Recombination landscape dimorphism and sex chromosome evolution in the dioecious plant Rumex hastatulus

Rifkin, Joanna L.

2022-05-09

Rifkin , J L , Hnatovska , S , Yuan , M , Sacchi , B M , Choudhury , B , Gong , Y , Rastas , P , Barrett , S C H & Wright , S 2022 , ' Recombination landscape dimorphism and sex chromosome evolution in the dioecious plant Rumex hastatulus ' , Philosophical Transactions of the Royal Society. Biological Sciences , vol. 377 , no. 1850 , 20210226 . https://doi.org/10.1098/rstb.

http://hdl.handle.net/10138/353828
https://doi.org/10.1098/rstb.2021.0226

PHILOSOPHICAL TRANSACTIONS
OF THE ROYAL SOCIETY B

BIOLOGICAL SCIENCES

Recombination landscape dimorphism contributes to sex chromosome evolution in the dioecious plant Rumex hastatulus

Journal:	*Philosophical Transactions B*
Manuscript ID	Draft
Article Type:	Research
Date Submitted by the Author:	n/a
Complete List of Authors:	Rifkin, Joanna; Univ of Toronto, Ecology & Evolutionary Biology Hnatovska, Solomiya; Univ of Toronto, Ecology & Evolutionary Biology Yuan, Meng; Univ of Toronto, Ecology & Evolutionary Biology Sacchi, Bianca; Univ of Toronto, Ecology & Evolutionary Biology Choudhury, Baharul; Queen's University Gong, Yunchen; Univ of Toronto, Ecology & Evolutionary Biology Rastas, Pasi; University of Helsinki, Institute of Botechnology Barrett, Spencer; Univ of Toronto, Ecology & Evolutionary Biology; Wright, Stephen; University of Toronto,
Issue Code (this should have already been entered and appear below the blue box, but please contact the Editorial Office if it is not present):	SEXPLANTS
Subject:	Evolution < BIOLOGY, Genetics < BIOLOGY, Plant Science < BIOLOGY, Genomics < BIOLOGY
Keywords:	dioecy, evolution, gametophytic competition, heterochiasmy, recombination, sex chromosomes

Author-supplied statements

Relevant information will appear here if provided.

Ethics

Does your article include research that required ethical approval or permits?:
This article does not present research with ethical considerations

Statement (if applicable):
CUST_IF_YES_ETHICS :No data available.

Data

It is a condition of publication that data, code and materials supporting your paper are made publicly available. Does your paper present new data?:
Yes

Statement (if applicable):
CUST_IF_YES_DATA :No data available.

Conflict of interest

I/We declare we have no competing interests

Statement (if applicable):
CUST_STATE_CONFLICT :No data available.

Authors' contributions

This paper has multiple authors and our individual contributions were as below

Statement (if applicable):
Joanna L. Rifkin coordinated and performed descriptive genomics analyses and linear models and participated in conceiving and writing the paper
Solomiya Hnatovska performed the TE annotation
Meng Yuan performed differential expression analyses
Bianca Sacchi performed differential expression analyses and SNP-calling
Baharul Choudhury coordinated plant growth and nucleic acid expression
Yunchen Gong created the transcriptome annotation
Pasi Rastas generated the linkage maps and the improved genome assembly
Spencer C.H. Barrett provided funding, development, and editing
Stephen I. Wright conceived, funded, coordinated, and coauthored the paper

Phil. Trans. R. Soc. B. article template

PHILOSOPHICAL
TRANSACTIONS B

Phil. Trans. R. Soc. B.
doi:10.1098/not yet assigned

Recombination landscape dimorphism contributes to sex chromosome evolution in the dioecious plant *Rumex hastatulus*

Joanna L. Rifkin[1*], Solomiya Hnatovska[1], Meng Yuan[1], Bianca M. Sacchi[1], Baharul I. Choudhury[1,2], Yunchen Gong[3], Pasi Rastas[4], Spencer C.H. Barrett[1], Stephen I. Wright[1,3*]

1 – Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, ON Canada

2 – Department of Biology, Queen's University, Kingston, ON, Canada

3 – Centre for Analysis of Genome Evolution and Function, University of Toronto, Toronto, ON Canada

4 – Institute of Biotechnology, University of Helsinki, Finland

Keywords: dioecy, evolution, gametophytic competition, heterochiasmy, recombination, sex chromosomes

*Corresponding authors' contact information: joanna.rifkin@utoronto.ca, joannarifkin@gmail.com, stephen.wright@utoronto.ca

Summary

There is growing evidence across diverse taxa for sex differences in the genomic landscape of recombination, but the causes and consequences of these differences remain poorly understood. Strong recombination landscape dimorphism between the sexes could have important implications for the dynamics of sex chromosome evolution and turnover because low recombination in the heterogametic sex can help favour the spread of sexually antagonistic alleles. Here, we present a sex-specific linkage map and revised genome assembly of *Rumex hastatulus*, representing the first characterization of sex differences in recombination landscape in a dioecious plant. We provide evidence for strong sex differences in recombination, with pericentromeric regions of highly suppressed recombination in males that cover over half of the genome. These differences are found on autosomes as well as sex chromosomes, suggesting that pre-existing differences in recombination may have contributed to sex chromosome formation and divergence. Analysis of segregation distortion suggests that haploid selection due to pollen competition occurs disproportionately in regions with low male recombination. Our results are consistent with the hypothesis that sex differences in the recombination landscape contributed to the formation of a large heteromorphic pair of sex chromosomes, and that pollen competition is an important determinant of recombination dimorphism.

Introduction

The distribution of rates of recombination along chromosomes (recombination landscape [1]) shapes many aspects of evolutionary genetics, including the efficacy of natural selection [2], genome structure [3], and the dynamics of reproductive isolation [4]. Rates of recombination can vary between species, between and within chromosomes, and between male and female meiosis in both dioecious/gonochoristic and hermaphroditic species [5–7]. We refer to this phenomenon as 'sex differences in the recombination landscape' [1]. Although sex differences in the rate and distribution of recombination appear to be widespread and variable, the causes and consequences of this variation have only recently been investigated in detail [1,6,8].

