Purifying and Positive Selection Influence Patterns of Gene Loss and Gene Expression in the Evolution of a Plant Sex Chromosome System

Daisy Crowson,*[,1] Spencer C.H. Barrett,[1] and Stephen I. Wright[1]
[1]Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, ON, Canada
*Corresponding author: E-mail: daisy.crowson@mail.utoronto.ca.
Associate editor: Hideki Innan

Abstract

Sex chromosomes are unique regions of the genome, with a host of properties that distinguish them from autosomes and from each other. Although there is extensive theory describing sex chromosome formation and subsequent degeneration of the Y chromosome, the relative importance of processes governing degeneration is poorly understood. In particular, it is not known whether degeneration occurs solely as a direct result of inefficient selection due to loss of recombination, or whether adaptive gene silencing on the Y chromosome results in most degeneration occurring neutrally. We used comparative transcriptome data from two related annual plants with highly heteromorphic sex chromosomes, *Rumex rothschildianus* and *Rumex hastatulus*, to investigate the patterns and processes underlying Y chromosome degeneration. The rate of degeneration varied greatly between the two species. In *R. rothschildianus*, we infer widespread gene loss, higher than previously reported for any plant. Gene loss was not random: genes with lower constraint and those not expressed during the haploid phase were more likely to be lost. There was indirect evidence of adaptive evolution on the Y chromosome from the over-expression of Y alleles in certain genes with sex-biased gene expression. There was no complete dosage compensation, but there was evidence for targeted dosage compensation occurring in more selectively constrained genes. Overall, our results are consistent with selective interference playing the dominant role in the degeneration of the Y chromosome, rather than adaptive gene silencing.

Key words: **dosage compensation, gene loss, *Rumex*, selective interference, sex-biased gene expression, Y chromosome degeneration.**

Article

Introduction

Heteromorphic sex chromosomes are a striking example of convergent evolution: they are found in taxa as diverse as mammals, arthropods, algae, fish, and flowering plants, and in some clades have arisen independently many times (Bull 1983; Bachtrog et al. 2011). This distribution allows powerful comparisons between closely related taxa to investigate the mechanisms governing sex chromosome evolution. The currently accepted model for the evolution of heteromorphic sex chromosomes (Rice 1987a; Jordan and Charlesworth 2011) involves the accumulation of sexually antagonistic mutations (alleles that increase fitness in one sex but have detrimental effects in the other) in the partially sex-linked pseudo-autosomal region. This process drives selection for suppression of recombination, resulting in complete linkage between the sexually antagonistic allele and the sex-determination locus.

As well as a pre-requisite for divergence between the X and Y chromosomes, loss of recombination drives further divergence due to Hill–Robertson interference between genetically linked sites, in which selection at one site interferes with selection at other linked sites (Hill and Robertson 1966). This process is most easily understood in terms of a reduction in the local effective population size (N_e; Comeron et al. 2008), caused by selection at one locus increasing variance in the 'reproductive success' of genetically linked sites. The overall effect is a reduced efficacy of selection, ultimately leading to genetic degeneration of the sex-limited sex chromosome (Y or W, hereafter referred to as the Y chromosome) due to accumulation of deleterious mutations and repetitive DNA, loss of gene function, and eventually widespread gene loss (Charlesworth and Charlesworth 2000; Bachtrog 2013).

Despite these basic model predictions, highly heteromorphic sex chromosomes with a fully degenerated Y chromosome are not the only outcome of sex chromosome evolution. A lack of extensive degeneration is particularly notable in flowering plants, where highly heteromorphic sex chromosomes are rare (Ming et al. 2011). Instead, plant sex chromosomes are often homomorphic or only moderately differentiated, with recombination occurring over most of their length, as in papaya (VanBuren et al. 2015), *Populus* (Geraldes et al. 2015) and persimmon (Akagi et al. 2014). Moreover, in-depth studies on two plant species with highly heteromorphic sex chromosomes, the unrelated *Rumex hastatulus* and *Silene latifolia*, reveal that although Y chromosome degeneration is occurring, the extent of gene loss is comparatively low, estimated to be $<$30% based on transcriptome data (Hough et al. 2014; Bergero et al. 2015, respectively), although a recent study in *S. latifolia* using a genomic

approach estimated gene loss to be 45% (Papadopulos et al. 2015). Although these cases are often viewed as an intermediate step in the evolution of a fully degenerate Y chromosome, partial degeneration may be the long-term stable outcome of sex chromosome evolution in some taxa.

There are several models describing how non-independence of sites in non-recombining regions could drive genetic degeneration. Background selection reduces the effective population size of linked regions due to the removal of continuously arising deleterious mutations (Charlesworth et al. 1993); Muller's ratchet drives stochastic fixation of deleterious mutations due to the loss of variants without them via genetic drift (Muller 1964); and selective sweeps lower the effective population size of linked regions as a result of the fixation of beneficial mutations (Rice 1987b). These three processes are not mutually exclusive—on the contrary, they are closely related and are all likely to operate jointly to some extent. Indeed, there is evidence for all three processes occurring on the Y chromosome (Bachtrog 2004; Kaiser and Charlesworth 2010; Singh et al. 2014). Moreover, their relative importance is predicted to vary among taxa, sex chromosome systems, and over time (Bachtrog and Gordo 2004; Bachtrog 2008). Processes requiring the action of strong selection (background selection and selective sweeps) might be particularly relevant in plants, as it has been estimated that ~2/3 of all genes are expressed in the male haploid gametophyte (pollen; Honys and Twell 2004), which should lead to strong selection against the degeneration of these genes (Chibalina and Filatov 2011), potentially at the expense of an accelerated rate of degeneration of genes not expressed in the haploid phase.

Crucially, although the relative importance of these three processes is unknown, if they are the drivers of Y chromosome degeneration then degeneration should occur despite a fitness cost. These models predict an inexorable accumulation of mildly deleterious mutations at a rate that directly determines the rate of Y chromosome degeneration. They also predict that degeneration should affect less important genes first, because these are more likely to be subject to mildly deleterious fitness effects.

However, it is also possible that most genetic degeneration occurs neutrally following adaptive gene silencing. Adaptive suppression of expression could occur for two reasons. First, the Y chromosome is likely to experience a lower rate of positive selection compared with the X chromosome, which may select for a suppression of expression of the Y allele in favor of the better-adapted X allele (Orr and Kim 1998). Second, mildly deleterious mutations may accumulate on the Y due to inefficient selection, which may have strong enough fitness costs to actively select for suppressed expression (Charlesworth 1978). In both cases, following adaptive silencing the Y gene copy may degenerate neutrally given its lack of expression. Preferential expression of X over Y alleles in genes that still retain a Y copy has been found in many taxa with partially degenerated Y chromosomes (e.g., Muyle et al. 2012; Hough et al. 2014; White et al. 2015), which could be evidence for adaptive silencing. However, reduced expression of the Y allele is also observed in the very young neo-Y chromosome of *Drosophila albomicans*, which displays very little other signs of degeneration (Zhou and Bachtrog 2012). Therefore, silencing of the Y allele could be a result of the degeneration of regulatory regions through inefficient selection, rather than adaptive suppression of expression.

If adaptive silencing is an important force in Y chromosome degeneration, genes undergoing more positive selection (if the driver was positive selection on the X) or, possibly, more selectively constrained genes (if the driver was deleterious mutations on the Y) should be lost from the Y chromosome first. Adaptive silencing is also more likely if overall gene dosage is not highly affected, either due to dosage compensation or transcriptional buffering (see below), as this would make the fitness cost of expressing the Y allele more likely to outweigh the fitness cost of suppressing its expression.

Degeneration of the Y chromosome has genome-wide consequences. The reduction in gene dosage caused by the loss of a gene copy from the Y should select for the evolution of dosage compensation to restore optimal expression levels within the genome (Vicoso and Bachtrog 2009). However, how and to what extent dosage compensation occurs varies greatly among taxa (Mank 2013). In some taxa chromosome-wide changes have resulted in equal expression between the sexes, although exactly how this occurs differs: for example, in Diptera the single X chromosome in males has up-regulated expression (Vicoso and Bachtrog 2015); in mammals one of the two X chromosomes is inactivated in females (Pessia et al. 2012); and in *Caenorhabditis elegans* both X chromosomes are down-regulated in females (Ercan et al. 2007). Other taxa also have chromosome-wide changes, but fail to fully restore equal expression between the sexes (e.g., *Heliconius* butterflies, Walters et al. 2015). Finally, in many taxa there appears to be patchy dosage compensation, with only some genes having equal expression levels between males and females (e.g., Uebbing et al. 2013; Mahajan and Bachtrog 2015; Papadopulos et al. 2015).

Distinguishing between different models of Y chromosome degeneration and dosage compensation requires comparisons between sex chromosomes that vary in their rate and extent of degeneration in partially degenerated Y chromosomes. This situation may be found in the large dioecious clade of the plant genus *Rumex* (Polygonaceae), which is divided into two smaller monophyletic clades with distinct sex chromosome systems. One clade has an XX/XY sex chromosome system and a Y-based sex-determining mechanism, whereas the other clade has an XX/XY_1Y_2 sex chromosome system and a sex-determining system based on X/autosome balance (Smith 1969; Parker and Clark 1991). The second neo-Y chromosome in this latter clade may have arisen either via an ancient X-autosome fusion event, where the autosomal homologue in males formed the Y_2 chromosome, or a fission in the Y chromosome. The reduction in chromosome number in this clade (Smith 1969; Navajas-Pérez et al. 2005) supports the former idea, although evidence from repetitive DNA accumulation is inconclusive (Steflova, Hobza, et al. 2013; Steflova, Tokan, et al. 2013). Overall, we expect that the XX/XY_1Y_2 clade may have had a larger non-recombining region over much of its history.

Studies investigating the accumulation and distribution of repetitive DNA report a different degree of heteromorphism between sex chromosomes in these two groups of *Rumex* species. In the XX/XY_1Y_2 group the Y chromosomes are heterochromatic, with Y-specific satellite DNA families and recent expansions of repetitive DNA (Kejnovský et al. 2013; Steflova, Tokan, et al. 2013). In the XX/XY group, however, there appears to be less differentiation between sex chromosomes and Y-specific repetitive DNA (Cuñado et al. 2007; Quesada del Bosque et al. 2011). If the sex chromosomes in these two clades have the same origin, these results would suggest that degeneration of the Y chromosome is occurring at different rates in the two sex chromosome systems. However, it is also possible that the XY system is younger than the XY_1Y_2 system, either because sex chromosomes arose independently in the two lineages, or because of sex chromosome turnover.

