The "two rules of speciation" in species with young sex chromosomes

Dmitry A. Filatov

Email for correspondence: Dmitry.Filatov@plants.ox.ac.uk

Address:
Department of Plant Sciences,
University of Oxford,
South Parks Road,
Oxford OX1 3RB
United Kingdom

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

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

Abstract

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

Introduction

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

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

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

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

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

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

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

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

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

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

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

Materials and Methods

Plant material

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

RNA extraction and sequencing

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

Expression analyses

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

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

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

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

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

Results

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

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

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

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

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

Y-degeneration accelerates the evolution of X-linked gene expression

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

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

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

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

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

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

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

Does haploid expression affect species divergence?

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

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

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

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

Discussion

Is the Silene Y-chromosome degenerate "enough"?

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

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

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

The theory expressing the conditions for Haldane's rule as a function of p_x, the proportion of hybrid incompatibilities that are X-linked hemizygous ((Turelli & Orr 1995) equation B2), shows that with a smaller p_x the conditions for HR evolution become more restrictive (e.g. see fig. 1 in (Orr & Turelli 1996)). The X-chromosome in *S. latifolia* and *S. dioica* contains about 10% of the Silene genome (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 ~ 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 ~ 0.05), but smaller than *Drosophila melanogaster*, where p_x ~
382 0.36 (Orr & Turelli 1996). As both *Drosophila* and mammals comply with HR, it appears that the
383 extent of hemizygosity 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 $X_{noYexpr}$ genes. This is consistent
433 with the observation of faster gene expression divergence at $X_{noYexpr}$ compared to X_{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_{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.

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

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

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

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

It is possible that the faster-X evolution in Silene has a composite nature, with accelerated positive 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

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

Acknowledgements

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

Data Accessibility

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

Author Contributions

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

References

Ahmed-Braimah YH, Unckless RL, Clark AG (2017) Evolutionary dynamics of male reproductive genes in the *Drosophila virilis* subgroup. *G3 (Bethesda)* **7**, 3145-3155.

Armstrong SJ, Filatov DA (2008) A cytogenetic view of sex chromosome evolution in plants. *Cytogenet Genome Res* **120**, 241-246.

Bachtrog D (2008) The temporal dynamics of processes underlying Y chromosome degeneration. *Genetics* **179**, 1513-1525.

Bachtrog D, Hom E, Wong KM, Maside X, de Jong P (2008) Genomic degradation of a young Y chromosome in *Drosophila miranda*. *Genome Biol* **9**, R30.

Bachtrog D, Mank JE, Peichel CL*, et al.* (2014) Sex determination: why so many ways of doing it? *PLoS Biol* **12**, e1001899.

Bellott DW, Hughes JF, Skaletsky H*, et al.* (2014) Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators. *Nature* **508**, 494-499.

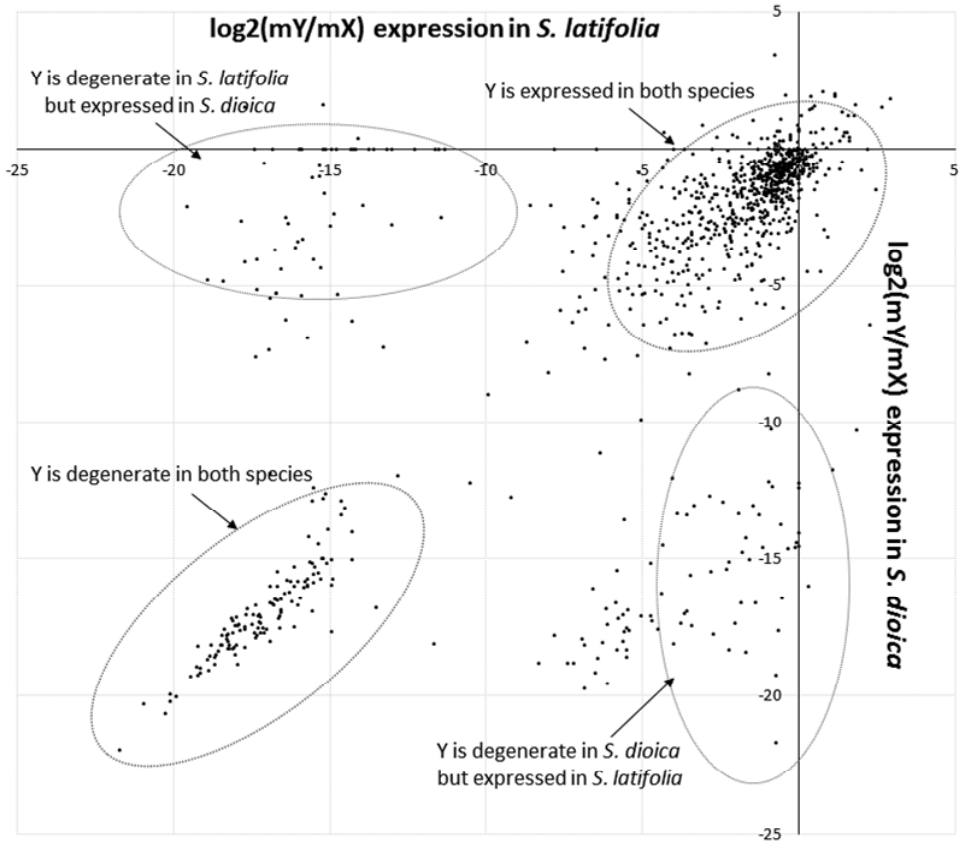
Bergero R, Charlesworth D (2011) Preservation of the Y transcriptome in a 10-million-year-old plant sex chromosome system. *Curr Biol* **21**, 1470-1474.

Bergero R, Qiu S, Charlesworth D (2015) Gene loss from a plant sex chromosome system. *Curr Biol* **25**, 1234-1240.