Many components of evolutionary processes depend on the sex-averaged rate of recombination. Nevertheless, sex differences in recombination (heterochiasmy) in dioecious populations can have important consequences for the evolution of sex chromosomes. This is because on the sex chromosome restricted to the heterogametic sex (i.e. the Y or W chromosome), sex-specific recombination landscapes entirely control the rate of recombination, and thereby influence the scale of recombination suppression surrounding a sex-determining region (SDR) [1]. Recently, heterochiasmy has been proposed as an important factor in maintaining sexually antagonistic (SA) variants on the sex chromosomes even in the absence of recombination modifiers [9]. In particular, SA alleles can spread more easily through populations because of pre-existing male-specific suppression of recombination, rather than recombination suppression evolving as a secondary consequence of the segregation of SA alleles [9]. Variation among species in patterns of heterochiasmy could thus be an important general determinant of the evolution of sex chromosomes and their turnover [1], potentially contributing to differences among lineages in the likelihood of the formation of heteromorphic sex chromosomes, the maintenance of sexually antagonistic polymorphisms and the size of the SDR.

Several patterns are evident in the characteristics of sexual dimorphism in the recombination landscape. First, although data are limited, many eukaryotes have recombination rates biased towards the tips of chromosomes in male meiosis, whereas female recombination rates are more likely to be either elevated towards the centromeres or are more uniform across the chromosome [1]. In hermaphroditic plants, three of five taxa studied show this pattern [10–12], although in maize there was limited evidence for large-scale differences in recombination between male and female meiosis [13], and in an interspecific cross between *Solanum esculentum and S. pennellii* recombination in male gametes was reduced genome-wide compared with female gametes [14]. Preliminary analysis of genetic maps in the dioecious *Mercurialis annua* do not suggest major sex differences in recombination rates [15], although the genomic context of these maps is still being investigated. In general, however, recombination rate landscape dimorphism has not yet been investigated in dioecious plants, limiting our understanding of its potential role in the evolutionary dynamics of plant sex chromosomes.

Many species have convergently evolved tip-biased recombination in male meiosis [1], but the reasons for this pattern are unclear. Both non-adaptive and adaptive explanations have been proposed. If recombination landscape differences are not adaptive, they may simply result from mechanistic differences between the process of female and male meiosis [1]. Adaptive hypotheses include sexually antagonistic selection favoring tighter linkage between sex-specific genes and regulatory elements [1], selection favoring recombination near the centromere in female meiosis to remove meiotic drive alleles [7], and epistatic haploid selection on male gametes and gametophytes [6]. In plants, evidence that female recombination rates are elevated relative to male recombination rates in outcrossing species compared with selfing species [6] is consistent with models of both female meiotic drive and male gametophytic selection, as both are expected to be more intense with higher rates of outcrossing [6,7]. Disentangling these alternatives is challenging and will require more comparative information on sex-specific recombination in both hermaphroditic and dioecious taxa.

Rumex hastatulus is a dioecious, wind-pollinated plant with heteromorphic sex chromosomes [16,17]. Recent genome sequencing combined with high marker-density linkage mapping has revealed that the SDR is embedded within a very large genomic region of highly suppressed recombination [18]. Evidence for similarly large non-recombining regions in the pericentromeric regions of autosomes suggested that pre-existing recombination suppression may have contributed to the formation of large heteromorphic sex chromosomes in *R. hastatulus* [18]. However, this study measured sex-averaged recombination rates, limiting our ability to investigate the potential role of heterochiasmy in sex chromosome formation and maintenance. With earlier evidence for an important role for gametophytic selection on the sex ratio in this species [19,20], the influence of pollen competition in the evolution of the sex chromosomes [21], and indications of frequent male and female transmission distortion in related dioecious *Rumex* taxa [22], there is a strong likelihood that haploid selection in males and/or females may contribute to sex-specific selection favouring sexual dimorphism in recombination landscapes in this system.

Here, we explore the potential importance of heterochiasmy for the evolution of sex chromosomes and test hypotheses concerning the evolutionary forces favouring sex-specific recombination rate differences in *R. hastatulus*. Using a sex-specific linkage map and corrected draft genome assembly, we first determine whether *R. hastatulus* shows evidence for heterochiasmy and other sex differences in recombination landscape and compare this pattern between the sex chromosome and the autosomes. We then examine the correlates of male and female recombination rates genome-wide and use this information to explore the potential role of sexual antagonism, haploid selection, and meiotic drive as evolutionary drivers of sexual dimorphism in the recombination landscape.

Methods

Linkage mapping and genome assembly

We generated a mapping population from a cross between a male and female derived from a single population collected in Rosebud, TX [20]. Seeds from the field collection were grown in the glasshouse and at onset of flowering one male and one female individual were randomly paired for a controlled cross to develop the F_1 generation. Paired plants were immediately moved into miniature crossing chambers [23] to avoid pollen contamination from other plants growing in the same glasshouse, and F_1 seeds were harvested after maturation. To obtain tissue from F_1 plants, we sterilized seeds using 5% (V/V) bleach and germinated them on filter paper in refrigerated petri dishes. After germination, we transplanted seedlings into six-inch plastic pots containing a 3:1 ratio of Promix soil and sand and a slow-release fertilizer (Nutricote, 14:13:13, 300mL per 60lbs) and placed them in a glasshouse at the University of Toronto, St. George campus. We watered plants every other day, and their positions on benches were randomized weekly. On day 43 or 44 after transplant, between 10:00 and 12:00, we collected and flash-froze 30mg of leaf tissue for RNA extraction using liquid nitrogen. When plants flowered, we phenotyped for sex. We used Spectrum Plant Total RNA Kits (Sigma Aldrich) for RNA extraction. The sequenced library included 188 individuals: 102 female offspring, 84 male offspring, and three replicate samples of each parent.