Here, we use comparative transcriptome data from two species of *Rumex* to investigate the processes that govern sex chromosome evolution. We use transcriptome sequencing to identify and characterise sex-linked genes in *R. rothschildianus*, a species belonging to the XY_1Y_2 clade (see supplementary table S1, Supplementary Material online) with highly heteromorphic sex chromosomes (Wilby and Parker 1988) and compare it to previously generated data from *R. hastatulus* (see Hough et al. 2014), a species that belongs to the XX/XY clade but that has very recently acquired a neo-Y chromosome, also resulting in an XX/XY_1Y_2 system. Using this comparison, we ask: 1) Does the extent of degeneration differ between species? 2) What types of genes experience silencing and loss? 3) Is there global or targeted dosage compensation, and which types of genes show evidence of gene expression evolution? Our results provide novel insights into the factors driving Y chromosome degeneration, and highlight the potential for different rates of degeneration even among closely related species.

Results and Discussion

The Sex Chromosomes Are Unlikely to Be Homologous between *Rumex* Species

We used transcriptome data from two independent crosses of *R. rothschildianus* mapped to a reference transcriptome assembled *de novo* to identify genes located on the sex chromosomes from their diagnostic segregation patterns. From these segregation patterns we were able to identify genes that still have an expressed Y allele (hereafter 'sex-linked XY') as well as genes that have been lost or silenced on the Y chromosome (hereafter 'sex-linked hemizygous'). The number of sex-linked XY genes and sex-linked hemizygous genes identified in *R. rothschildianus* are shown in table 1, as well as a comparison with those numbers for the XY_1Y_2 race of *R. hastatulus* (Hough et al. 2014) using the same identification criteria. The number of sex-linked genes identified using different SNP cut-offs, and the number of autosomal genes identified, are shown in supplementary table S2, Supplementary Material online. Overall, we identified a total of 553 sex-linked genes in *R. rothschildianus*, compared with 1,686 sex-linked genes in *R. hastatulus* (table 1).

Table 1. Number of Genes Identified as Sex-Linked XY and Sex-Linked Hemizygous in *Rumex rothschildianus* and *R. hastatulus*.

	Rumex rothschildianus	*Rumex hastatulus*
Sex-linked XY	**304 (23)**	**1544 (633)**
Sex-linked hemizygous	**249**	**142**
Estimated gene loss[a]	**92%**	**18%**

NOTE.—Sex-linked XY genes are still expressed on the Y chromosome, whereas sex-linked hemizygous genes are lost/unexpressed on the Y chromosome. Values in parentheses represent the number of sex-linked XY genes that were identified using SNPs that constituted polymorphisms on the X as opposed to SNPs between the X and the Y.

[a]The estimated percent gene loss is based on the number of sex-linked XY genes identified using SNPs that constituted polymorphisms on the X.

To be able to make appropriate comparisons between species about patterns of Y chromosome degeneration, it is essential to know whether the evolution of sex chromosomes has occurred independently or not. To determine whether the sex chromosomes in the *Rumex* clades are homologous, we investigated the overlap between genes identified as sex-linked and autosomal in the two species. Surprisingly, the number of orthologous genes that are sex-linked in both species was no different than the number expected to overlap by chance (observed = 40; expected = 40.7; permutation test, $P = 0.99$). Because both species have neo-Y chromosomes, even a single origin of the ancestral sex chromosomes would only imply homology of the older sex chromosomes. Therefore, to increase the power of the comparison we also assessed the overlap of a subset of older sex-linked genes by calculating synonymous site divergence between X and Y chromosomes (d_S), and examining genes where pairwise X–Y $d_S > 0.1$. However, we identified no overlap in old sex-linked genes in the two species (observed = 0; expected = 0.165; permutation test, $P = 0.99$). The percent overlap of sex-linked genes did not vary with the exact criteria used to classify genes into particular datasets, and there was no effect of the number of SNPs used to define a gene as sex-linked on the percent overlap, which would be expected if older sex-linked genes were shared between the species (supplementary table S3, Supplementary Material online).

Therefore, these comparisons provide no evidence of homology between the sex chromosomes. Because old sex-linked genes are the most likely to be detected, due to a higher degree of X–Y divergence increasing the number of sex-linked SNPs, a low detection rate is unlikely to be causing this pattern. Overall, these data, taken together with the phylogeny of these species (supplementary table S1, Supplementary Material online), are consistent with the hypothesis that the two *Rumex* clades have different sex chromosomes.

Extent of Gene Loss Differs between *Rumex* Species

Despite the smaller total number of sex-linked genes in *R. rothschildianus*, we identified considerably more hemizygous genes compared with *R. hastatulus* (table 1). To estimate the percent gene loss, we corrected for the fact

that sex-linked XY genes are easier to detect than hemizygous genes. This is because hemizygous genes can only be identified using X polymorphisms, whereas sex-linked XY genes can be detected using both X polymorphisms and accumulated differences between the X and Y chromosome. We therefore estimated percentage hemizygosity using the number of sex-linked genes that were identified from X polymorphisms. Strikingly, we found over 90% of sex-linked genes are hemizygous in *R. rothschildianus* (table 1), which is 2 or 3 times higher than estimates from other plants with heteromorphic sex chromosomes, such as *R. hastatulus* and *S. latifolia* (table 1; Hough et al. 2014; Papadopulos et al. 2015). Note that estimated percent gene loss in *R. hastatulus* will be affected by the recent acquisition of a neo-Y chromosome, which increases the number of sex-linked XY genes while presumably not immediately increasing the number of sex-linked hemizygous genes. However, previous work on *R. hastatulus* estimated gene loss to be below 30% in the race that lacks the neo-Y chromosome (Hough et al. 2014). Therefore, although the exact magnitude of the difference is unclear, the results show that *R. rothschildianus* has ~3 times higher percentage gene loss compared with *R. hastatulus*.

It is important to note that transcriptome data cannot distinguish between genes that have been lost from the Y from those that are just unexpressed. However, two studies explicitly comparing lack of expression with DNA sequence in *S. latifolia* concluded that in most cases lack of expression of the Y allele was due to loss of the gene itself (Bergero et al. 2015; Papadopulos et al. 2015). These studies also resulted in a higher estimate of gene loss than those based on transcriptome data, suggesting our methods are conservative. Similarly, genomic evidence suggests that in *R. hastatulus* the majority of genes identified as hemizygous using transcriptome data are deleted from the genome (Beaudry et al. unpublished data).

The absence of plant species with completely degenerated Y chromosomes has led to the idea that the Y chromosome does not degenerate as much or as quickly in plants, compared with animals. The reason invoked is that a large proportion of the plant genome (~2/3) is expressed in the male haploid gametophyte (Honys and Twell 2004), which could mean that loss is only tolerated for ~1/3 of Y chromosome genes. There are several possible explanations for the greater extent of gene loss in *R. rothschildianus* (~90%; table 1) than this theoretical maximum. First, it is possible that not all of the genes expressed in the male gametophyte are fundamental to its function, thus freeing up a larger proportion of the genome to degeneration. Second, although unexpressed in diploid tissue, these genes might still be present in the genome and capable of being expressed in the haploid phase. Last, lack of recombination might be causing genetic degeneration despite selection against it, resulting in important genes being lost. This would result in Y-bearing pollen being less competitive than X-bearing pollen, a phenomenon called certation (Correns 1928). There is indirect evidence that this might occur from the widespread female-biased sex ratios in plants with heteromorphic sex chromosomes (Field et al. 2012a), which in *Rumex* has been shown to be determined at least partly at the pre-zygotic level (Stehlik and Barrett 2006; Błocka-Wandas et al. 2007; Field et al. 2012b). This could result in an interesting balance between selection for equal sex ratios (Fisher 1930) and the degeneration of the Y chromosome, potentially resulting in stable female-biased sex ratios (Hough et al. 2013).

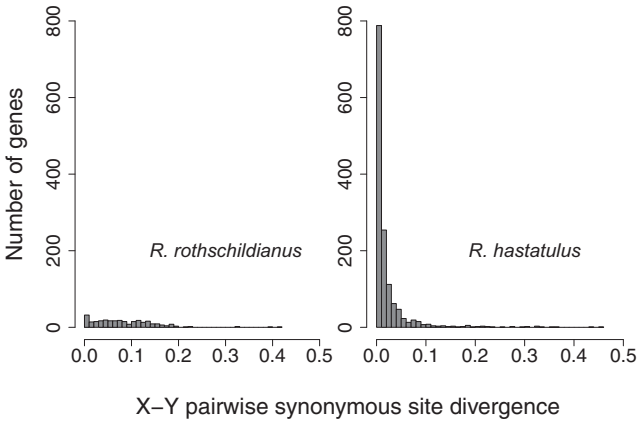

Fig. 1. Pairwise synonymous site divergence between X and Y sequences in sex-linked genes that still retain expression of the Y copy in *Rumex rothschildianus* and *Rumex hastatulus*.

Rate of Gene Loss Differs between *Rumex* Species

The difference in extent of gene loss between the two species must be due to either a difference in the rate of Y chromosome degeneration or a difference in the age of the sex chromosomes. Therefore, to estimate the age of the sex chromosomes, we examined the distribution of pairwise X–Y synonymous sequence divergence (d_S) for each gene (fig. 1). Median X–Y divergence in *R. rothschildianus* was 0.076, whereas in *R. hastatulus* it was 0.008. However, this difference is due to the large number of young genes with very low divergence in *R. hastatulus*, as opposed to any difference in the range of d_S values (fig. 1).

The age of sex chromosomes can be estimated using synonymous site divergence if a molecular clock is assumed. Although the highest d_S values should correspond to the oldest XY genes and therefore provide an estimate of when recombination was initially suppressed, synonymous site divergence is stochastic, and using extreme values in calculations based on average rates introduces bias. We therefore used the median value (plus and minus their associated error in the estimation; Yang and Nielsen 2000) of an arbitrarily defined old set of XY genes ($d_S > 0.1$), which included >50 genes in both species. The median X–Y d_S value for old genes was 0.12–0.16 for *R. rothschildianus* and 0.13–0.22 for *R. hastatulus*. Dividing these values by half (as divergence is occurring along both the X and Y branches) and using the direct estimate of spontaneous mutation rate in *Arabidopsis thaliana* of 7×10^{-9} per site per generation (Ossowski et al. 2010) gives an estimate of recombination suppression along the ancestral sex chromosomes having occurred 8–11 million generations ago in *R. rothschildianus* and 9–16 million in *R. hastatulus*. Although there is uncertainty associated with these calculations,

Table 2. Molecular Evolution on the X and Y Chromosomes in *Rumex rothschildianus* and *R. hastatulus*.