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

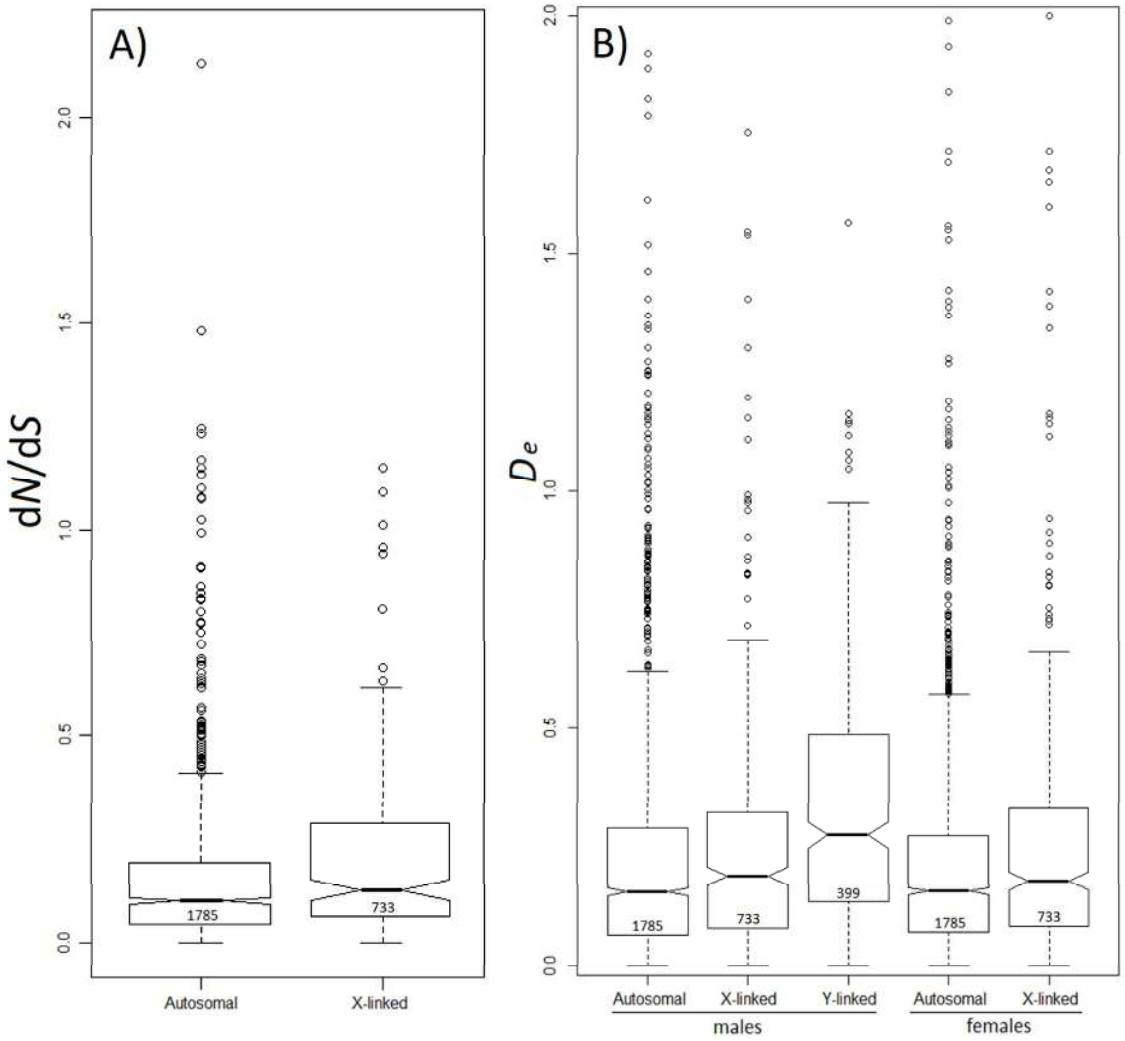


639

640 Figure 1. Loss of gene expression on the Y-chromosome. The plot shows log2-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 (d*N*) to synonymous (d*S*) 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

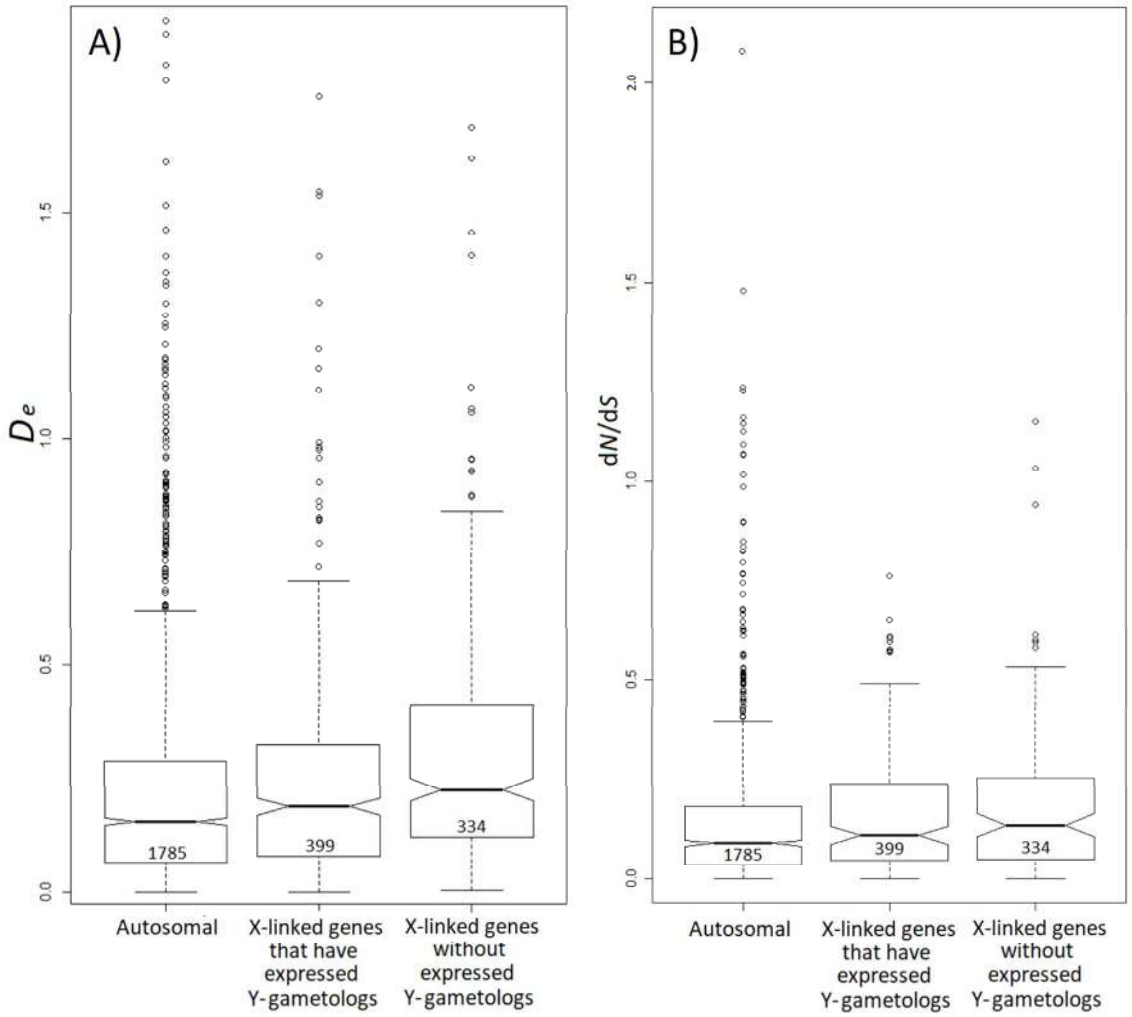


Figure 3. Faster-X evolution for gene expression divergence in males (A) and sequence divergence (d*N*/d*S*, panel B) between *S. latifolia* and *S. dioica*. X-linked genes are split into two categories for X-linked genes with expressed (X_{Yexpr}) and non-expressed ($X_{noYexpr}$) Y-linked gametologs. The numbers of genes analysed are shown inside the boxes.

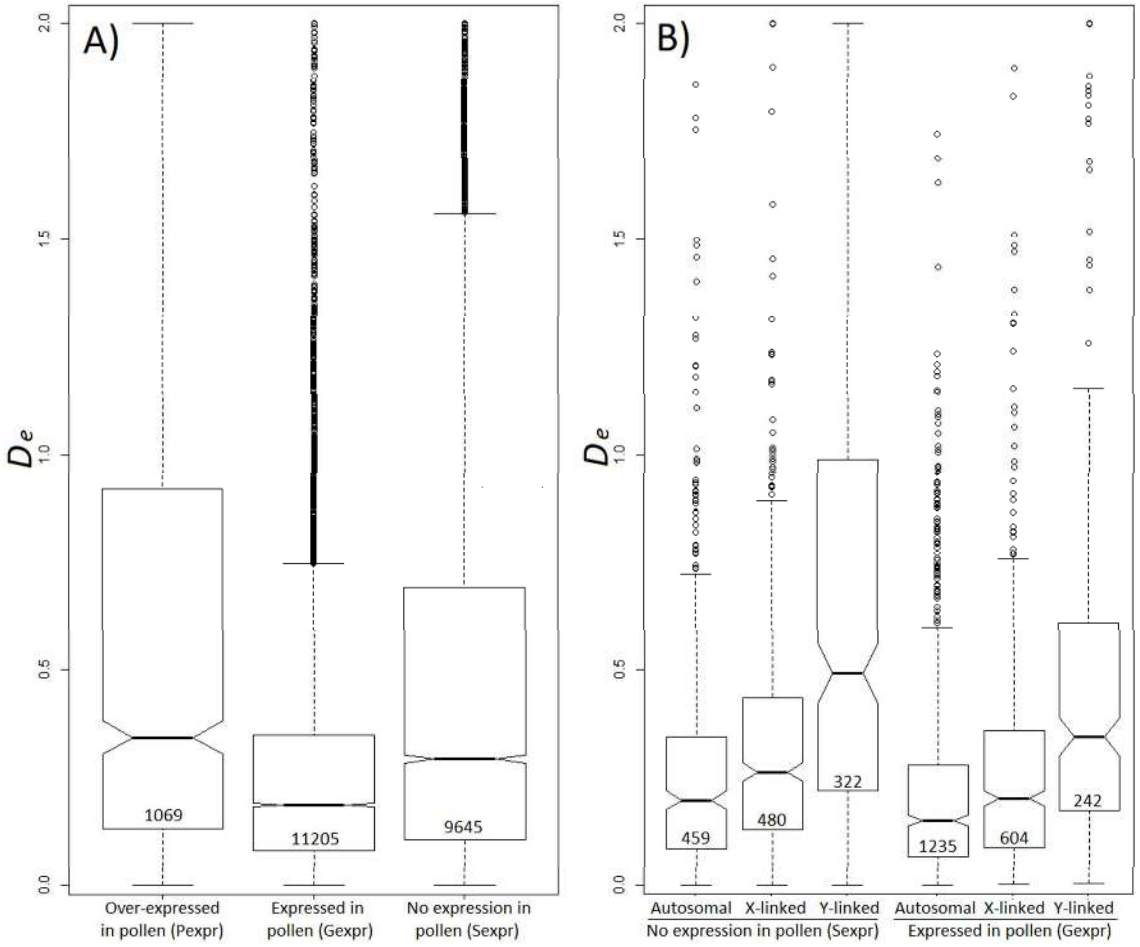


Figure 4. Gene expression divergence between *S. latifolia* and *S. dioica* males for genes predominantly expressed in the sporophyte (S_{expr}) and the gametophyte (G_{expr} and P_{expr}). A) All genes regardless of linkage; B) A comparison of autosomal and sex-linked genes. The numbers of genes analysed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