For library preparation and sequencing, we sent our RNA samples to the Centre d'expertise et de services Génome Québec (CES, McGill University, Montréal, QC, Canada). CES prepared libraries using NEBNext library prep kits and sequenced them on a NovaSeq6000 S4 PE100. Sequencing resulted in a total of 3.1 billion reads (3,060,570,370), with between 10 and 49 million reads per sample (mean 15,940,471, median 14,548,490). Raw sequence has been deposited on the Sequence Read Archive (SRA) under the accession number PRJNA692236 (embargoed until July 1, 2022 or publication).

We aligned our raw sequencing reads to the *R. hastatulus* Dovetail draft genome assembly [18] using Star 2-pass version 2.7.6 [24,25]. We processed the aligned files to sort, mark PCR duplicates, and (splitNCigar reads) using PicardTools (http://broadinstitute.github.io/picard/) and the Genome Analysis Tool Kit [26].

We initially generated a linkage map using Lep-Map3 [27] from reads aligned to the original *R. hastatulus* draft assembly. The markers could be split into five linkage groups using a LOD score limit of 30 for the initial split followed by a LOD score limit of 34 for the resulting largest linkage group (SeparateChromosomes2). Most of the remaining single markers were put into these groups using a LOD score limit of 25 (JoinSingles2All), totaling about 120,000 markers. We then calculated the marker order for each linkage group with OrderMarkers2 (with default settings).

To improve our linkage map, we reduced redundancy in our genome assembly and constructed a new pseudo-chromosome assembly using the Lep-Anchor [28] software. To obtain reliable linkage maps, we removed the 13 most-recombining individuals from the maps and constructed three independent linkage maps (Lep-MAP3: OrderMarkers2), using only male informative markers (parameter informativeMask=1), only female informative markers (informativeMask=2) and all markers. These maps were used by the Lep-Anchor software.

To reduce redundancy in the genome assembly, we first split the existing Dovetail assembly into contigs based on assembly gaps. Due to a high number of contigs (>44,000), we removed all contigs of < 500bp, full length haplotypes and joined partial haplotypes in windows of five adjacent contigs (link strength was 6 - |distance| - |difference in orientations|, where distance between contig i and j is |i-j| and difference is 0-2 based on how the orientations differ: same=0, one different=1, both different=2). This was done by Lep-Anchor giving it only the alignment chain computed by Haplomerger2 [29] on the WindowMasker [30] soft-masked (contig-split) genome. This allowed us to reduce the number of contigs to about 33,000. All data were lifted to the new contig assembly coordinates using the liftover script and LiftoverHaplotypes module in Lep-Anchor.

We then ran Lep-Anchor (lepanchor_wrapper2.sh) on the final contig assembly using the three linkage maps, new alignment chains (HaploMerger2) and alignments of raw Pacbio sequence aligned by minimap2 [31]. The resulting pseudo-chromosome assembly was used to calculate physical order of linkage map markers and the maps were evaluated (OrderMarkers2 parameter evaluateOrder) in this order to obtain the final linkage maps. After assembly improvement, we compared contig orders between our previous [18] and new maps using custom R scripts incorporating Plotly [32] interactive plotting.

Recombination rates and transmission distortion

We quantified recombination rates in two ways. First, we described recombination using map lengths in centimorgans (cM) from the maps produced by Lep-Map3. Based on the scale of recombination observed in previous work [18] and the current map, we performed all downstream analyses using 1Mb windows. We also calculated recombination rates as the sum of crossover events per 1Mb window. We first calculated the number of crossovers per site from cM differences using the inverse of the Haldane mapping function [33], then summed crossovers in 1Mb windows. To describe the extent of recombination suppression, we identified the total number of consecutive windows with zero crossovers. We estimated transmission ratio distortion as a likelihood ratio of the deviation from 1:1 transmission of haplotypes from the male and female parent using custom scripts by PR.

Gene and TE content

We also developed a new annotation using MAKER version 3.01.03 [34]. For our MAKER annotation, we used the soft-masked [30] genome integrated with previously published floral transcriptomes from six individuals [21] and leaf transcriptomes from six populations [17]. Transcripts were assembled with IDBA-tran version 1.2.0 [35] and annotated in four rounds, using the transcripts and the Tartary buckwheat annotation version FtChromosomeV2.IGDBv2 [36] as the evidences for MAKER. We functionally annotated the final annotation based on homology using BLAST version 2.2.28+ [37] and InterProScan 5.52-86.0 [38]. This annotation resulted in 59,121 genes. We also annotated the locations of rDNA repeats using rnammer-1.2 [39]. The parameters used, '-S euk' and '-m tsu,ssu,lsu', indicate that the input reference is a eukaryote, and that we are annotating 5/8s, 16/18s, and 23/28s rDNA.

We produced the TE annotation using the EDTA (Extensive de-novo TE Annotator) version 1.9.7 pipeline [40]. This pipeline combines the best-performing structure- and homology-based TE finding programs (LTR_FINDER_parallel [41], LTR_HARVEST_parallel [42], LTR_retriever [43], TIR-Learner2.5 [44], HelitronScannerv1.1[45], Repeatmodeler-2.0.1 [46] and RepeatMasker-4.1.1 [47] and filters their results to produce a comprehensive and non-redundant TE library [40]. The optional parameters '--sensitive 1' and '--anno 1' were used to identify remaining unidentified TEs with RepeatModeler and to produce an annotation. The 'EDTA.TEanno.split.gff3' output file was used as our non-overlapping TE annotation. This file is produced by EDTA by removing overlaps according to the following priorities: structure-based annotation > homology-based annotation, longer TE > shorter TE > nested inner TE > nested outer TE [40].