	Rumex rothschildianus		*Rumex hastatulus*	
	X	Y	X	Y
Number passing BM LRT (%)	33 (16%)	71 (35%)	56 (5%)	172 (16%)
% faster than background	21%	94%	67%	93%
Mean (SE) d_N/d_S	0.23 (0.02)	0.39 (0.02)	0.22 (0.02)	0.48 (0.03)
Median d_N/d_S	0.15	0.31	0.09	0.32

NOTE.—BM LRT, branch model likelihood ratio test. SE, standard error. PAML branch models were used to determine whether the rate of X or Y varied significantly from the background rate. The d_N/d_S values were estimated from the 'free ratios' model in PAML.

and we cannot rule out the presence of additional evolutionary strata due to multiple time points of recombination suppression, our results suggest the two sex chromosome systems arose at similar times, which means it is likely that the Y chromosome is experiencing a faster rate of gene loss in *R. rothschildianus* compared with *R. hastatulus*.

Genes Retained on the Y Are Also Degenerating

Degeneration is also likely to occur for genes remaining on the Y chromosome, resulting in an accumulation of deleterious mutations in expressed genes. We therefore further investigated Y-chromosome degeneration by comparing lineage-specific estimates of d_N/d_S between X and Y sequences for each species using various models of molecular evolution implemented in PAML (table 2). 'Branch models' allow d_N/d_S to vary on particular specified branches (here the X or Y branch of either *R. rothschildianus* or *R. hastatulus*) compared with a background rate estimated for the rest of the tree, which also included the hermaphrodite *R. bucephalophorus*. The lack of homology between the sex chromosome systems means this background rate represents the 'ancestral' evolutionary rate prior to sex linkage.

We estimated average branch-specific d_N/d_S using the 'free-ratios' model, which allows the rate to vary freely across the entire tree. Genes with very low X–Y synonymous divergence ($d_S < 0.001$) were removed from the calculations, as they give highly skewed estimates of d_N/d_S. The average d_N/d_S was significantly higher on the Y branch compared with the X branch in both *R. rothschildianus* and *R. hastatulus* (Wilcoxon test, $P < 10^{-12}$ for both comparisons; table 2). The d_N/d_S ratio of the X branch for each gene in both *R. rothschildianus* and *R. hastatulus* was not significantly different from the rate of evolution of these same genes in *R. bucephalophorus*, where they are not sex-linked (Wilcoxon test, $P \geq 0.54$ in both comparisons), whereas the Y was (Wilcoxon test, $P < 10^{-15}$ in both comparisons). These results confirm that the changes in d_N/d_S have occurred on the Y branch as opposed to the X branch.

An increase in d_N/d_S ratio can be driven by either a higher non-synonymous or lower synonymous rate of evolution. The synonymous site divergence (d_S) is also significantly different in *R. rothschildianus* between X and Y branches (mean d_S X $= 0.037 \pm 0.002$; mean d_S Y $= 0.057 \pm 0.003$; Wilcoxon test, $P < 10^{-8}$), which implies either a higher mutation rate on the Y chromosome or a reduced efficacy of selection on codon usage bias. However, this difference would result in a lower d_N/d_S ratio on the Y if it were the main driver of the difference in molecular evolution between the X and Y. Indeed, the difference in non-synonymous site divergence (d_N) is much larger (mean d_N X $= 0.006 \pm 0.0005$, mean d_N Y $= 0.021 \pm 0.002$; Wilcoxon text, $P < 10^{-15}$). This contrasts with *R. hastatulus*, where there is no evidence for elevated d_S on the Y (see Hough et al. 2014).

Degree of Constraint of Genes Retained on the Y Varies between the *Rumex* Species

The greater extent of gene loss in *R. rothschildianus* compared with *R. hastatulus* suggests a higher rate of Y chromosome degeneration in the former, which may also be reflected in genes that still retain expression of the Y allele. To investigate this, we compared the rate of evolution of sex-linked XY genes between the species. In contrast to our prediction, pairwise X–Y d_N/d_S was significantly higher in *R. hastatulus* compared with *R. rothschildianus* (median *R. rothschildianus* $d_N/d_S = 0.322$; median *R. hastatulus* $d_N/d_S = 0.498$; Mann–Whitney *U* test, $P < 10^{-7}$; supplementary fig. S1, Supplementary Material online). Thus, contrary to estimates of gene loss in the two systems, this latter result suggests that degeneration of sex-linked XY genes is occurring at a faster rate in *R. hastatulus* than *R. rothschildianus*.

The sex chromosomes of the two *Rumex* species are not homologous, which means underlying differences in gene content might be driving the increased rate of non-synonymous evolution in *R. hastatulus* relative to *R. rothschildianus*. Moreover, the extensive gene loss in *R. rothschildianus* could mean sex-linked XY genes are no longer representative of the way most genes degenerated prior to being lost. To investigate whether there is some underlying difference in the sex-linked XY genes between the two species we compared the d_N/d_S ratio of their orthologs in *R. bucephalophorus*. *R. bucephalophorus* genes that are sex-linked XY in *R. rothschildianus* were significantly more constrained than genes that are sex-linked XY in *R. hastatulus* (median d_N/d_S 0.16 and 0.20; Mann–Whitney *U* test, $P < 0.02$). When the same comparison was done for all sex-linked genes, including genes that have been lost from the Y chromosome in both species, the difference between the two sets of genes disappeared (median d_N/d_S 0.18 and 0.20; Mann–Whitney *U* test, $P = 0.67$). Finally, the negative correlation between X–Y d_N/d_S and ancestral gene expression (supplementary fig. S2, Supplementary Material online) in both *R. rothschildianus* ($r_s = -0.22$, $P < 0.001$) and *R. hastatulus* ($r_s = -0.17$, $P < 10^{-6}$), which suggests that highly expressed genes are more constrained than those that are expressed at lower levels, allowed us to perform a second comparison of constraint between the two sets of genes. We compared average ancestral expression and found that it was lower for genes that are sex-linked XY in *R. hastatulus* than *R. rothschildianus* (Mann–Whitney *U* test, $P < 0.003$), again suggesting that these genes are under less selective constraint.

These results are consistent with the hypothesis that there are differences in the particular genes that are sex-linked XY in the two species, suggesting that widespread gene loss is changing the composition of sex-linked XY genes in *R. rothschildianus*, resulting in them being on average more constrained (see below). Taken together with the estimates of gene loss, our results suggest that there is no clear difference between the two species in the nature of degeneration, or in the overall selective pressure on genes on the respective sex chromosomes. Instead, there is only a difference in the amount of degeneration that has already occurred. That is, *R. rothschildianus* has lost less constrained genes from the Y chromosome, whereas in *R. hastatulus* these are still present but are accumulating a high frequency of deleterious mutations.

Although it is clear from comparative studies (e.g., Bachtrog et al. 2011; Bachtrog 2013; Vicoso et al. 2013; Zhou et al. 2014) that the rate and extent of degeneration of the sex-limited sex chromosome vary considerably, comparisons are often between taxa that diverged many millions of years ago, with fundamental differences in their life history and, in some cases, major differences in sex chromosome systems and in their ages. The contrasting results shown here between *R. rothschildianus* and *R. hastatulus* demonstrate that even closely related taxa with similar life histories and sex chromosomes of roughly similar ages can have very different rates of degeneration. The evidence for a faster rate of degeneration in *R. rothschildianus* is consistent with the prediction that a larger non-recombining sex chromosome will experience stronger effects of linked selection. However, we cannot rule out alternative explanations, such as differences in ancestral population size or a large-scale deletion/silencing event in *R. rothschildianus*. Genome sequencing and a larger comparative genomics effort across the genus is necessary to fully understand the causes of the contrasting rates of degeneration.

Gene Loss from the Y Chromosome Is Not Random

The large number of genes in both sex-linked XY and sex-linked hemizygous datasets in *R. rothschildianus* allowed us to investigate the processes governing gene loss from the Y—in particular, what kinds of genes are lost, and what implications this has about the various models describing Y chromosome degeneration.

The rate of molecular evolution of hemizygous genes in *R. rothschildianus* (table 3) was significantly higher than autosomal genes (Mann–Whitney *U* test, $P < 0.01$), although not significantly different from genes on the X chromosome that still retain a Y copy (Mann–Whitney *U* test, $P = 0.12$). To investigate whether hemizygous genes were ancestrally under less constraint prior to the evolution of sex chromosomes, we compared evolutionary rates of genes orthologous to hemizygous genes in *R. rothschildianus* to genes that are orthologs to autosomal genes and sex-linked genes retained on the Y (table 3). Orthologs of genes that are hemizygous in *R. rothschildianus* have a significantly faster rate of evolution than orthologs of the other two classes in both *R. bucephalophorus* (Mann–Whitney *U* test, $P < 0.05$ for both comparisons) and *R. hastatulus* (Mann–Whitney *U* test, $P < 0.02$ for both comparisons). The fact that these genes show faster rates of molecular evolution prior to becoming sex-linked suggests that the lower constraint is a cause rather than a consequence of hemizygosity. Although the difference between sex-linked hemizygous and sex-linked genes that retain a copy on the Y chromosome in *R. rothschildianus* is not significant, this could be due to greater exposure of deleterious mutations in hemizygous genes driving more similar levels of constraint. To further compare the likely degree of constraint experienced by genes that have been lost compared with genes that retain a Y copy, we compared ancestral expression of these genes. The ancestral expression of hemizygous genes is significantly lower than for genes that remain on the Y (median normalised ancestral expression was 614.5 for sex-linked XY genes and 456.3 for sex-linked hemizygous genes, Mann–Whitney *U* test, $P < 0.02$). These results suggest that less constrained genes are more likely to be lost from the Y chromosome.

Table 3. Molecular Evolution of Different Gene Sets in *Rumex rothschildianus* and Their Orthologs in *R. hastatulus* and *R. bucephalophorus*.

Gene Set in *Rumex rothschildianus*	*Rumex rothschildianus*	*Rumex hastatulus*	*Rumex bucephalophorus*
Hemizygous	**0.225 (0.012)**	**0.236 (0.012)**	**0.231 (0.012)**
Sex-linked XY	**0.201 (0.010)**	**0.206 (0.010)**	**0.197 (0.010)**
Autosomal	**0.155 (0.006)**	**0.208 (0.006)**	**0.196 (0.006)**

NOTE.—All values are mean d_N/d_S and associated standard error in parentheses. Sex-linked XY genes are still expressed on the Y chromosome, whereas sex-linked hemizygous genes are lost/unexpressed on the Y chromosome.