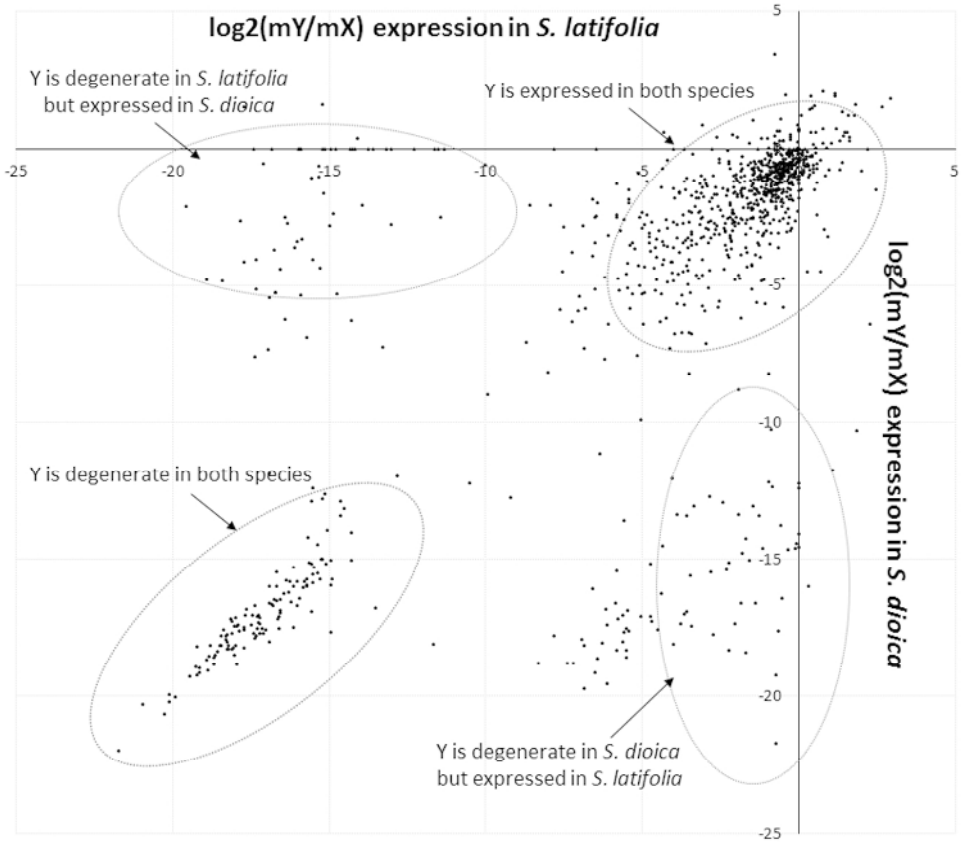


Figure 1

119x103mm (300 x 300 DPI)

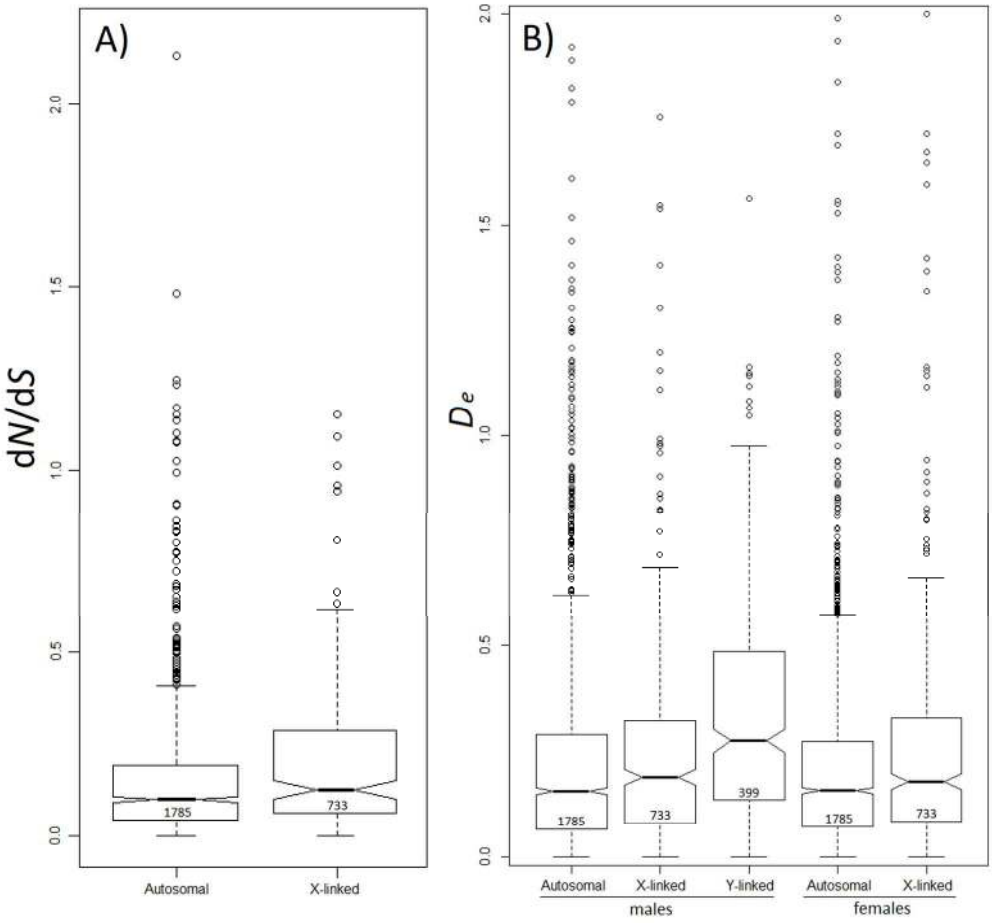


Figure 2

146x136mm (300 x 300 DPI)

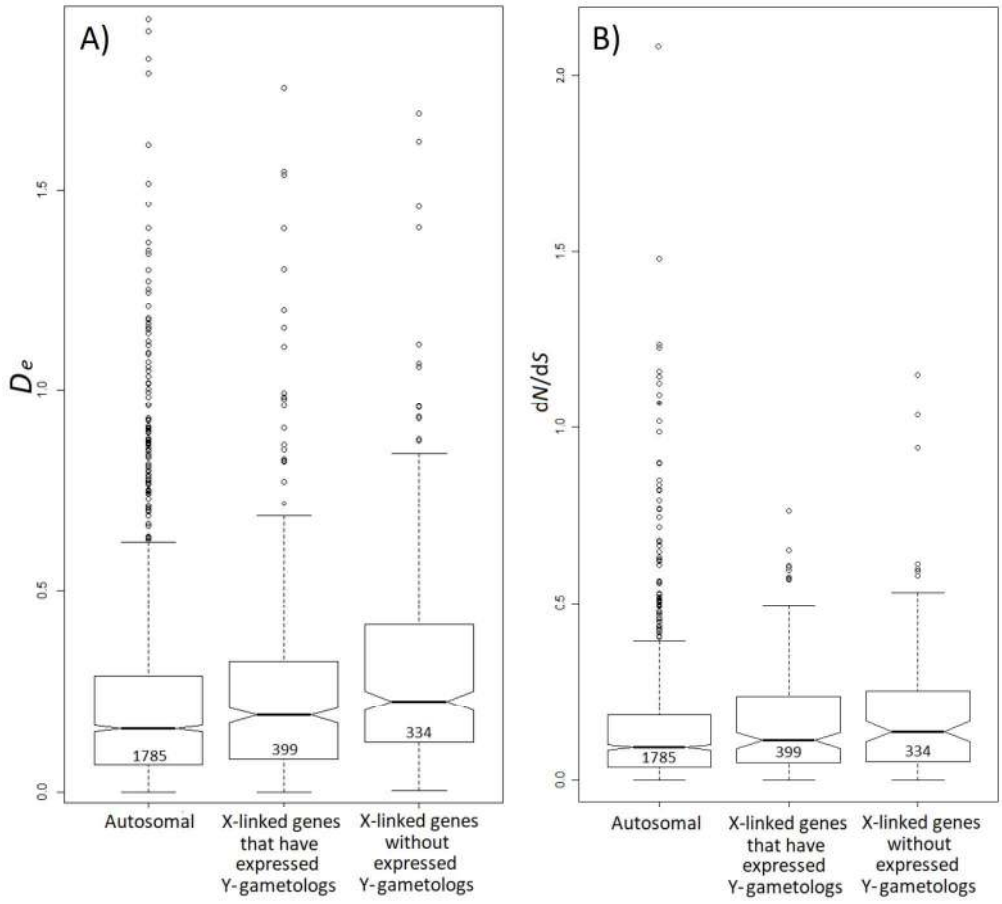


Figure 3

167x151mm (300 x 300 DPI)