For all gene content analyses, we used a stringently filtered set of genes to remove gene annotations associated with transposable elements. We first used BEDtools [48] to remove any exons that overlapped a TE, although genes containing both exons that overlapped TEs and exons that did not overlap TEs were retained. We then removed any gene functionally annotated with 'transpos*' (transposon, transposase, etc.), 'ribonuclease H,' 'pol poly,' 'mitochondri*,' 'chloroplast,' or 'retrovirus.' This filtered annotation contained 30,641 genes.

Differential expression and SNP calling

We performed differential expression analyses using DESeq2 (v. 1.28.1) [49] and our new annotation. For DESeq2 analyses, we aligned reads to the new genome pseudomolecules using STAR version 2.7.6a [49] and generated readcounts using featureCounts (2.02) [50]. Cutoffs for differential expression were as follows: adjusted p-value <0.1, absolute Log2Fold change >1. We identified genes that were differentially expressed between male and female leaf tissues using published leaf RNA sequence data from population samples of the XY cytotype [17], and between male and female floral tissue using published RNA sequence data from the XY cytotype [21]. Genes with fewer than 20 reads across all samples were removed from these analyses. We also identified sequences that were differentially expressed in pollen tissue compared to male leaf tissue, using published sequence data [21]. Finally, we identified sequences differentially expressed in pollen tubes compared to pollen, using pollen from the individuals in the mapping population described [18]. We collected pollen using a kief box (Wacky Willy’s, Victoria, BC, Canada), germinated and grew it in 100µL of media [51] for 24 hours, and flash froze it in LN2. After removing media, we lysed cells and extracted total RNA using Spectrum Plant Total RNA Kits (Sigma Aldrich) for RNA extraction. To identify sex-linked SNPs and fixed differences between the X and Y chromosome (all females homozygous reference or non-reference, all males heterozygous) for our new assembly, we used FreeBayes v0.9.10-3-g47a713e [52] to call SNPs from the population transcriptome data from the XY cytotype (six males and six females) [17] and the crossing transcriptome data from the same cytotype (six male and six female offspring, plus parents) [17]. We filtered the SNPs to exclude any with a SNP quality score of lower than 60, any sites with missing data, and fixed heterozygous SNPs across all samples that likely reflected paralogous mapping.

Linear modelling predictors of recombination rate

We combined our linkage map data with our annotation, TE annotation, differential expression data and summed and averaged variables in 1-Mb windows to perform analyses of recombination landscape, gene content, and differential expression in R version 4.1.0 [53] in RStudio version 1.4.1717 [54] using the packages dplyr version 1.0.7 [55] and stringr version 1.4 [56]. We performed correlations using R's built-in cor function, and estimated partial correlations using the package ppcor version 1.1 [57].

To identify factors associated with genome structure that predicted recombination rates and recombination rate differences [58], we created linear models with the following response variables: female crossovers per window, male crossovers per window, sex-averaged crossovers per window, crossover number sex difference per window, and female vs. male biased recombination across window. We fit all responses using generalized linear models with either negative binomial or Tweedie distributions except for female versus male biased recombination, for which we used logistic regression. We performed linear models in glmmtmb version 1.12 [59], evaluated fit using DHARMa version 0.4.3 [60], and compared models using ANOVAs. We performed separate models for each response variable on each chromosome. Scripts are available at https://github.com/joannarifkin/Rumex-sex-specific.

Results

Linkage mapping and genome assembly improvement

We identified five linkage groups, consistent with both the karyotype of the XY cytotype of this species [16, 61] and our previous sex-averaged linkage mapping results [18] (table 1). We again identified two apparently metacentric linkage groups (A1 and A2) and three apparently submetacentric linkage groups (A3, A4, XY) based on the patterns of recombination across the chromosomes (figure 1A) and the identities of the scaffolds constituting the linkage groups. We have retained the same autosomal labels across both maps, and they continue to reflect chromosome sizes from largest (A1) to smallest (A4).

Our larger genetic mapping population and improved genome assembly led to considerable improvement in higher-order chromosome-scale scaffolding of the genome of *R. hastatulus*. Our improved genome assembly contained 1.45Gb, a reduction of 0.2Gb from our previous 1.65Gb assembly [18] due to the collapsing of redundant haplotypes (see Methods). For this assembly, 1.212Gb (84%) is now grouped in the five linkage groups (previously 1.08GB, 65% of the previous primary assembly), with the remaining 0.23Gb in smaller contigs. These corrections have substantially increased the size of the assembled sex chromosome, with an additional 88.6 MB of sequence assembled on the sex chromosome, the largest increase of any of the chromosomal scaffolds (table 1). Consistent with this increase, analysis of our past genome assembly showed that only 52% of sex-linked SNPs mapped to the assembled sex chromosome; our new assembly integrated with transcriptomes from independent crossing data [17] now shows that 94% of SNPs showing X-Y segregation patterns map to the sex chromosome.

Recombination rates

As in our previous study, we found that recombination was unevenly distributed across the genome, with very large non-recombining regions on all chromosomes (figure 1). We identified clear evidence of heterochiasmy (table 1, figure 1A). Male map lengths were shorter than female map lengths: across chromosomes, female map length was 1.4x male map length (table 1, figure 1A), and the sex chromosome was not an obvious outlier for this metric. Males also had longer blocks of non-recombining windows across all chromosomes. To summarize this pattern, we identified the longest stretches of markers on each chromosome with zero crossovers. On the autosomes, males had runs of non-recombining windows approximately twice as long as those of females, with male-specific non-recombining regions as large as 238MB (table 1). By this measure, the sex chromosomes were an exception: the largest run of male-specific non-recombining windows on the sex chromosome was four times the length of the longest run of female-specific non-recombining sequence (table 1). Thus, although all chromosomes exhibited reduced male recombination rates, the XY chromosome showed the largest region of differentially suppressed recombination between the sexes despite not being the largest chromosome. Overall, male and female recombination rates in *R. hastatulus* were only weakly correlated (r=0.333, correlation of male and female crossover number across 1Mb windows across all chromosomes; figure 2, table S1).