It is also possible that the higher rate of molecular evolution is due to prevalent positive selection on hemizygous genes, either as a cause of their hemizygosity, under a model of adaptation on the X driving silencing of the Y (Orr and Kim 1998), or as a consequence of hemizygosity, due to more efficient positive selection of recessive alleles in hemizygous genes (Faster-X effect; Charlesworth et al. 1987). To investigate these hypotheses, we fitted a site model in PAML, which allows rates to vary between sites to provide a test for positive selection ($d_N/d_S > 1$) operating on particular sites. The proportion of genes passing the LRT for the presence of positive selection in hemizygous genes (17%) was similar to the proportion of autosomal genes (15%) and sex-linked XY genes (21%). This result suggests that the higher average d_N/d_S of hemizygous genes is not caused by widespread positive selection on these genes, but is instead a result of lower selective constraint.

Genetic degeneration of the Y chromosome in plants may be limited by extensive gene expression in the male gametophyte. To test whether selection during the haploid pollen phase of the life cycle limits the extent of gene loss from the Y chromosome, we used pollen and sperm transcriptome data from *A. thaliana* and *Nicotiana tabacum* to identify genes expressed in the haploid phase of the life cycle, and tested for an underrepresentation of haploid-expressed genes in hemizygous genes compared with genes that remain on the Y (supplementary table S4, Supplementary Material

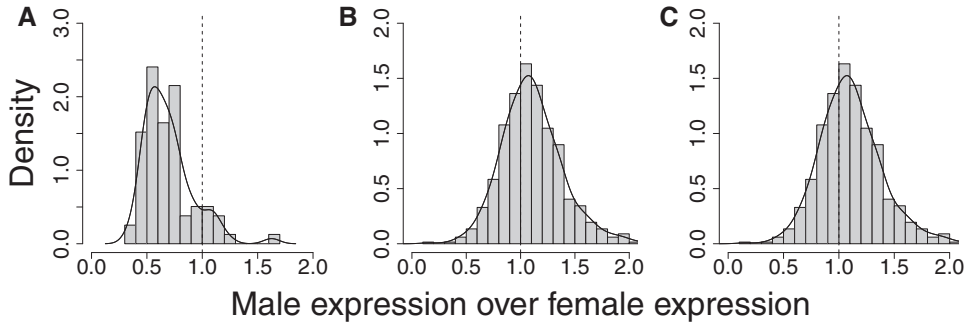

FIG. 2. Probability density histograms representing male gene expression relative to female gene expression in different sets of genes in *Rumex rothschildianus*. Sex-linked hemizygous genes (*A*) have only a single expressed copy in males and two expressed copies in females, whereas sex-linked XY genes (*B*) and autosomal genes (*C*) have two expressed copies in both males and females. Dotted line represents null expectation of 1:1 male/female expression ratio.

online). There was a significant underrepresentation of haploid-expressed genes in the set of genes that have been lost from the Y chromosome (Fisher's exact test, $P < 0.001$), whereas the difference between all sex-linked genes and autosomal genes was not significant (Fisher's exact test, $P = 0.67$). This suggests genes expressed during the haploid phase of the life cycle are less likely to be lost from the Y chromosome than genes that are only expressed in diploid tissue.

A source of bias could arise if genes expressed during the haploid phase are on average more constrained than genes that are only expressed in the sporophyte, as we have already shown less constrained genes are more likely to be lost. Indeed, genes of *R. rothschildianus* expressed in the haploid phase are significantly more constrained than genes not expressed in the haploid phase (median 0.160 and 0.168 respectively, Mann–Whitney *U* test, $P < 0.03$). However, the difference is small, and the same genes are not significantly different from each other in *R. bucephalophorus* (Mann–Whitney *U* test, $P = 0.34$). Therefore, it seems unlikely that this is the only factor driving the pattern of gene loss, but rather that haploid expression affects rate of gene loss in addition to degree of constraint.

We also carried out a functional enrichment analysis of genes that remain on the Y chromosome compared with genes that are lost from the Y chromosome, to see whether some types of genes are more likely to be lost than others. Sex-linked XY genes were enriched for various fundamental categories, such as intracellular transport, protein transport, post-embryonic development, and RNA binding, splicing and processing (supplementary table S5, Supplementary Material online), whereas sex-linked hemizygous genes were enriched for fewer categories of probably less fundamental processes (supplementary table S6, Supplementary Material online).

Taken together, these results suggest that although gene silencing and/or loss in *R. rothschildianus* is extensive, and has occurred rapidly, it is not a rampant or random process. A similar pattern has been observed on the neo-sex chromosome of *Drosophila miranda*, where genes that have been lost were ancestrally expressed at lower levels (Kaiser et al. 2011), and in threespine stickleback, where genes retained on the oldest stratum are under stronger purifying selection than the rest of the X chromosome (White et al. 2015).

The genes that still remain on the Y chromosome in *R. rothschildianus* could reflect a core set of genes that cannot be lost, similar to that found in old, fully degenerated sex chromosomes, although the number of genes involved would be ~10 times higher than in mammals (~36 retained genes; Bellott et al. 2014) or *Drosophila* species (<20 retained genes; Carvalho et al. 2009). This difference could be the result of more extensive haploid gene expression in plants, which might increase the number of genes that cannot be lost from the Y chromosome because they are essential for the development of Y-bearing pollen. However, the relative youth of sex chromosomes in *R. rothschildianus* (~8–10 million generations compared with, for example, ~160–180 million years in mammals; Potrzebowski et al. 2008; Veyrunes et al. 2008) suggests gene loss is probably ongoing. This implies purifying selection on the Y chromosome has a greater role than just maintaining a core set of genes. Instead, purifying selection determines the order in which genes are lost, as opposed to gene loss being a random mutation-driven process.

There Is No Complete Dosage Compensation

To investigate whether *R. rothschildianus* shows evidence of dosage compensation, we compared the relative expression levels of different sets of genes in males and females (fig. 2). Sex-linked hemizygous genes have only a single expressed copy in males and two expressed copies in females, and the distribution of total male/female expression is centred around 0.5, although there is a shoulder representing genes that are expressed similarly in both sexes (fig. 2*A*), whereas sex-linked genes that still retain expression of the Y copy (fig. 2*B*) and autosomal genes (fig. 2*C*) have a male/female expression ratio centred around the null expectation of 1, corresponding to equal male and female gene expression. These results suggest that, despite widespread gene loss, complete dosage compensation has not evolved in *R. rothschildianus*, although a small subset of hemizygous genes is expressed at equal levels between males and females.

This result indicates that for a large number of genes male expression levels are lower than female, presumably closer to optimal, expression levels. This is in accord with reports of the lack of complete dosage compensation from a wide range of species, including birds, many insects, fish and snakes

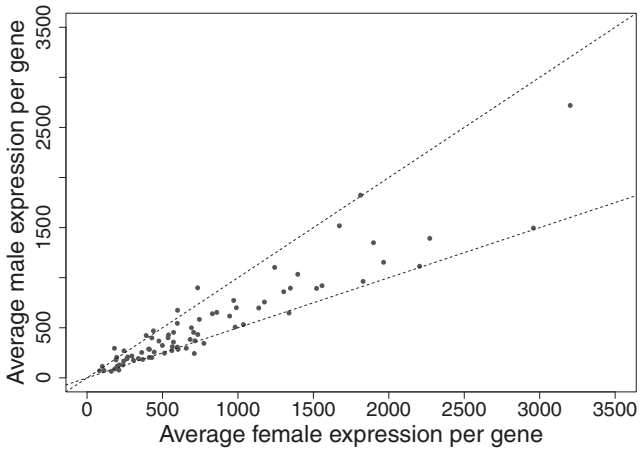

Fig. 3. Male versus female expression level for hemizygous genes in *Rumex rothschildianus*. Lines showing 1:1 expression levels and 0.5:1 expression levels expected under scenarios of full dosage compensation (equal expression in both sexes), or no dosage compensation (half the expression level in males compared with females).

(e.g., Ellegren et al. 2007; Uebbing et al. 2013; Vicoso et al. 2013; Chen et al. 2014; Mahajan and Bachtrog 2015; Walters et al. 2015). The fact that full dosage compensation fails to occur in many species implies either the fitness cost of a reduced expression level in some genes is not large or, alternatively, that dosage compensation is difficult to achieve. In either case, it suggests that widespread gene degeneration did not occur neutrally following widespread dosage compensation.

There Is Evidence for Targeted Dosage Compensation

It is possible that particular genes, rather than the whole chromosome, have been targeted for dosage compensation. This is perhaps more likely in *R. rothschildianus* because many genes present on the Y chromosome may not require correction of dosage. To investigate this hypothesis, we looked at the male/female ratio on a gene-by-gene basis (fig. 3). In a scenario of full dosage compensation most genes should fall on the 1:1 line, whereas with no dosage compensation they should fall close to the 0.5:1 line. Although in some cases hemizygous gene expression in males is close to the null expectation of half the expression level found in females, there are many genes that are above the line. However, between-individual variance introduces noise and makes it difficult to distinguish when a gene is likely to have been 'dosage compensated' from random noise. We therefore directly tested for statistically significant differential expression (see Methods) to more robustly categorise the gene set into 'dosage compensated' or 'not dosage compensated' depending on whether males had expression levels that were significantly higher than half the value of female expression (table 4). Approximately a quarter of hemizygous genes are expressed at significantly higher levels than half the value of female expression, and a subset of these (~14% of all hemizygous genes) also show no significant difference in expression between males and females.

To investigate whether genes defined as 'dosage compensated' were likely to be actively targeted for dosage compensation we compared the male/female expression ratio to the degree of constraint of hemizygous genes, using this as a proxy for its overall importance. Because dosage compensation may be more difficult to accomplish in genes expressed at higher levels, and more constrained genes have higher expression (supplementary fig. S3, Supplementary Material online), we included ancestral gene expression in a partial correlation test of degree of compensation and d_N/d_S ratio, which was significant ($r_s = -0.30$, $P < 0.05$; supplementary fig. S4, Supplementary Material online). This suggests that targeted dosage compensation is occurring in males, up-regulating the expression of more selectively constrained, and presumably more important, genes.

