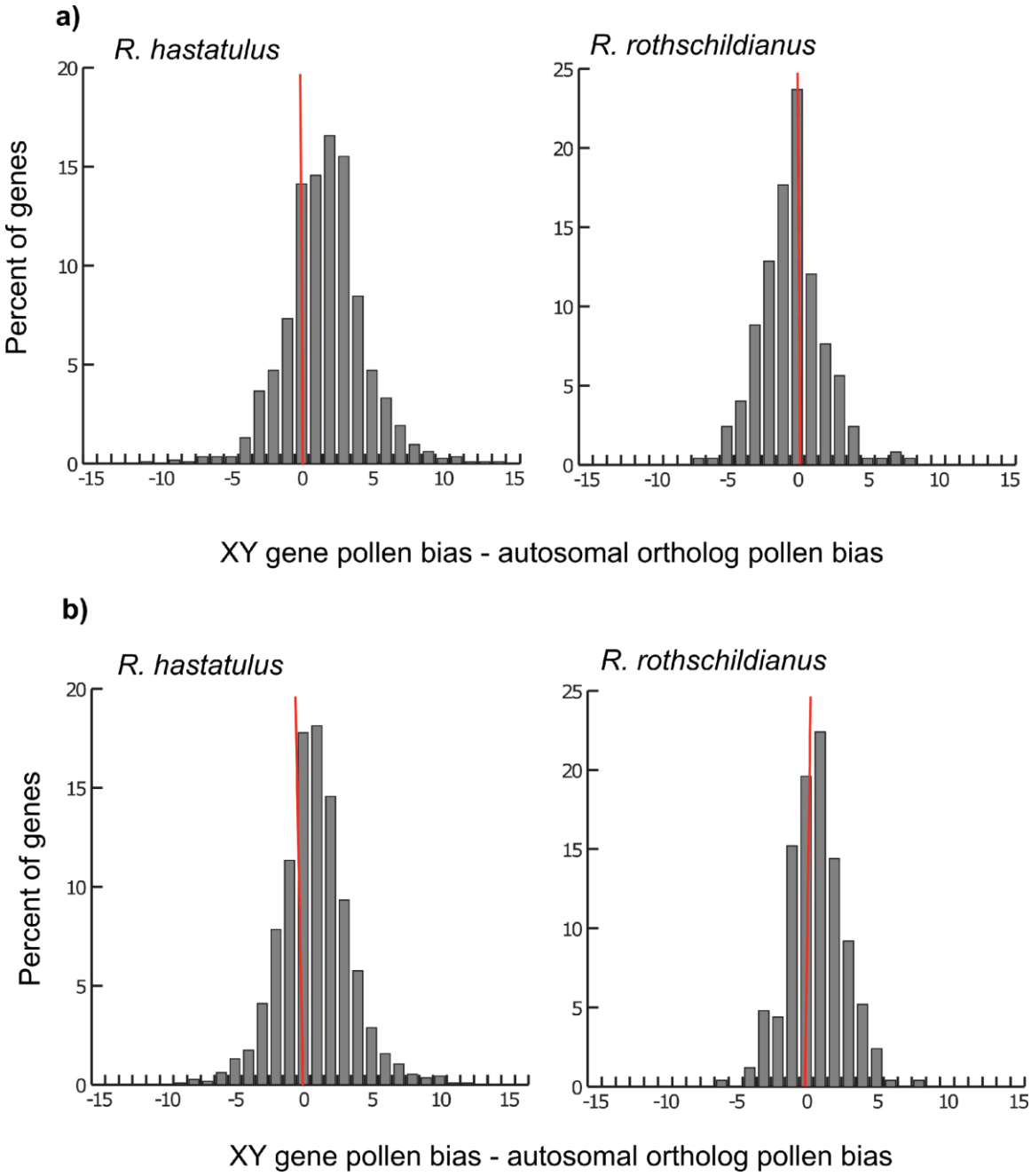
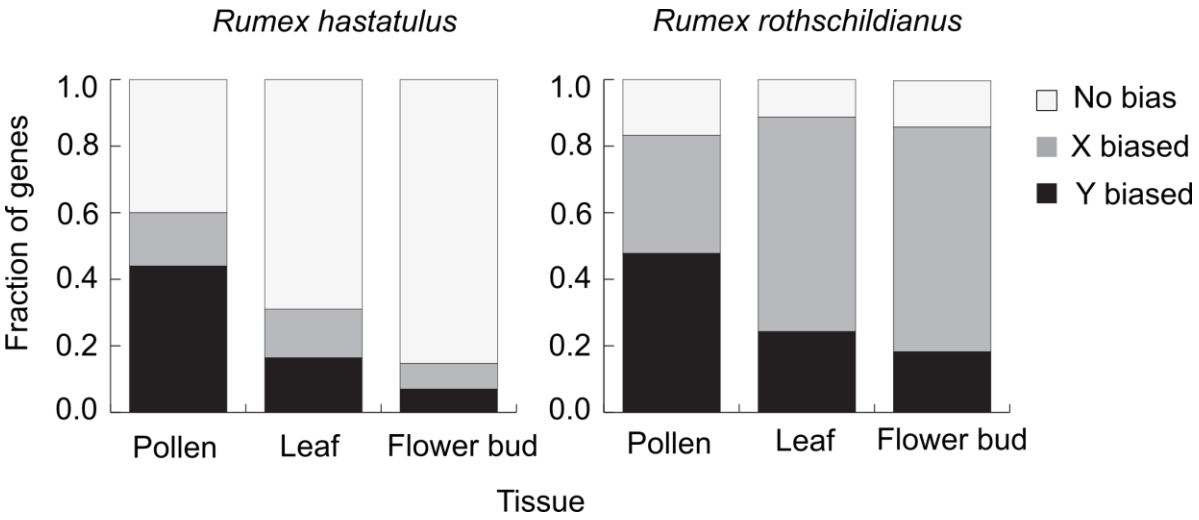


Fig. 3: Differences in normalized tissue expression bias of XY genes and their autosomal orthologs in two species of *Rumex*. The magnitude and direction of the differences are related to the evolution of tissue expression in XY genes after their linkage to the sex chromosomes. Tissue expression data used in the comparisons include a) leaf/pollen and b) flower bud/pollen. Positive values indicate greater pollen

bioRxiv preprint doi: https://doi.org/10.1101/264382; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

608 overexpression in XY genes relative to their autosomal orthologs. For details of

609 normalization see methods.



610

611 Fig. 4: Allele specific expression bias of X- and Y-linked genes in two species of

612 *Rumex.* Bar segments represent the percent of XY genes with no allelic bias (white),

613 significant X-overexpression (grey), and significant Y-overexpression (black).

614

615

616