The extent of sex differences in recombination varied both along and between chromosomes (figure 1B). Chromosomes A1, A2, and the sex chromosome conformed to the common pattern of more tip-focused recombination in males, but the submetacentric chromosome A3 showed female-biased recombination in the more highly recombining end, and A4 appeared to have low-recombination regions at both ends. This differs from the previous linkage map, likely because of the difficulty of positioning low-recombination regions. In general, pericentromeric regions showed reduced recombination in both sexes, but female map lengths were larger in these regions and showed apparent hotspots of recombination with large jumps in centimorgan position (figure 1). All five chromosomes had regions of female-biased and male-biased recombination, as well as shared recombining and non-recombining regions (figure 1B). This complex pattern creates a highly heterogeneous recombination landscape.

Table 1: The linkage groups of the revised *Rumex hastatulus* genome assembly, including sex-averaged, male, and female map lengths, length in Mb, gene content, and extent of non-recombining region

LG	Sex-averaged map length(cM)	Male map length (cM)	Female map length (cM)	Mb*	Number of genes	Largest size of markers with 0 crossover events (Mb), in males/females
A1	104.57	94.624	114.516	388.3386 (344.5)	7512	238.6 / 94.3
A2	91.67	79.032	104.301	278.2766 (260.43)	7301	33.4 / 22.9
A3	69.09	48.387	89.785	171.7219 (175.01)	3923	91.5 / 44.3
A4	61.29	52.688	69.892	135.2128 (158.24)	3502	62.4 / 30.2
XY	79.03	64.516	93.548	239.0056 (150.39)	4752	212.1 / 55.3

*Values in brackets indicate the lengths from the previous assembly

Gene and TE content

As in our previous study, we found that genes were generally concentrated in high-recombination regions (figure 1D). However, A1 contained one gene-dense yet low-recombination region (~100Mb-200Mb). RNA (class 1) TEs were concentrated in low-recombination regions, whereas DNA (class 2) TEs were concentrated in high-recombination regions (figure 1E). Despite the reduced gene density in low-recombination regions, the extent of high recombination-suppressed regions means that a large fraction of genes in the genome (approximately 37%) are in these large regions with no male recombination. Ribosomal genes were concentrated mostly on A3 (62 rDNA features annotated) and A4 (28 rDNA features annotated). rDNA genes occurred in the first 50Mb of A3 (5S subunit sequence) with additional rDNA sequence located around 130Mb of A3 (18S and 28S subunit sequence) and in the first 3 Mb of A4 (18S and 28S subunit sequence). These rDNA locations are consistent with past cytological findings [61], further confirming our identification of the A3 and A4 chromosomes.

Characterization of the sex-determining (SDR) and pseudo-autosomal regions

With female and male recombination separated, it is clear that the male-specific non-recombining region on the sex chromosome in *R. hastatulus*, i.e. the SDR, is extensive. In particular, our linkage mapping suggests that the SDR is as large as 209MB (14% of the total assembly: table 1, figure 1), including as many as 3595 genes (12% of the total filtered annotated genes). The gene-dense recombining pseudoautosomal region of the sex chromosome is similarly the smallest male recombining segment of any chromosome, representing only approximately 13% of the physical size of the chromosome. Note that the ‘true’ SDR may be narrower and the pseudoautosomal region larger, since rare male recombination may have gone unobserved in our cross. However, BLAST [37] searches of our sex-linked transcripts that have at least one fixed difference between the X and Y chromosomes from a population sample [17] are found across most of the length of this nonrecombining region (from 1.6 MB to 208.5 MB), and fixed differences mapped onto the chromosome are common across the first 210 MB (figure S1) suggesting that most of this region is nonrecombining and linked to the SDR.

Transmission ratio distortion

We identified loci with biased haplotype transmission through either paternal or maternal inheritance based on the haplotype reconstructions from Lep-Map 3. Transmission ratio distortion varied between chromosomes (figure 1C, figure S2, table S2). More sites experienced biased transmission through maternal (276) than paternal (48) inheritance, across a larger fraction of the genome. On A1, 18 sites were significantly distorted in transmission from males using a 0.05 cutoff in a chi-squared distribution (LOD > 3.841) and on A4, 30 sites were significantly distorted in transmission from males with a 0.05 cutoff. Although deviations from 1:1 male haplotype transmission occurred on other chromosomes, there were no significant male-distorted sites on A2, A3, or the sex chromosome. In contrast, female haplotype distortion was extensive on A2 (91 sites in females at 0.05, 38 at 0.01 cutoff of LOD > 6.635), A3 (108 sites at .05 cutoff, 88 at 0.01), and the sex chromosome (76 sites at a 0.05 cutoff, 26 at 0.01) but negligible on A1 (0 sites) and A4 (1 site at 0.01).