We carried out two further tests to confirm that these genes are likely to have been targeted for dosage compensation. First, we compared their ancestral expression levels, because genes with low expression are more likely to have buffered expression resulting in 'automatic' dosage compensation (Malone et al. 2012). However, there was no significant difference between ancestral expression levels between genes defined as 'dosage compensated' compared with those defined as 'not dosage compensated' (Mann–Whitney *U* test, $P = 0.92$) and there was no significant correlation between male/female expression ratio and ancestral gene expression ($r_s = 0.1$, $P = 0.47$). It is also clear from figure 3 that not all genes with male/female ratios close to 1 are expressed at low levels. Second, we tested for a correlation between male/female expression ratio and the between-individual variance, observed in some studies reporting patchy dosage compensation (e.g., Uebbing et al. 2013), which if significant would suggest apparent dosage compensation was a statistical artefact. However, there was no significant correlation between degree of dosage compensation and between-individual variance ($r_s = -0.09$, $P = 0.40$). Therefore, similar to reports from chicken sex chromosomes, where dosage compensation is occurring only in dosage sensitive genes (Zimmer et al. 2016), our results suggest that active targeted dosage compensation of particular important genes is occurring in *R. rothschildianus*.

Table 4. Number of Genes With Significantly Different Expression Levels between Males and Females for Different Gene Sets in *Rumex rothschildianus*.

	Hemizygous Sex-Linked	XY	Autosomal
% significantly different	76%	51%	14%
Number lower in males/total	58/58	56/156	40/91
% significantly different from 1/2 female expression	24%		
Number lower in males from 1/2 female expression/total	3/22		

Note.—Sex-linked XY genes are still expressed on the Y chromosome, whereas sex-linked hemizygous genes are lost/unexpressed on the Y chromosome. Significance was calculated using differential expression analysis based on a negative binomial distribution model.

Changes in Gene Expression from Ancestral Levels

A comparison between male and female individuals is a simplistic way of testing for dosage compensation, because

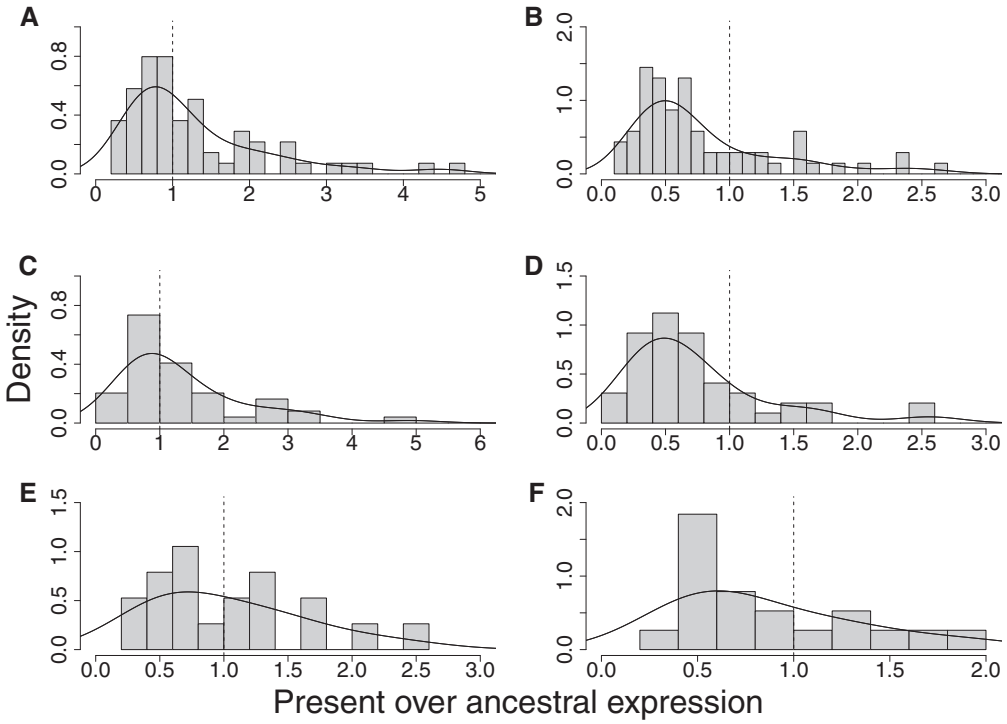

Fig. 4. Probability density histograms representing the ratio of present day expression in hemizygous genes in *Rumex rothschildianus* females and males over an estimate of ancestral expression given by expression levels of putatively autosomal orthologs in *Rumex hastatulus* in females (*A*) and males (*B*). Female and male hemizygous genes were further divided into statistically 'not dosage compensated' (*C* and *D*, respectively) versus 'dosage compensated' (*E* and *F*, respectively). Dotted line shows null expectation of 1:1 present/ancestral expression.

changes in expression can occur in both sexes. We therefore used ancestral expression levels to determine the direction of change in gene expression in *R. rothschildianus*, to assess whether the overall difference between male and female hemizygous gene expression is coming from changes to male or female expression. Whereas females showed similar expression levels to ancestral expression, with a female/ancestral ratio centred around 1 (fig. 4*A*), males had lower expression than ancestral levels, with male/ancestral expression centred around 0.5 (fig. 4*B*). This result indicates that the differences between male and female expression of hemizygous genes in *R. rothschildianus* are arising due to the loss of one copy of the gene in males, as opposed to changes in female expression of these genes.

We divided hemizygous genes into statistically 'dosage compensated' (those with male expression significantly above half the value of female expression) and 'not dosage compensated' (those with male expression not significantly above half the value of female expression) and compared expression levels of these two groups to ancestral expression, to address two questions. First, we asked whether genes in which male expression is not significantly different from half the value of females ('not dosage compensated') are expressed at ancestral levels in females, or whether some up-regulation of male transcription has led to a corresponding increase in female gene expression, keeping the male/female expression close to 0.5 but shifting the female expression away from optimal (Mullon et al. 2015). The distribution of female/ancestral expression was centred around 1 (fig. 4*C*), whereas the distribution of male/ancestral expression was centred around 0.5 (fig. 4*D*). Moreover, a more rigorous test for differential expression between current and ancestral levels showed that current female expression varied equally in both directions relative to ancestral levels (50 genes significant, 46% of them lower than ancestral) whereas male expression was biased towards being significantly lower than ancestral (62 genes significant, 74% of them lower than ancestral). These results suggest that the lower expression levels in males are not ancestral, and females have not moved away from optimal expression.

Second, we asked whether genes that have been compensated have simply been up-regulated in males, resulting in equal gene expression compared with ancestral expression levels. Statistically 'dosage compensated' genes had a distribution of expression shifted below ancestral expression in both females (fig. 4*E*) and males (fig. 4*F*). Supporting this, compensated genes that showed significantly different expression between current and ancestral levels are biased towards lower current expression in both females (17 genes significant, 10 of these lower than ancestral) and males (15 genes significant, 11 of these lower than ancestral). These results suggest that dosage compensation has not occurred entirely through up-regulation of male transcription, but also through down-regulation of female expression. This pattern has also been reported in other species that exhibit partial dosage compensation, such as Strepsiptera beetles and *Heliconius* butterflies (Mahajan and Bachtrog 2015; Walters et al. 2015), as well as species

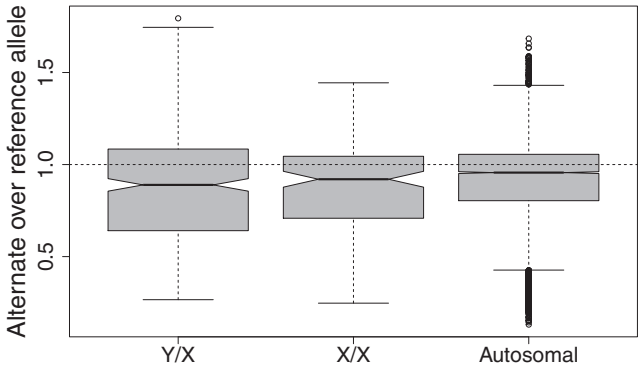

FIG. 5. Allele-specific expression for sex-linked genes that still retain the Y copy and for autosomal genes in *Rumex rothschildianus*. Ratio of alternate/reference is shown for males (corresponding to the ratio of Y allele over X allele), females and autosomal genes in both males and females as a control. Line represents theoretical expectation of equal expression of both alleles, although mapping bias can shift the expected ratio to slightly below one.

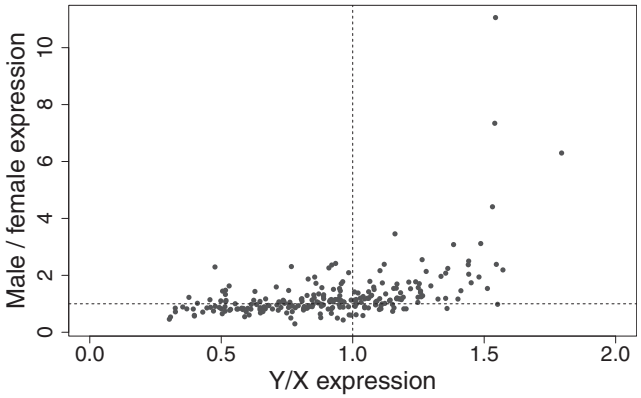

FIG. 6. Ratio of male gene expression over female gene expression plotted against Y/X allele-specific expression ratio in *Rumex rothschildianus*, showing that higher expression levels in males are driven by over-expression of the Y allele. Lines represent neutral expectation of 1:1 expression between males and females and between X and Y alleles.

with complete dosage compensation, such as mammals and *C. elegans* (Ercan et al. 2007; Pessia et al. 2012). In mammals, it has been shown that autosomal genes that interact with genes on the X chromosomes are down regulated accordingly (Julien et al. 2012). Presumably, something similar is occurring in other taxa that exhibit this pattern. A reduction in gene expression levels in the homogametic sex therefore appears to be a common, albeit counterintuitive, way of balancing within-genome expression.