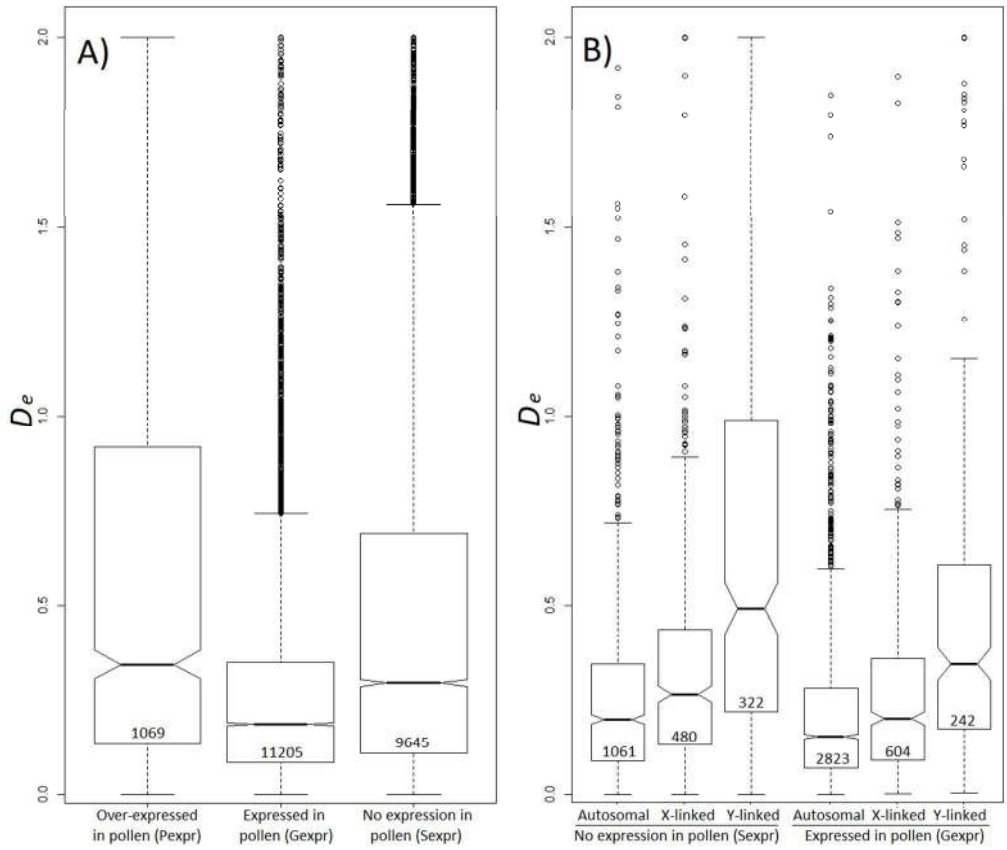


Figure 4

108x91mm (300 x 300 DPI)

The "two rules of speciation" in species with young sex chromosomes

Dmitry A. Filatov

Email for correspondence: Dmitry.Filatov@plants.ox.ac.uk

Address:
Department of Plant Sciences,
University of Oxford,
South Parks Road,
Oxford OX1 3RB
United Kingdom

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

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

Abstract

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

Introduction

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

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

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

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

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

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

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

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

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

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

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

Materials and Methods

Plant material

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

RNA extraction and sequencing

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

Expression analyses

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

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

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

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

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

Results

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

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

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

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

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

Y-degeneration accelerates the evolution of X-linked gene expression

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

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

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

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

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

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

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

Does haploid expression affect species divergence?

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

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

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

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

Discussion

Is the Silene Y-chromosome degenerate "enough"?

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

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

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

The theory expressing the conditions for Haldane's rule as a function of p_x, the proportion of hybrid incompatibilities that are X-linked hemizygous ((Turelli & Orr 1995) equation B2), shows that with a smaller p_x the conditions for HR evolution become more restrictive (e.g. see fig. 1 in (Orr & Turelli 1996)). The X-chromosome in *S. latifolia* and *S. dioica* contains about 10% of the Silene genome (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 ~ 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 ~ 0.05), but smaller than *Drosophila melanogaster*, where p_x ~
382 0.36 (Orr & Turelli 1996). As both *Drosophila* and mammals comply with HR, it appears that the
383 extent of hemizygosity 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

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

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

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

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

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

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

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

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

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

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

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

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

Conclusions

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

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

Acknowledgements

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

Data Accessibility

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

Author Contributions

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

References

Ahmed-Braimah YH, Unckless RL, Clark AG (2017) Evolutionary dynamics of male reproductive genes in the *Drosophila virilis* subgroup. *G3 (Bethesda)* **7**, 3145-3155.

Armstrong SJ, Filatov DA (2008) A cytogenetic view of sex chromosome evolution in plants. *Cytogenet Genome Res* **120**, 241-246.

Bachtrog D (2008) The temporal dynamics of processes underlying Y chromosome degeneration. *Genetics* **179**, 1513-1525.

Bachtrog D, Hom E, Wong KM, Maside X, de Jong P (2008) Genomic degradation of a young Y chromosome in *Drosophila miranda*. *Genome Biol* **9**, R30.

Bachtrog D, Mank JE, Peichel CL*, et al.* (2014) Sex determination: why so many ways of doing it? *PLoS Biol* **12**, e1001899.

Bellott DW, Hughes JF, Skaletsky H*, et al.* (2014) Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators. *Nature* **508**, 494-499.

Bergero R, Charlesworth D (2011) Preservation of the Y transcriptome in a 10-million-year-old plant sex chromosome system. *Curr Biol* **21**, 1470-1474.

Bergero R, Qiu S, Charlesworth D (2015) Gene loss from a plant sex chromosome system. *Curr Biol* **25**, 1234-1240.