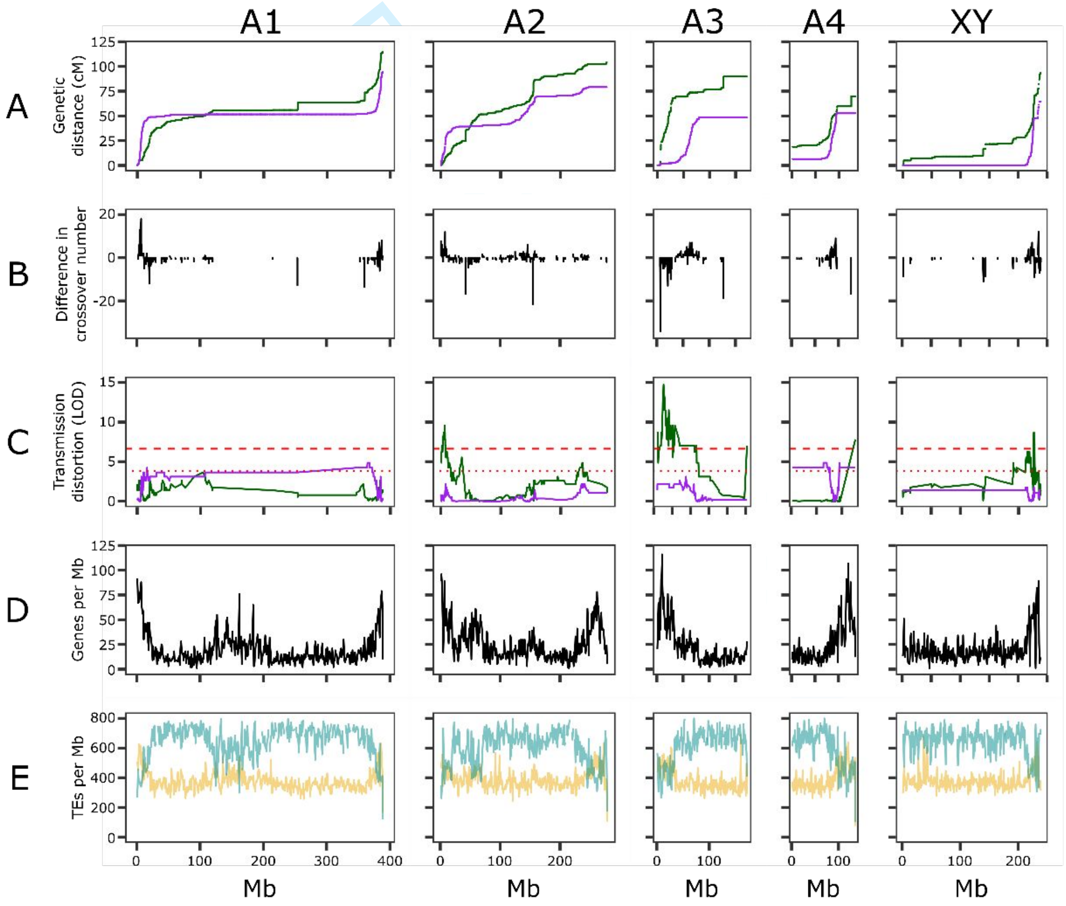


Figure 1. Distribution of recombination, segregation distortion, and gene content in *Rumex hastatulus*. A. Male (purple) and female (green) Marey maps of the chromosomes. B. Difference in crossover number for 1Mb windows along the chromosome (male crossovers per window - female crossovers per window). Positive: male crossover excess. Negative: female crossover excess. C. Segregation distortion for male (purple) and female (green) haplotypes. Dashed red lines indicate significance at 0.05 and 0.01 levels according to a chi-squared test. D. Genes per 1Mb window along the genome. E. TEs per 1Mb window along the genome. Yellow: DNA TEs. Blue: RNA TEs.

Correlates of recombination rate differences

Across the genome, the number of genes, leaf- and flower-expressed genes, and DNA (Class 2) TE density were all positively correlated with recombination rate, and RNA (Class 1) TE density was negatively correlated with recombination rate (figure 2, table S1, figure S3-S6). Male crossover number and female crossover number were both positively correlated with gene density, but this correlation was stronger for male crossovers (male Pearson's *r*=0.390, female Pearson's *r*=0.278). However, the correlations between transmission ratio distortion and crossover number were in opposite directions: in females, more distorted regions were also more recombining (*r*=0.111) whereas in males distortion was negatively correlated with crossover number (*r*=-0.197), reflecting the fact that male transmission distortion signals were enriched in the large non-recombining regions of male meiosis (figure 1). Correlations varied in strength between chromosomes, with notable differences in correlates of transmission ratio distortion and recombination rate difference, both of which varied in magnitude, position, and direction along and between chromosomes (figure 2, figure S7, table S3). The difference in crossover number between the sexes was most strongly correlated with signals of male distortion (figure 2), where regions of particularly low male crossover number represented regions with larger signals of male distortion. We also estimated partial correlation coefficients controlling for gene density, which was consistently correlated with many genomic variables (figure S8, figure S9, tables S4-S5). Correlations between distortion and crossover number in both males and females persisted after controlling for gene density, and varied in direction across chromosomes (figure S9). However, all chromosomes except A2 showed a consistent negative correlation between male-biased recombination and male transmission distortion, even when gene density was controlled.

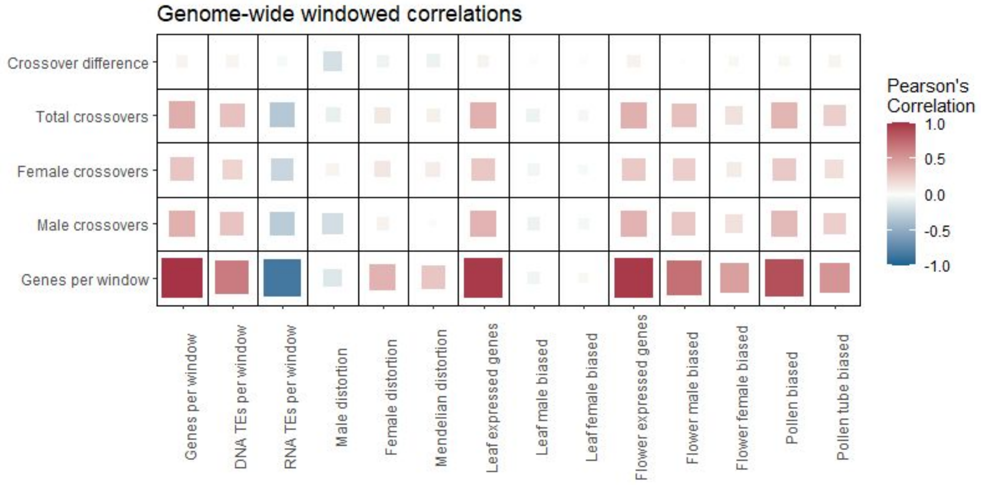


Figure 2: Correlations among genome window characteristics across the whole genome of *Rumex hastatulus*. Colours and sizes correspond to the strength and direction of the correlations.