Evolution of Allele-Specific Gene Expression

Under a model of adaptive gene silencing, the expression of Y alleles should be actively suppressed relative to X alleles, due either to adaptive evolution on the X, or slightly deleterious mutations on the Y. To investigate these possibilities, we measured allele-specific expression for X and Y alleles in males, as well as reference and alternate alleles in sex-linked genes in females and autosomal genes (fig. 5). In all cases, the ratio of alternate over reference allele is shifted below the 1:1 line due to mapping bias, but the difference is greatest between the X and Y alleles. If this was due to an active suppression of the Y due to either of the two models outlined above, the Y/X expression is expected to be lower in genes with higher X–Y d_N/d_S (where more deleterious mutations have accumulated), or in genes that showed evidence of positive selection occurring on the X chromosome. However, there was no correlation between Y/X expression ratio and X–Y d_N/d_S ($r_s = 0.01$, $P = 0.86$) and no difference in Y/X expression ratio between genes that showed evidence of positive selection on the X chromosome compared with genes that did not (based on results from PAML branch-site model likelihood ratio test for positive selection; Mann Whitney U test, $P = 0.96$). Additionally, there was no correlation between Y/X expression and time since recombination suppression, estimated by X–Y synonymous divergence ($r_s = -0.05$, $P = 0.41$). These findings suggest that the lower expression of the Y allele relative to the X is more likely to be a passive consequence of degeneration, rather than a result of adaptive suppression of Y chromosome expression.

A reduction in expression of Y alleles could lead to lower overall expression of sex-linked XY genes in males, if it was not accompanied by an equal increase in expression of the X allele, that is, dosage compensation of genes that have retained a Y copy. To investigate this, we looked at overall expression differences between males and females in XY genes (table 4). Sex-linked genes had a much higher percentage of genes showing significant differences in expression between the sexes compared with autosomal genes. However, differentially expressed genes are biased towards being higher in males (table 4 and supplementary fig. S5, Supplementary Material online). Moreover, the increase in gene expression in males did not arise from expression of a higher-quality X allele, but was instead the result of over-expression of the Y allele compared with the X allele (fig. 6). This suggests that the observed difference in levels of gene expression between the sexes in XY genes may be a result of male-specific adaptive Y alleles. To further investigate this possibility, we conducted a functional enrichment analysis of genes with biased Y-allele expression, and found that they were enriched for categories of genes that are likely to have different optima between the two sexes, such as negative regulation of flower development and reproductive structures (supplementary table S7, Supplementary Material online).

These putatively adaptive Y alleles did not necessarily arise on the Y chromosome, as they might have been present in the ancestral pseudo-autosomal region and therefore contributed to the suppression of recombination due to sexually antagonistic effects (Rice 1987a; Jordan and Charlesworth 2011). However, the highly skewed expression favoring the Y allele over the X allele must have occurred after recombination was suppressed. Although it is possible that mutations in regulatory regions are causing deleterious increases in expression level, as opposed to the more commonly reported

reduction in expression, high expression levels of degenerating genes are likely to be strongly selected against. Moreover, the enrichment of these Y-biased genes for functions that are likely to have different optima between the sexes suggests that this is an adaptive bias in allele expression. Overall, our results suggest that the Y chromosome is not only passively degenerating, but that it can also respond to male-specific selection. These processes also provide a mechanism by which selective sweeps, the last of the non-independence models, could be operating.

Conclusions

We observed extensive gene loss (or silencing) from the Y chromosome in *R. rothschildianus*, and we inferred that Y chromosome degeneration has occurred at a significantly faster rate in *R. rothschildianus* compared with *R. hastatulus*. This result demonstrates that closely related taxa with similar life histories and sex chromosomes of roughly similar ages can have contrasting rates of degeneration. We hypothesise that these differences may be driven by the presence of an ancient neo-Y chromosome in *R. rothschildianus* driving stronger selective interference over a longer timescale than in *R. hastatulus*.

The patterns observed in the loss of genes and changes in gene expression in *R. rothschildianus* can be used to evaluate possible models of Y chromosome degeneration. In particular, they do not support the idea that degeneration is driven by active silencing of the Y allele due to either adaptive evolution of the X or mildly deleterious mutations on the Y. These processes would be expected to suppress Y allele expression in genes undergoing extensive positive selection in the former or, arguably, more important genes in the latter, which is opposite to the pattern we found. Moreover, although analysis of allele-specific expression suggests Y alleles are expressed at a lower level than the X allele on average, there is no clear correlation with age of the sex-linked gene or its degree of constraint, which suggests this pattern may be a direct cause of deleterious mutations in regulatory regions, as opposed to any active suppression of particular genes. Finally, adaptive silencing would be expected to be accompanied by extensive dosage compensation, as otherwise the fitness cost of the loss of a gene copy would likely outweigh any benefit arising from suppression of a less fit allele. In *R. rothschildianus*, however, only a subset of genes showed evidence of dosage compensation.

Instead, the pattern observed in the types of genes that are lost or retained in *R. rothschildianus*, and in their expression, suggests Y chromosome degeneration is a direct consequence of interference between sites reducing the efficacy of selection. In particular, background selection could have a prominent role, as the genes that remain on the Y show evidence of being under stronger purifying selection. These genes are ancestrally more constrained and have higher expression than those that have been lost from the Y chromosome, they are more likely to be expressed in the haploid phase of the life cycle, and they are enriched for fundamental functional categories. There is also evidence of male-specific adaptation on Y allele expression, which could have resulted in selective sweeps. Therefore, both purifying and positive selection appear to be operating on the Y chromosome of *R. rothschildianus*, reducing the effective population size of non-recombining regions and increasing the speed of Muller's ratchet, ultimately leading to the degeneration and loss of more weakly selected linked genes.

Materials and Methods

RNA Sequencing and Transcriptome Assembly

Rumex rothschildianus is a rare dioecious annual, endemic to Israel (Rottenberg and Parker 2003). Material of *R. rothschildianus* used in this study was obtained from a bulk seed collection from the Tel Aviv Botanical Garden. We collected leaf tissue from the parents and from six female and six male offspring of two independent crosses of *R. rothschildianus*. We used leaf tissue as opposed to floral tissue because the latter involves considerable sex-specific gene expression associated with the development of male versus female flowers, which would complicate analyses unnecessarily. The crosses were carried out in isolation under glasshouse conditions to prevent cross-fertilisation from an individual other than the focal male parent. The second cross only had four male offspring due to an extreme sex ratio bias. We extracted total RNA using RNAeasy plant kit (Qiagen) following manufacturer recommended protocols. Illumina Hi-Seq 2,500 sequencing of 100 bp paired-end reads was carried out at The Centre for Applied Genomics (Toronto, Canada) multiplexed across four lanes, with one lane having fewer samples allowing increased coverage of the two individuals (one male and one female parent) subsequently used for *de novo* transcriptome assembly.

We assembled separate female and male reference transcriptomes *de novo* using one female and one male parent. Prior to assembly the reads were passed through a quality filter to remove low quality read pairs with over 10% Ns, over 50% low quality bases, or shorter than 50 bp long. We used the perl script VelvetOptimiser.pl (v. 2.2.4) available from GitHub to determine the optimal kmer size (45 for the female assembly; 27 for the male assembly) and the assembly was carried out using Velvet (v. 1.2.09; Zerbino and Birney 2008) followed by Oases (v. 0.2.08; Schulz et al. 2012). The output from Oases contained several isoforms per transcript, and in each case we chose the longest as the representative transcript. This pipeline resulted in a female reference transcriptome of 31,408 contigs ($N50 = 1{,}894$) and a male reference transcriptome of 58,358 contigs ($N50 = 1{,}766$). We obtained reference transcriptomes for *R. hastatulus* and the hermaphrodite *R. bucephalophorus* from Hough et al. (2014).

Identification of Sex-Linked Genes

To identify genes located on the sex chromosomes in *R. rothschildianus* we mapped all individuals to the female reference transcriptome using Burrows–Wheeler Aligner (v. 0.7.8-r455; Li and Durbin 2009) followed by Stampy (v.1.0.20; Lunter and Goodson 2011) to map more divergent reads. We carried out a pre-processing pipeline of

the mapped bam files using Picard tools (http://picard.sourceforge.net) as described in the GATK 'best practices' (DePristo et al. 2011; Van der Auwera et al. 2013). We then used GATK haplotype caller followed by unified genotyper to call variants, implemented with the recommended quality filters (McKenna et al. 2010).

The resulting VCF files contained variant information for the parents and male and female offspring of the two crosses. We parsed these files searching for SNPs segregating in diagnostic patterns to form three different datasets: genes located on the X and Y chromosomes (hereafter 'sex-linked XY genes'); genes located on the X and missing from the Y chromosome ('sex-linked hemizygous genes'); and genes located on the autosomes (for details see supplementary material S1, Supplementary Material online). The hemizygosity of genes classified as sex-linked hemizygous was confirmed in several ways (for details see supplementary material S1, Supplementary Material online). We obtained lists of sex-linked XY genes, sex-linked hemizygous and autosomal genes for *R. hastatulus* from Hough et al. (2014).

Identification of Orthologs, Construction of Alignments and Phylogenetic Trees

To identify orthologs between *R. rothschildianus* and two closely related species *R. hastatulus* and *R. bucephalophorus*, we first identified the longest open reading frame (ORF) of each transcript from each species using the program getorf from the software suite EMBOSS (v. 6.4.0.0; Rice et al. 2000). We then performed a reciprocal nucleotide BLAST search (Altschul et al. 1990) between each species pair, and used three-way best hits to determine orthology.

To place *R. rothschildianus* in the *Rumex* phylogeny we downloaded 10 *R. acetosa* genes and the transcriptome of *Fagopyrum esculentum* (Polygonaceae), for use an out-group, from the GenBank nucleotide database. We identified the longest ORF for each sequence of both species and used reciprocal nucleotide blast searches of *R. acetosa* and reciprocal protein blast of *F. esculentum* against the longest ORFs of the three reference transcriptomes to determine orthology.

We constructed alignments of orthologous ORFs by first aligning the amino acid sequences using the program Muscle (v3.8.3; Edgar 2004), which were then used to guide the alignment of nucleotide sequences using the program RevTrans (v. 1.4; Wernersson and Pedersen 2003). We constructed maximum likelihood phylogenetic trees using the program RAxML (v. 8.2.4; Stamatakis 2014) and bootstrapped the trees to get measures of support for each node using the same program. The trees were rooted using *F. esculentum* as an outgroup.

Molecular Evolution on the X and Y Chromosomes

To investigate molecular evolution of the X and Y sequences, it is necessary to have phased consensus sequences. This was carried out using a likelihood approach developed and tested in Hough et al (2014). Briefly, this involves an R script that assesses sex-linkage of SNPs using a likelihood ratio, and generates a consensus Y sequence given the female reference as a consensus X sequence and the VCF file of variants mapped to it. We modified this script slightly to include the evolution of indels, but other details remain the same and are described in Hough et al. (2014). Because of the modification of the script we generated consensus phased X and Y sequences for both *R. rothschildianus* and *R. hastatulus*.