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

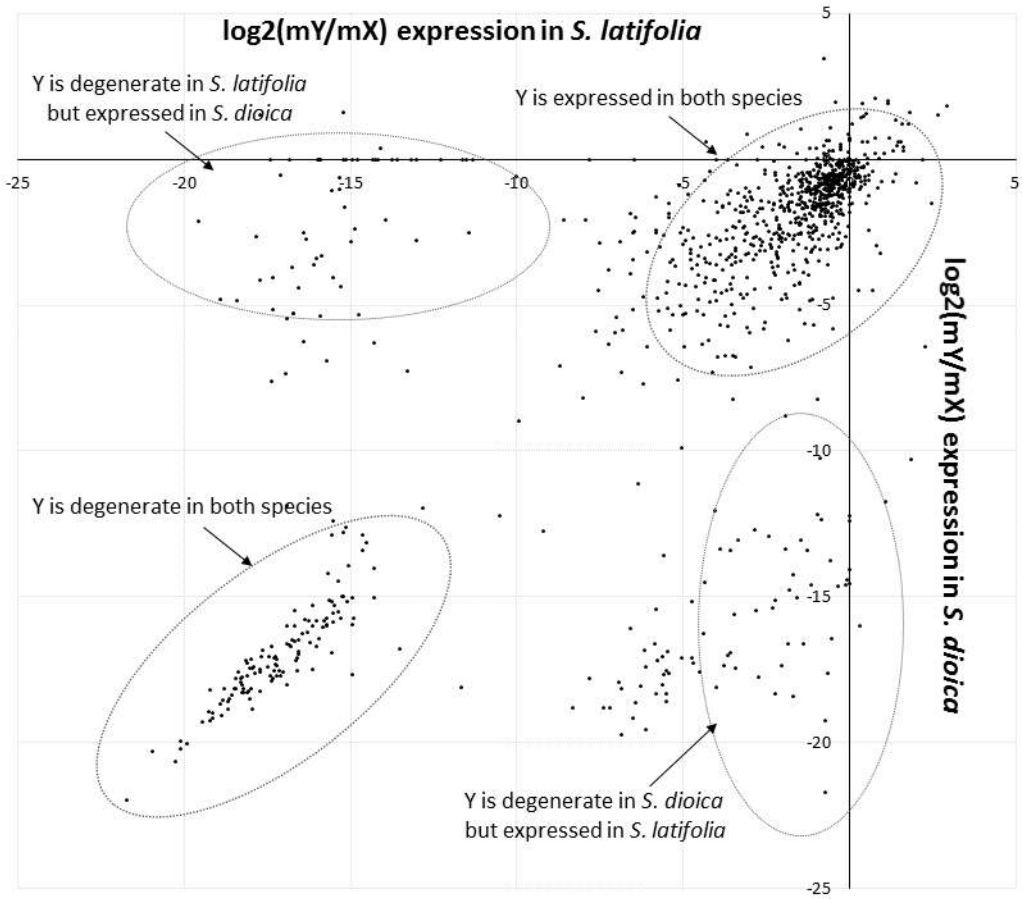
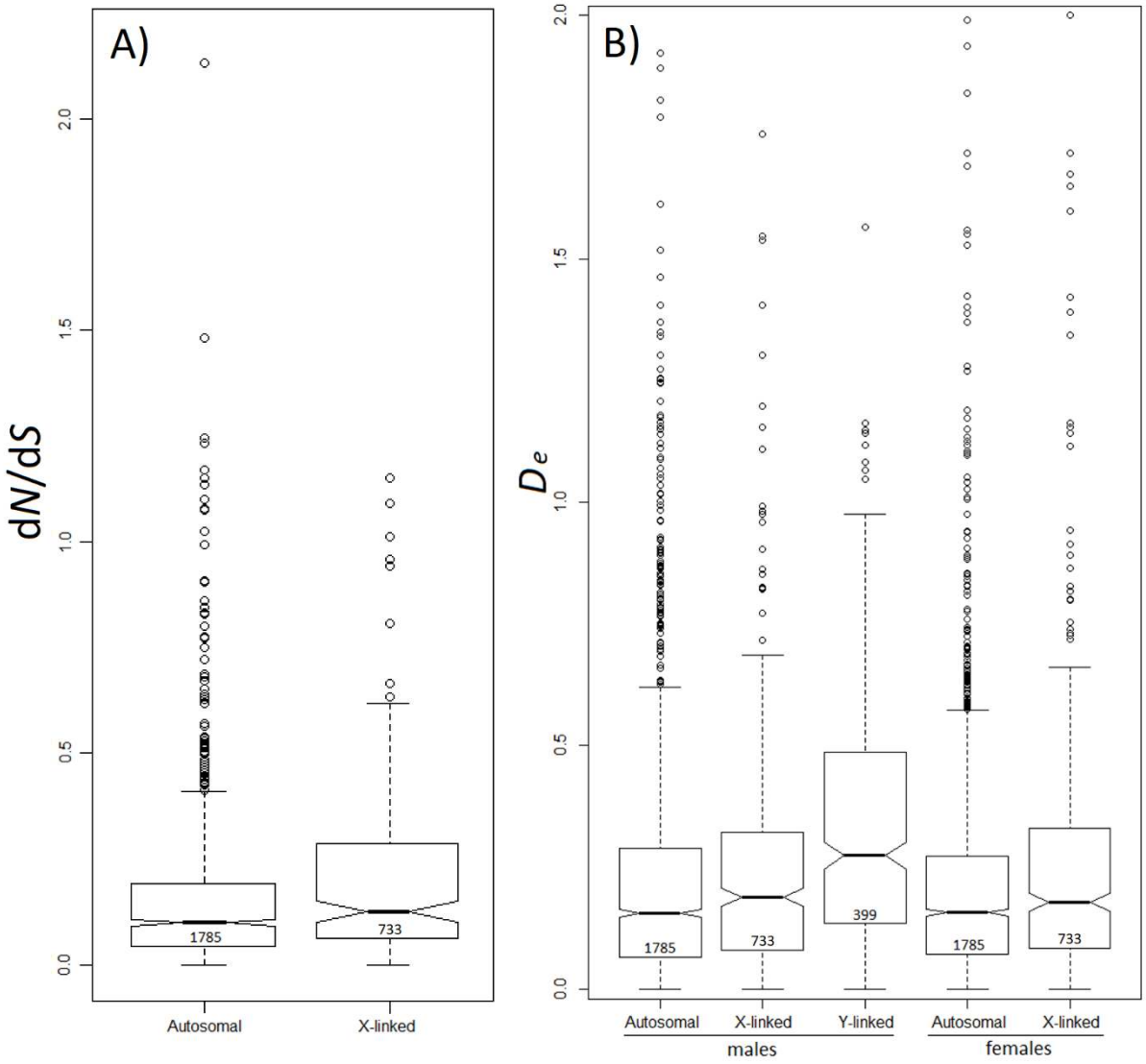


Figure 1. Loss of gene expression on the Y-chromosome. The plot shows log2-transformed ratios of relative Y/X gene expression of individual sex-linked genes in *S. latifolia* (X-axis) and *S. dioica* (Y-axis) males. To avoid infinity and division by zero errors for non-expressed genes, a small number (0.0001) was added to all expression values. The ellipses were drawn by hand to highlight different parts of the plot.