Linear models of recombination rate differences

We used generalized linear models to identify predictors of sex-specific and sex-averaged recombination rates, whether recombination was male- or female-biased, and the magnitude of the recombination rate difference in *R. hastatulus*. Full modeling results are available in table S6 and table S7.

In our linear models, the variables that significantly predicted male recombination rate varied between chromosomes (table 2A). Number of genes predicted increased male recombination rate on three chromosomes (as well as a fourth in some models), and position along the genome emerged as an important predictor on two submetacentric chromosomes, suggesting that distance from centromere predicts male recombination (figure 1). RNA TE count also appeared to play a role, but in inconsistent directions, predicting increased male recombination on two chromosomes and decreased male recombination on a third; this pattern also appears in the partial correlations with gene density removed (figure S9). Transmission ratio distortion predicted reduced male recombination rate on three chromosomes (A2, A3, and A4), and predicted increased male recombination on A1. Either pollen bias or pollen tube bias predicted male recombination rate on three chromosomes (negatively on A4 and the sex chromosome, positively on A3).

In contrast, the predictors of female recombination are more consistent across chromosomes (table 2B). Number of genes per window positively predicted female recombination rate on all five chromosomes, and position relative to centromere affected female recombination rate on four chromosomes. Only one other variable, pollen tube-biased expression, predicted female recombination rate on more than one other chromosome. Female transmission ratio distortion did not emerge as a significant predictor of recombination rate on any chromosome.

Predictors of both sex-averaged recombination and of the magnitude of the recombination rate difference between males and females reflected predictors of male and female recombination rate independently. Number of genes, position along the chromosome, number of RNA TEs, female-biased floral expression, pollen tube-biased expression all appeared as significant predictors, but their importance and direction varied between chromosomes (table S6A, S6B).

Finally, we used logistic regression to identify variables that predicted whether windows exhibited female or male recombination bias (table S6C). Our logistic regressions also suggested considerable variation in the factors that predicted sex differences in recombination. Number of genes per window was an important predictor for four out of five chromosomes, but in variable directions, predicting both male bias (A1, XY) and female bias (A2, A3) in recombination.

Table 2
Variables identified as significant predictors of recombination in linear models of windows of the *Rumex hastatulus* genome.

Male recombination rate predictors
'+' - variable predicts increased male recombination. '-' - variable predicts decreased male recombination. * - 95% confidence interval for slope estimate includes zero, chiefly because term is significant in interaction term.

	# genes	Position	# RNA TEs	Male distortion	Pollen tube bias	Pollen bias
A1	+		+	+		
A2	+		+*	-*		
A3 (neo)	~ +*	-*		-*	+	
A4			-	-	-	
XY	+*	+				-

Interaction effects in best model:
A1: n RNA TEs:male distortion (-)
A2: n genes:male distortion (+ ns)
A3: male distortion:position window (+)
A3 (more complex model): position window:n genes (- ns), male distortion:position window (+), n genes:male distortion, n genes:position window:male distortion (+ns)
A4: none
XY: position window:n genes (-)

Female recombination rate predictors
'+' - variable predicts increased female recombination. '-' - variable predicts decreased female recombination. * - 95% confidence interval for slope estimate includes zero, chiefly because term is significant in interaction term.

	# genes	Position	# RNA TEs	Flower female bias	Pollen tube bias	Leaf expression
A1	+*		-*		-	
A2	+*	-*				
A3 (neo)	+	-*				-
A4	+	+				
XY	+*	+		+	+*	

Interaction effects in best model:
A1: n RNA TEs:pollen tube bias (+)
A2: n genes: position window (-ns)
A3: n genes: position window (-), leaf expression: position window (+)
A4: n genes: position window (- ns)
XY: position window:female flower bias (-)
XY (more complex model): position window:female flower bias (-), position window:pollen tube bias (-ns), position window:pollen tube bias:flower female bias (+)

Discussion

The main findings of this study are consistent with the general pattern of extensive recombination suppression that we previously inferred based on sex-averaged recombination in *R. hastatulus* [18]. However, our sex-specific maps show that recombination suppression is not evenly distributed between males and females, and that the observed very large pericentromeric regions of suppressed recombination are particularly influenced by highly suppressed male recombination. Across all chromosomes, females recombine more frequently than males and males have much larger non-recombining blocks than females. Our results are in line with several studies of hermaphroditic plants, as well as other eukaryotes [1], with the very large male-specific non-recombining regions that we report making this an extreme case.

However, *R. hastatulus* does not strictly follow the common eukaryotic pattern of tip-focused male recombination [1] across all chromosomes. The recombination landscapes of both the metacentric and submetacentric chromosomes suggest greater variation in the distribution of recombination than simply less male recombination in the centres and more at the chromosome ends, with highly recombining regions scattered along chromosomes (particularly A2, A3, and A4). Thus, male recombination is more concentrated, but not always at the tips of the chromosome. These differences among chromosomes may reflect an ongoing history of chromosomal rearrangements in the genus and additional patterns of chromosome structure such as the locations of centromeres and rDNA clusters [61].