We identified the longest ORF separately for the X and Y consensus sequences and constructed pairwise alignments of the X and Y sequences following the same method described above. We calculated pairwise X–Y sequence divergence using the yn00 program implemented in PAML (v. 4.8; Yang 2007). For each species, we aligned the phased XY sequences to their orthologs in the other species, and analysed these four-way alignments using the codeml program in PAML. Input tree files were unrooted, and in each case the X and Y sequences formed a monophyletic group, with a trifurcation at the root. We used an initial run of the M0 'one-ratio' model to provide the initial branch lengths for the trees used in the subsequent analyses. We ran several models: 1) Branch models allowing the rate of the X or Y sequence to vary compared with the 'background' rate; 2) A free ratio model allowing rates to vary freely for each branch; 3) Site models allowing rates to vary between sites, which is used to detect signals of positive selection when particular sites (codons) have d_N/d_S ratio >1.

We ran the same PAML models on sex-linked hemizygous genes and autosomal genes in *R. rothschildianus* and their orthologs in the other species. The analyses looking for evidence of positive selection included the out-group *F. esculentum*, so were also based on alignments of four sequences. We generated tree files separately for each gene using the program RAxML, as the phylogenetic relationship between *R. rothschildianus*, *R. hastatulus* and *R. bucephalophorus* is not fully resolved.

Assessing Inter-Specific Homology of Sex Chromosomes

To assess homology of sex chromosomes in *R. rothschildianus* and *R. hastatulus* we carried out a thorough comparison of all sex-linked genes in the two species. We used datasets of different stringency (with different numbers of SNPs required to define the gene as sex-linked) and different age (using X–Y synonymous divergence as a measure of age), looking for the presence of more orthologous genes present in each dataset than one would expect by chance. We calculated expected values simply by random expectation given the number of orthologs between the two species, and the number of orthologs in each gene class. We calculated significance by two-tailed random permutation tests.

Analyses of Gene and Allele-Specific Expression

We measured gene expression with the python package HTSeq (Anders et al. 2015) using the longest ORF as the defined region for expression counts. We analysed the output of raw counts using the R package DESeq2 (v. 1.11.45; Love et al. 2014), which uses a negative binomial distribution model to test for significantly different levels of gene expression between groups. We identified genes with significantly different expression between: 1) males and females of *R.*

rothschildianus, and 2) present expression levels in *R. rothschildianus* and ancestral expression levels, estimated by using the average of male and female expression in *R. hastatulus*. Because hemizygous genes are expected to have lower levels of expression in males relative to females, we carried out the between-individual normalisation using only the autosomal genes, which have no reason to be biased overall in their expression between the sexes. We then used the individual size factors obtained from the autosomal data to transform the raw counts of sex-linked genes into normalised counts. None of the hemizygous genes in *R. rothschildianus* are hemizygous in *R. hastatulus*, allowing *R. hastatulus* to be used for estimation of ancestral expression levels. The close correlation between expression in *R. rothschildianus* and *R. hastatulus* ($r_s = 0.60$, $P < 10^{-15}$) validates this approach.

We obtained allele-specific expression from allele read counts in the VCF files. SNPs with low coverage (< 8 reads) were removed, as these are more likely to give highly skewed ratios. We calculated the ratio of alternate over reference allele counts for each heterozygous individual: in males this corresponds to Y over X expression, whereas in females this corresponds to expression of the X inherited from the father over the X inherited from the mother. We then averaged these ratios across individuals and across SNPs in each gene, giving an average alternative/reference value per gene. We calculated the same alternate over reference allele count ratio for autosomal genes to provide the null expectation, as mapping bias lowers the actual expected ratio to below the theoretical null expectation of 1:1.

Generating Gene Lists for Gametophyte and Sporophyte Expression

Lists of genes expressed during different life cycle stages were obtained for *A. thaliana* and *N. tabacum* from Borges et al. (2008) and Hafidh et al. (2012). We divided these into two sets of genes: those expressed at any point during the haploid gametophyte phase of the life cycle, and those only expressed during the diploid sporophyte phase (for details see supplementary material S1, Supplementary Material online). We analysed the number of orthologous sex-linked hemizygous genes and sex-linked XY genes present in each dataset (expressed or not expressed during the haploid phase of the life cycle) using a Fisher's exact test. We also carried out a second test between all sex-linked genes compared with all autosomal genes.

Functional Enrichment Analysis

We used the Database for Annotation, Visualization and Discovery (DAVID; Huang et al. 2009) to identify significant overrepresentation of particular functional categories in sex-linked XY genes compared with hemizygous genes. We also carried out a second functional enrichment test on genes showing biased expression of the Y allele in males. Gene ontology was defined using orthology to *A. thaliana*, which has the most comprehensive gene annotation of any plant species. The background from which to base the enrichment analysis involved all *R. rothschildianus*–*A. thaliana* orthologs. DAVID provides functional annotation clustering to facilitate interpretation, with enrichment scores above 1 being significant. Significance of enrichment of particular functional categories was determined using a modified (more conservative) Fisher's exact test.

Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

Acknowledgments

We thank Yuval Sapir for providing seed of *Rumex rothschildianus*, Wei Wang and Jesse Hollister for sharing scripts, Bill Cole for glasshouse help, Yunchen Gong for server support, and two anonymous reviewers for helpful comments on the manuscript. This work was supported by Discovery Grants from the Natural Sciences and Engineering Research Council of Canada to S.C.H.B. and S.I.W. D.C. was supported by a Connaught Fellowship from the University of Toronto.

References

Akagi T, Henry IM, Tao R, Comai L. 2014. A Y-chromosome-encoded small RNA acts as a sex determinant in persimmons. *Science* 346:646–650.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–410.

Anders S, Pyl PT, Huber W. 2015. HTSeq – a Python framework to work with high-throughput sequencing data. *Bioinformatics* 21:166–169.

Bachtrog D. 2004. Evidence that positive selection drives Y-chromosome degeneration in *Drosophila miranda*. *Nat Genet* 36:518–522.

Bachtrog D. 2008. The temporal dynamics of processes underlying Y chromosome degeneration. *Genetics* 179:1513–1525.

Bachtrog D. 2013. Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. *Nat Rev Genet* 14:113–124.

Bachtrog D, Gordo I. 2004. Adaptive evolution of asexual populations under Muller's ratchet. *Evolution* 58:1403–1413.

Bachtrog D, Kirkpatrick M, Mank JE, McDaniel SF, Pires JC, Rice WR, Valenzuela N. 2011. Are all sex chromosomes created equal?. *Trends Genet* 27:350–357.

Bellott DW, Hughes JF, Skaletsky H, Brown LG, Pyntikova T, Cho T-J, Koutseva N, Zaghlul S, Graves T, Rock S, et al. 2014. Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators. *Nature* 508:494–499.

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

Błocka-Wandas M, Sliwinska E, Grabowska-Joachimiak A, Musial K, Joachimiak AJ. 2007. Male gametophyte development and two different DNA classes of pollen grains in *Rumex acetosa* L., a plant with an XX/XY_1Y_2 sex chromosome system and a female-biased sex ratio. *Sex Plant Reprod* 20:171–180.

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

Bull JJ. 1983. Evolution of sex determining mechanisms. Menlo Park (CA): Benjamin-Cummings Publishing Company.

Carvalho AB, Koerich LB, Clark AG. 2009. Origin and evolution of Y chromosomes: *Drosophila* tales. *Trends Genet* 25:270–277.

Charlesworth B. 1978. Model for evolution of Y chromosomes and dosage compensation. *Proc Natl Acad Sci USA* 75:5618–5622.

Charlesworth B, Charlesworth D. 2000. The degeneration of Y chromosomes. *Philos Trans R Soc Lond B* 355:1563–1572.

Charlesworth B, Coyne JA, Barton NH. 1987. The relative rates of evolution of sex chromosomes and autosomes. *Am Nat* 130:113–146.

Charlesworth B, Morgan MT, Charlesworth D. 1993. The effect of deleterious mutations on neutral molecular variation. *Genetics* 134:1289–1303.

Chen S, Zhang G, Shao C, Huang Q, Liu G, Zhang P, Song W, An N, Chalopin D, Volff J-N, et al. 2014. Whole-genome sequence of a flatfish provides insights into ZW sex chromosome evolution and adaptation to a benthic lifestyle. *Nat Genet* 46:253–260.

Chibalina MV, Filatov DA. 2011. Plant Y chromosome degeneration is retarded by haploid purifying selection. *Curr Biol* 21:1475–1479.

Comeron JM, Williford A, Kliman RM. 2008. The Hill–Robertson effect: evolutionary consequences of weak selection and linkage in finite populations. *Heredity* 100:19–31.

Correns C. 1928. Bestimmung, Vererbung und Verteilung des Geschlechtes bei den höheren Pflanzen. *Vererbungswissenschaft* 2:1–138.

Cuñado N, Navajas-Pérez R, de la Herrán R, Ruiz Rejón C, Ruiz Rejón M, Santos J, Garrido-Ramos M. 2007. The evolution of sex chromosomes in the genus *Rumex* (Polygonaceae): identification of a new species with heteromorphic sex chromosomes. *Chromosome Res* 15:825–832.

DePristo M, Banks E, Poplin R, Garimella K, Maguire J, Hartl C, Philippakis A, del Angel G, Rivas MA, Hanna M, et al. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 43:491–498.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucl Acids Res* 32:1792–1797.

Ellegren H, Hultin-Rosenberg L, Brunström B, Dencker L, Kultima K, Scholz B. 2007. Faced with inequality: chicken do not have a general dosage compensation of sex-linked genes. *BMC Biol* 5:40.

Ercan S, Giresi PG, Whittle CM, Zhang X, Green RD, Lieb JD. 2007. X chromosome repression by localization of the *C. elegans* dosage compensation machinery to sites of transcription initiation. *Nat Genet* 39:403–408.

Field DL, Pickup M, Barrett SCH. 2012a. Comparative analyses of sex-ratio variation in dioecious flowering plants. *Evolution* 67:661–672.

Field DL, Pickup M, Barrett SCH. 2012b. The influence of pollination intensity on fertilization success, progeny sex ratio, and fitness in a wind-pollinated, dioecious plant. *Int J Plant Sci* 173:184–191.

Fisher R. 1930. The genetical theory of natural selection. Oxford: Clarendon Press.

Geraldes A, Hefer CA, Capron A, Kolosova N, Martinez-Nuñez F, Soolanayakanahally RY, Stanton B, Guy RD, Mansfield SD, Douglas CJ, et al. 2015. Recent Y chromosome divergence despite ancient origin of dioecy in poplars (*Populus*). *Mol Ecol* 24:3243–3256.