644



645

646 Figure 2. Non-synonymous (d*N*) to synonymous (d*S*) 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

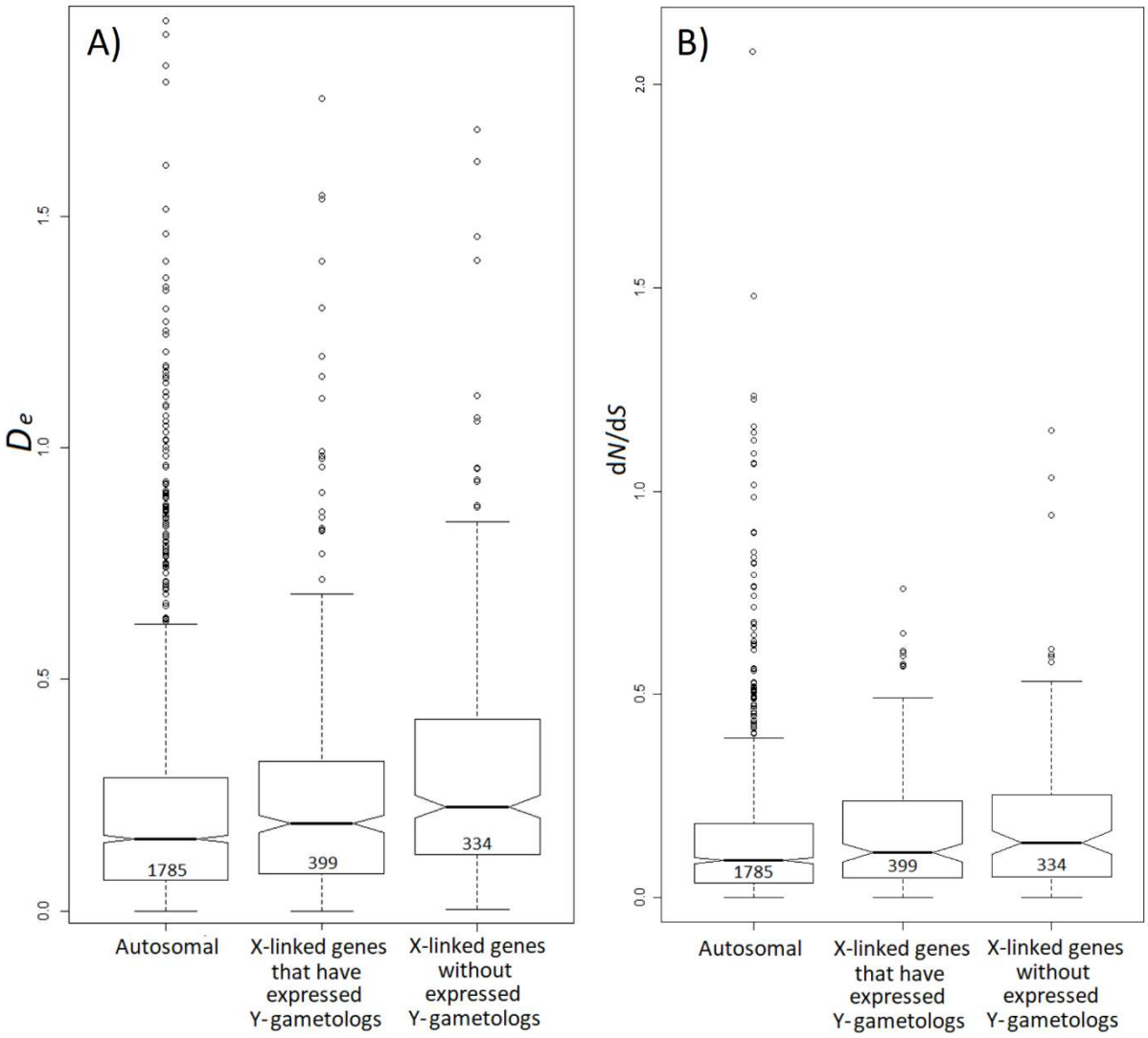


Figure 3. Faster-X evolution for gene expression divergence in males (A) and sequence divergence (d*N*/d*S*, panel B) between *S. latifolia* and *S. dioica*. X-linked genes are split into two categories for X-linked genes with expressed (X_{Yexpr}) and non-expressed ($X_{noYexpr}$) Y-linked gametologs. The numbers of genes analysed are shown inside the boxes.

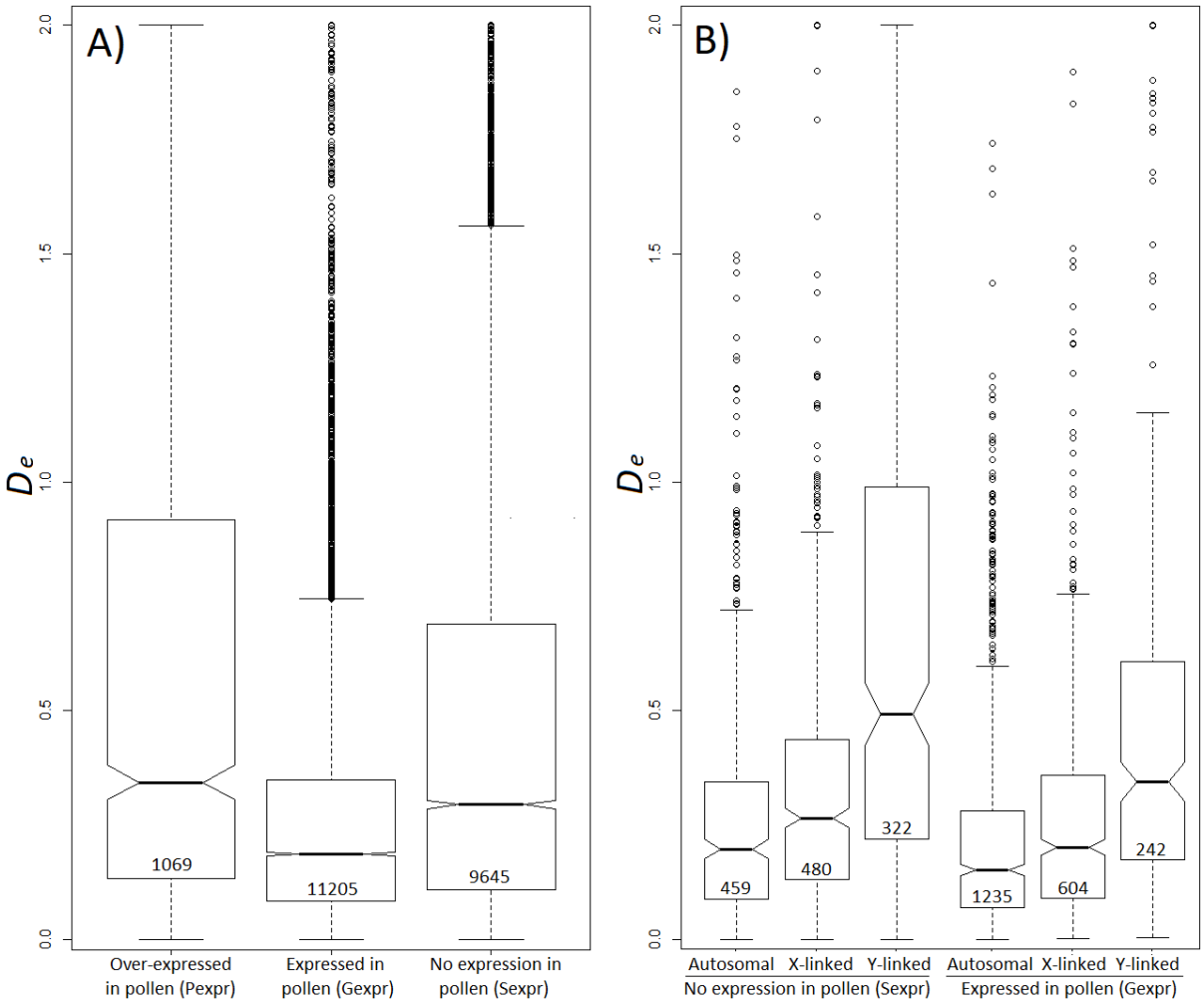


Figure 4. Gene expression divergence between *S. latifolia* and *S. dioica* males for genes predominantly expressed in the sporophyte (S_{expr}) and the gametophyte (G_{expr} and P_{expr}). A) All genes regardless of linkage; B) A comparison of autosomal and sex-linked genes. The numbers of genes analysed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

Suppl. Figures

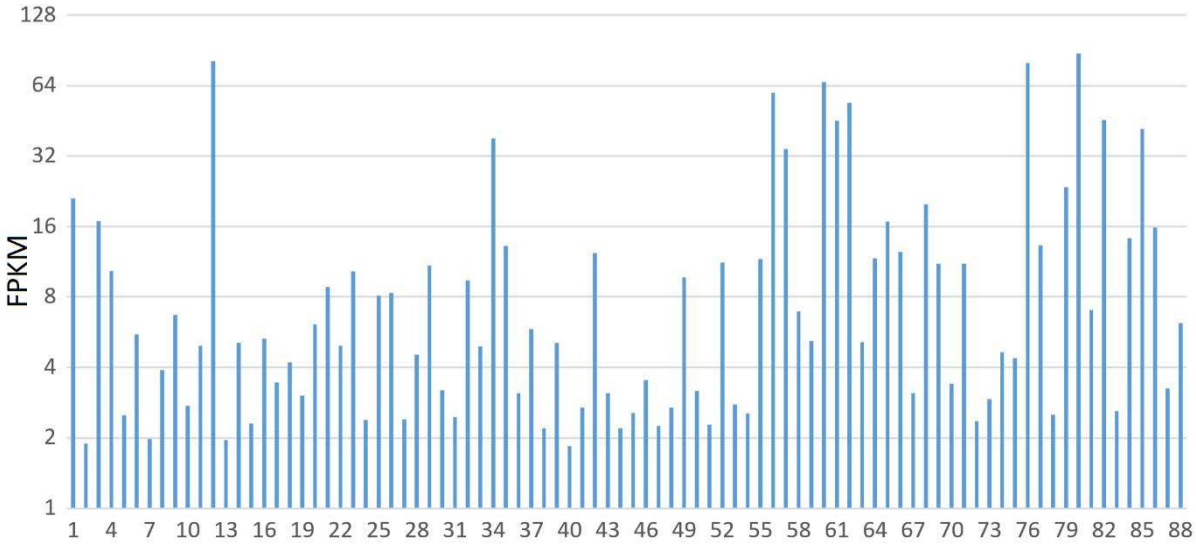


Figure S1 Expression of 88 *S. diclinis* Y-linked genes that lost expression in *S. latifolia* or *S. dioica*.