Hafidh S, Breznenová K, Růžička P, Feciková J, Čapková V, Honys D. 2012. Comprehensive analysis of tobacco pollen transcriptome unveils common pathways in polar cell expansion and underlying heterochronic shift during spermatogenesis. *BMC Plant Biol* 12:24.

Hill W, Robertson A. 1966. The effect of linkage on limits to artificial selection. *Genet Res* 8:269–294.

Honys D, Twell D. 2004. Transcriptome analysis of haploid male gametophyte development in *Arabidopsis*. *Genome Biol* 5:R85.

Hough J, Hollister JD, Wang W, Barrett SCH, Wright SI. 2014. Genetic degeneration of old and young Y chromosomes in the flowering plant *Rumex hastatulus*. *Proc Natl Acad Sci USA* 111:7713–7718.

Hough J, Immler S, Barrett SCH, Otto SP. 2013. Evolutionarily stable sex ratios and mutation load. *Evolution* 67:1915–1925.

Huang DW, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. *Nat Protoc* 4:44–57.

Jordan CY, Charlesworth D. 2011. The potential for sexually antagonistic polymorphism in different genome regions. *Evolution* 66:505–516.

Julien P, Brawand D, Soumillon M, Necsulea A, Liechti A, Schütz F, Daish T, Grützner F, Kaessmann H. 2012. Mechanisms and evolutionary patterns of mammalian and avian dosage compensation. *PLoS Biol* 10:e1001328.

Kaiser VB, Charlesworth B. 2010. Muller's ratchet and the degeneration of the *Drosophila miranda* neo-Y chromosome. *Genetics* 185:339–348.

Kaiser VB, Zhou Q, Bachtrog D. 2011. Nonrandom gene loss from the *Drosophila miranda* neo-Y chromosome. *Genome Biol Evol* 3:1329–1337.

Kejnovský E, Michalovova M, Steflova P, Kejnovska I, Manzano S, Hobza R, Kubat Z, Kovarik J, Jamilena M, Vyskot B. 2013. Expansion of microsatellites on evolutionary young Y chromosome. *PLoS ONE* 8:e45519.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler Transform. *Bioinformatics* 25:1754–1760.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:550.

Lunter G, Goodson M. 2011. Stampy: a statistical algorithm for sensitive and fast mapping of illumina sequence reads. *Genome Res* 21:936–939.

Mahajan S, Bachtrog D. 2015. Partial dosage compensation in Strepsiptera, a sister group of beetles. *Genome Biol Evol* 7:591–600.

Malone JH, Cho D-Y, Mattiuzzo NR, Artieri CG, Jiang L, Dale RK, Smith HE, McDaniel J, Munro S, Salit M, et al. 2012. Mediation of *Drosophila* autosomal dosage effects and compensation by network interactions. *Genome Biol* 13:R28.

Mank JE. 2013. Sex chromosome dosage compensation: definitely not for everyone. *Trends Genet* 29:677–683.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20:1297–1303.

Ming R, Bendahmane A, Renner SS. 2011. Sex chromosomes in land plants. *Annu Rev Plant Biol* 62:485–514.

Muller HJ. 1964. The relation of recombination to mutational advance. *Mutat Res* 1:2–9.

Mullon C, Wright AE, Reuter M, Pomiankowski A, Mank JE. 2015. Evolution of dosage compensation under sexual selection differs between X and Z chromosomes. *Nat Commun* 6:7720.

Muyle A, Zemp N, Deschamps C, Mousset S, Widmer A, Marais GAB. 2012. Rapid de novo evolution of X chromosome dosage compensation in *Silene latifolia*, a plant with young sex chromosomes. *PLoS Biol* 10:e1001308.

Navajas-Pérez R, de la Herrán R, López González G, Jamilena M, Lozano R, Ruiz Rejón C, Ruiz Rejón M, Garrido-Ramos MA. 2005. The evolution of reproductive systems and sex-determining mechanisms within *Rumex* (Polygonaceae) inferred from nuclear and chloroplastidial sequence data. *Mol Biol Evol* 22:1929–1939.

Orr HA, Kim Y. 1998. An adaptive hypothesis for the evolution of the Y chromosome. *Genetics* 150:1693–1698.

Ossowski S, Schneeberger K, Lucas-Lledó JI, Warthmann N, Clark RM, Shaw RG, Weigel D, Lynch M. 2010. The rate and molecular spectrum of spontaneous mutations in *Arabidopsis thaliana*. *Science* 327:92–94.

Papadopulos AST, Chester M, Ridout K, Filatov DA. 2015. Rapid Y degeneration and dosage compensation in plant sex chromosomes. *Proc Natl Acad Sci USA* 112:13021–13026.

Parker JS, Clark MS. 1991. Dosage sex-chromosome systems in plants. *Plant Sci* 80:79–92.

Pessia E, Makino T, Bailly-Bechet M, McLysaght A, Marais GAB. 2012. Mammalian X chromosome inactivation evolved as a dosage-compensation mechanism for dosage-sensitive genes on the X chromosome. *Proc Natl Acad Sci USA* 109:5346–5351.

Potrzebowski L, Vinckenbosch N, Marques AC, Chalmel F, Jégou B, Kaessmann H. 2008. Chromosomal gene movements reflect the recent origin and biology of therian sex chromosomes. *PLoS Biol* 6:709–716.

Quesada del Bosque M, Navajas-Pérez R, Panero J, Fernández-González A, Garrido-Ramos M. 2011. A satellite DNA evolutionary analysis in the North American endemic dioecious plant *Rumex hastatulus* (Polygonaceae). *Genome* 54:253–260.

Rice WR. 1987a. The accumulation of sexually antagonistic genes as a selective agent promoting the evolution of reduced recombination between primitive sex chromosomes. *Evolution* 41:911–914.

Rice WR. 1987b. Genetic hitchhiking and the evolution of reduced genetic activity of the Y sex chromosome. *Genetics* 116:161–167.
Rice P, Longden I, Bleasby A. 2000. EMBOSS: the European molecular biology open software suite. *Trends Genet* 16:276–277.
Rottenberg A, Parker JS. 2003. Conservation of the critically endangered *Rumex rothschildianus* as implied from AFLP diversity. *Biol Conserv* 114:299–303.
Schulz MH, Zerbino DR, Vingron M, Birney E. 2012. Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics* 28:1086–1092.
Singh ND, Koerich LB, Carvalho AB, Clark AG. 2014. Positive and purifying selection on the *Drosophila* Y chromosome. *Mol Biol Evol* 31:2612–2623.
Smith BW. 1969. Evolution of sex-determining mechanisms in *Rumex*. *Chromosomes Today* 2:172–182.
Stamatakis A. 2014. RAxML Version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313.
Steflova P, Hobza R, Vyskot B, Kejnovsky E. 2013. Strong accumulation of chloroplast DNA in the Y chromosomes of *Rumex acetosa* and *Silene latifolia*. *Cytogenet Genome Res* 141:59–65.
Steflova P, Tokan V, Vogel I, Lexa M, Macas J, Novak P, Hobza R, Vyskot B, Kejnovsky E. 2013. Contrasting patterns of transposable element and satellite distribution on sex chromosomes (XY_1Y_2) in the dioecious plant *Rumex acetosa*. *Genome Biol Evol* 5:769–782.
Stehlik I, Barrett SCH. 2006. Pollination intensity influences sex ratios in dioecious *Rumex nivalis*, a wind-pollinated plant. *Evolution* 60:1207–1214.
Uebbing S, Künstner A, Mäkinen H, Ellegren H. 2013. Transcriptome sequencing reveals the character of incomplete dosage compensation across multiple tissues in flycatchers. *Genome Biol Evol* 5:1555–1566.
Van der Auwera GA, Carneiro M, Hartl C, Poplin R, del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, et al. 2013. From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. *Curr Protoc Bioinf* 43:1–33.
VanBuren R, Zeng F, Chen C, Zhang J, Wai CM, Han J, Aryal R, Gschwend AR, Wang J, Na J, et al. 2015. Origin and domestication of papaya Y^h chromosome. *Genome Res* 25:524–533.
Veyrunes F, Waters PD, Miethke P, Murchison EP, Kheradpour P, Sachidanandam R, Park J, Semyonov J, Chang CL, Whittington CM, et al. 2008. Bird-like sex chromosomes of platypus imply recent origin of mammal sex chromosomes. *Genome Res* 18:965–973.
Vicoso B, Bachtrog D. 2009. Progress and prospects toward our understanding of the evolution of dosage compensation. *Chromosome Res* 17:585–602.
Vicoso B, Bachtrog D. 2015. Numerous transitions of sex chromosomes in Diptera. *PLoS Biol* 13:e1002078.
Vicoso B, Emerson JJ, Zektser Y, Mahajan S, Bachtrog D. 2013. Comparative sex chromosome genomics in snakes: differentiation, evolutionary strata, and lack of global dosage compensation. *PLoS Biol* 11:e1001643.
Walters JR, Hardcastle TJ, Jiggins CD. 2015. Sex chromosome dosage compensation in *Heliconius* butterflies: global yet still incomplete? *Genome Biol Evol* 7:2545–2559.
Wernersson R, Pedersen AG. 2003. RevTrans – constructing alignments of coding DNA from aligned amino acid sequences. *Nucl Acids Res* 31:3537–3539.
White MA, Kitano J, Peichel CL. 2015. Purifying selection maintains dosage-sensitive genes during degeneration of the threespine stickleback Y chromosome. *Mol Biol Evol* 32:1981–1995.
Wilby AS, Parker JS. 1988. Recurrent patterns of chromosome variation in a species group. *Heredity* 61:55–62.
Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24:1586–1591.
Yang Z, Nielsen R. 2000. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol Biol Evol* 12:32–43.
Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 18:821–829.
Zhou Q, Bachtrog D. 2012. Chromosome-wide gene silencing initiates Y degeneration in *Drosophila*. *Curr Biol* 22:522–525.
Zhou Q, Zhang J, Bachtrog D, An N, Huang Q, Jarvis ED, Gilbert MTP, Zhang G. 2014. Complex evolutionary trajectories of sex chromosomes across bird taxa. *Science* 346:1246338.
Zimmer F, Harrison PW, Dessimoz C, Mank JE. 2016. Compensation of dosage-sensitive genes on the chicken Z chromosome. *Genome Biol Evol* 8:1233–1242.