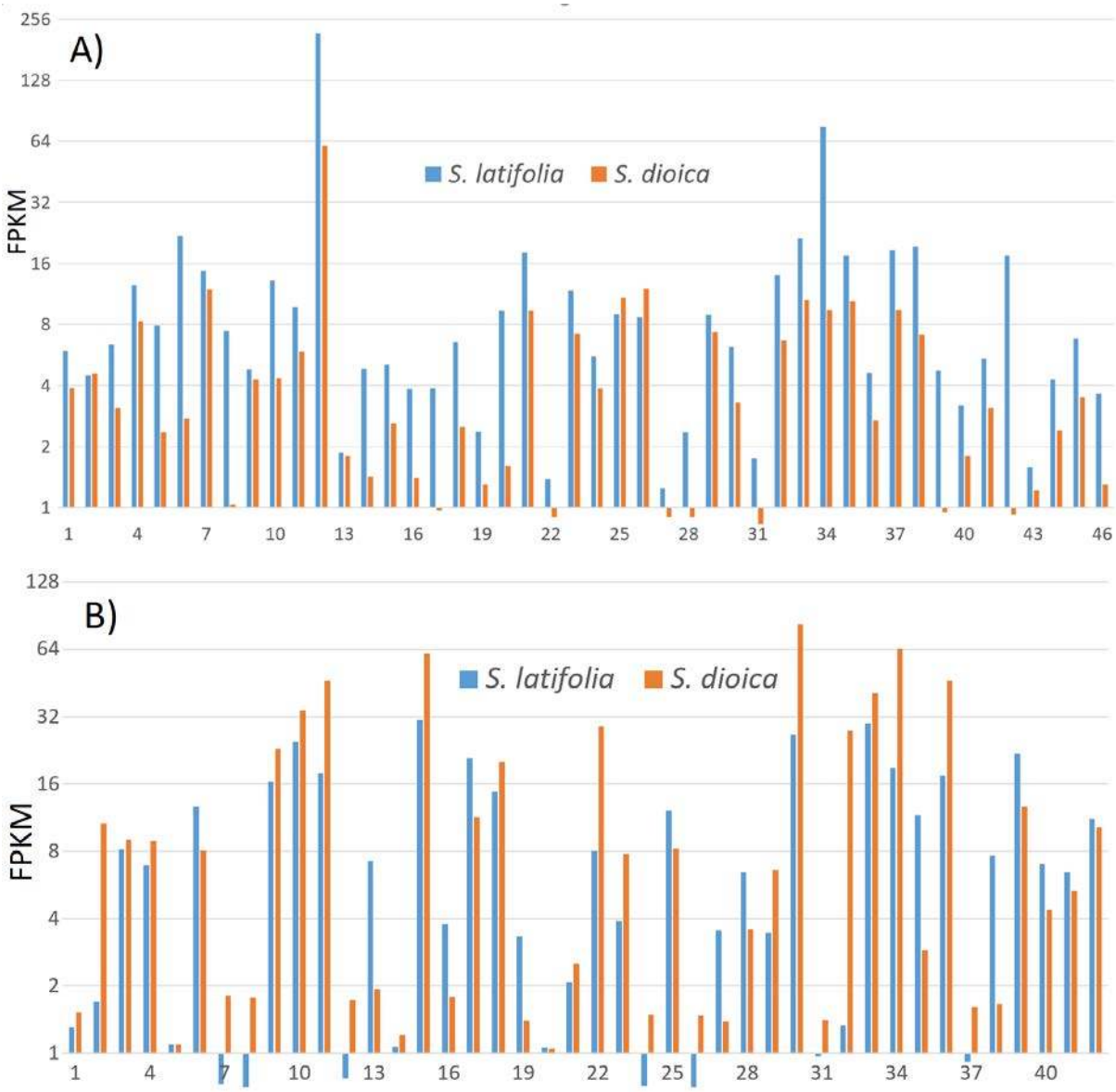
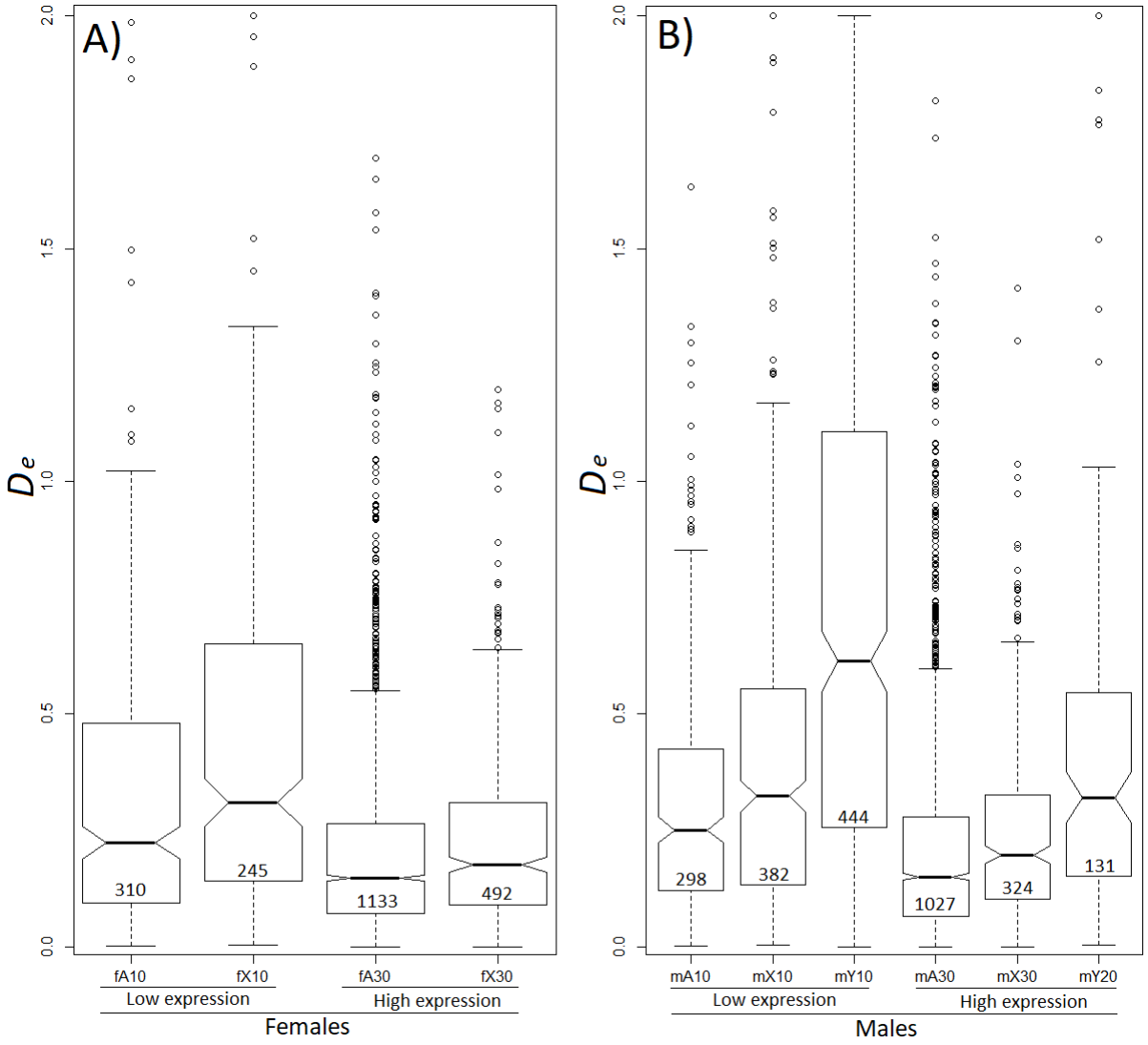


Figure S2. Expression of 46 and 42 X-linked genes that lost expression of the Y-linked gametologs ($X_{noYexpr}$) in *S. latifolia* (A) or *S. dioica* (B) males, respectively. Expression of $X_{noYexpr}$ genes in males is consistently higher in the species where the Y-copy is already degenerate.

675

676



677

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

684

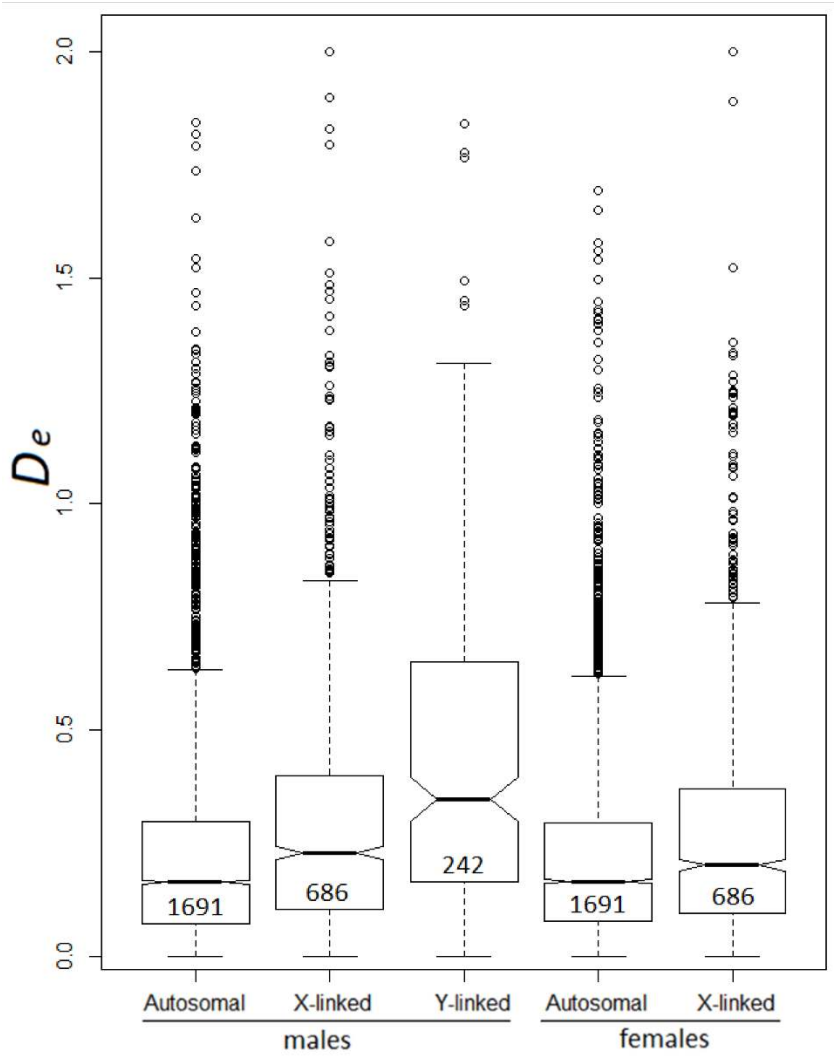


Figure S4. A faster expression divergence (D_e) of the sex-linked genes was detectable after exclusion of weakly expressed genes (fpkm<1). The numbers of genes analysed are shown inside the boxes.

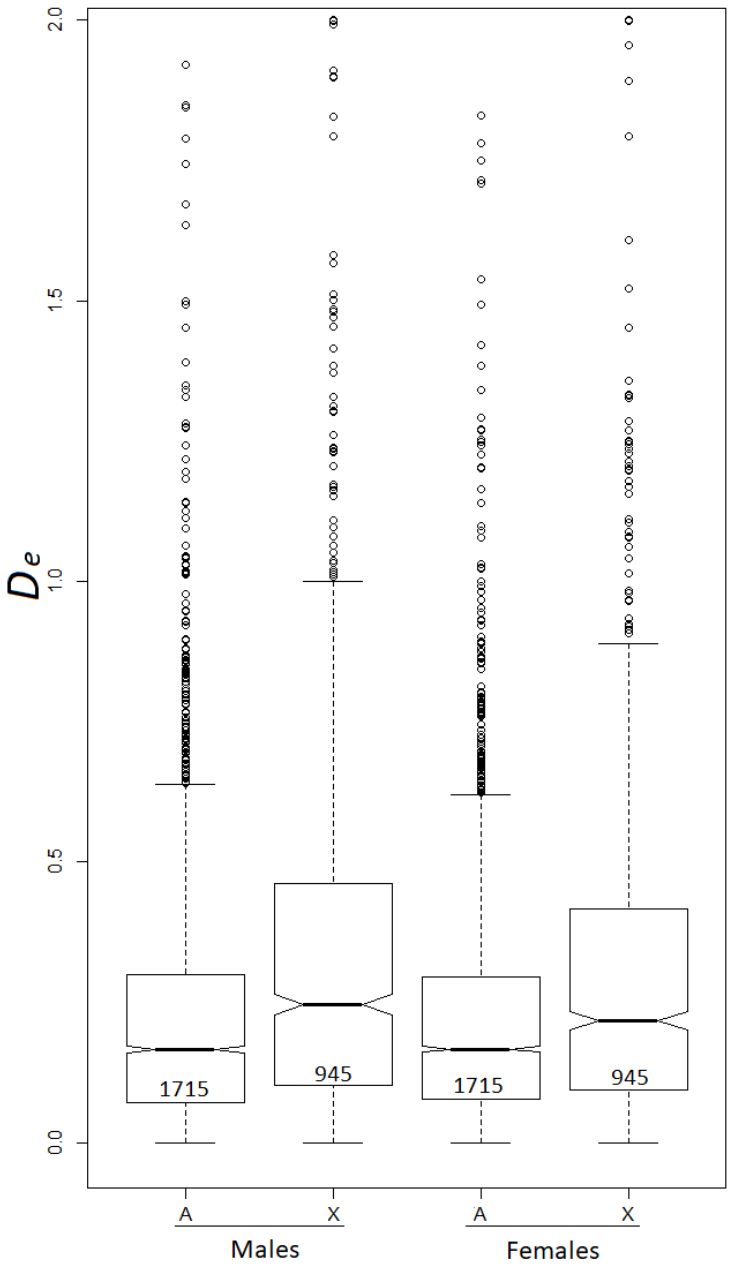


Figure S5. Gene expression divergence (D_e) between *S. latifolia* and *S. dioica* for X-linked (X) and autosomal (A) genes after exclusion of sex-biased genes. Y-liked genes are not shown because by definition all Y-linked genes are male-specific. The numbers of genes analysed are shown inside the boxes.

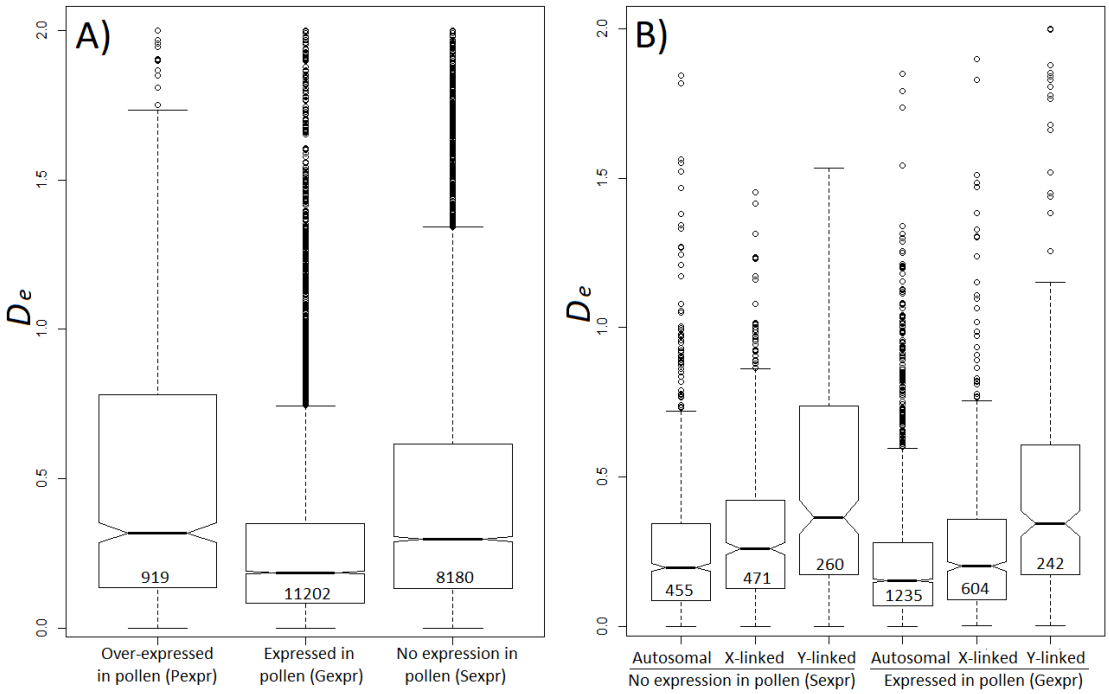


Figure S6. Gene expression divergence between *S. latifolia* and *S. dioica* males for genes predominantly expressed in the sporophyte (S_{expr}) and the gametophyte (G_{expr} and P_{expr}). A) All genes regardless of linkage; B) A comparison of autosomal and sex-linked genes. The numbers of genes analysed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