The pattern of larger non-recombining regions in males is consistent with an evolutionary bias toward the evolution of male heterogamety and XY sex chromosomes in *Rumex*. In particular, suppressed recombination can facilitate the maintenance and invasion of sexually antagonistic variants linked to sex-determining regions [62], and sex-determining regions that evolve in large pericentromeric non-recombining regions, especially sex-specific ones, may contribute to the evolution of sex chromosomes [63]. The existence of male-specific non-recombining regions may thus facilitate the evolution of XY rather than ZW sex chromosomes [1]. Given our observation of no crossovers in male meiosis over very large fractions of each chromosome, it is possible that recombination suppression was ancestral to the evolution of dioecy in the genus, and that subsequent recombination modifiers did not evolve following the origin of the SDR. The observed size of the SDR, which is over 200 MB using population-validated sex-linked genes and includes over 14% of the assembled genome and over 3500 genes, is much larger than those recently reviewed to date in plants [64], although larger population samples of *R. hastatulus* should be used to test for very rare recombination between the X and Y, and the sex-linked region of *S. latifolia* may be even larger [65]. Although we cannot rule out a role for subsequent recombination modifiers, particularly because the sex chromosomes show the most extreme size dimorphism of the nonrecombining region (table 1), our results do suggest that sex differences in heterochiasmy may have played an important role in determining the large size of the SDR facilitating the evolution of large heteromorphic sex chromosomes in this system. Comparative studies of heterochiasmy in both hermaphroditic and other dioecious species in *Rumex* will be important to further assess the extent to which ancestral heterochiasmy and derived changes in recombination rates have contributed to sex chromosome evolution in this lineage.

Models for the evolution of heterochiasmy due to male haploid selection and female meiotic drive both predict lower overall male recombination rates and higher female recombination near centromeres [1,6,7], as we observed in *R. hastatulus*. Our mapping population provided evidence for both male and female transmission ratio distortion, which varied within and among chromosomes and between the sexes. Overall, more sites displayed significant distortion in female than in male transmission, but significant regions of both types of distortion were observed across the genome. Transmission ratio distortion through female inheritance is generally thought to be consistent with female meiotic drive. In contrast, transmission ratio distortion through male inheritance may result from haploid (pollen) competition [66].

Nevertheless, zygotic distortion (i.e., differential seed germination or seedling survival) could also lead to transmission ratio distortion, and may result from alleles inherited from either parent [66]. In our study, we genotyped reproductive adults rather than pollen or seeds, which conflates several opportunities for natural and sexual selection causing biased transmission. Zygotic selection may be particularly likely to explain our observed female distortion, since these regions were not focused on low-recombination centromeric regions, where meiotic drive is expected to act [7]. In contrast, signals of male transmission distortion are particularly enriched in regions of low male recombination and high sex bias in recombination (figure 1 and figure 2, table 2), suggesting that haploid selection in males may be an important selective pressure for sex differences in recombination. Distorted regions can vary widely between populations of the same species [67], so distortion in a single cross should be interpreted with some caution. Furthermore, patterns of biased pollen or pollen tube expression do not show similarly consistent enrichment in regions of low male recombination except for the sex chromosome (table 2, figure S6), although direct observation of transmission distortion likely provides a stronger indicator of loci involved in pollen competition. Similarly, although we did not identify evidence consistent with ongoing recombination increases to counter meiotic drive, the overall pattern of increased and more centromere-biased recombination is still consistent with a history of selection eliminating meiotic drive alleles. With those caveats, our results do provide some evidence in accord with the hypothesis that pollen competition may play an important role in sex differences in recombination.

We used both pairwise correlations and regression models to investigate various genomic correlates of male and female recombination and sex bias in recombination, in order to further explore other possible factors favouring sex differences in recombination. On a genome-wide scale, both male and female recombination rates in *R. hastatulus* are consistent with widely observed patterns that genes and Class 2 DNA TEs concentrate in high-recombination regions and Class 1 RNA TEs concentrate in low-recombination regions (figure 2; [3], [56]). These patterns are consistent with recombination rates in plants occurring preferentially upstream of genes, and with epigenetic silencing of retrotransposable elements causing a reduction of recombination, although they could also be explained if transposable elements preferentially accumulate in regions of low sex-averaged recombination [3].

Because the recombination landscape varied widely between the chromosomes of *R. hastatulus*, we also identified chromosome-specific predictors of recombination rates and recombination rate bias using linear models and partial correlations controlling for gene density. At this more granular scale, a different and more complex picture emerges (table 2, figure S8, S9). In particular, our logistic regression models of recombination bias direction support the possibility that different mechanisms may be contributing to variation in recombination rates on different chromosomes. Both number of genes per window and position along the genome predicted female-biased recombination on some chromosomes and male-biased recombination on others. Aside from physical position and gene density, different diverse factors predicted both male and female bias on different chromosomes, including RNA TEs, number of genes expressed in different tissues, sex-biased floral expression, pollen-biased expression, and transmission ratio distortion (table 2C). This finding suggests a general picture in which gene density and proximity to centromere shape recombination on a 'global scale', but variation in gene content and haploid selection may lead to region-specific selective forces acting on both male and female recombination rates. Consistent with this, a recent comparative study in fish [68] found that sex differences in recombination are labile at the species level, but do not present clear trends consistent with adaptive hypotheses across species.

Conclusions

Our study has provided some of the first evidence of sex differences in recombination and identified one of the largest known SDRs in a dioecious plant species, or in fact in any eukaryote [65]. We identified both genome-wide and chromosome-specific factors predicting sex differences in recombination, and also found evidence consistent with a role for male gametophytic selection in driving these differences. Future work in this system will allow more precise dissection of the genetic and evolutionary mechanisms favouring sex differences in recombination landscapes. In particular, exploring both sex-specific eQTL positions and further study of transmission ratios in pollen and seeds will allow us to further differentiate between sexually antagonistic cis epistasis in diploids and epistasis in haploids [1]. Finally, characterizing sex-specific recombination landscapes of hermaphroditic *Rumex* species should make it possible to determine whether these sex differences in recombination landscape did indeed precede and promote the evolution of a heterogametic XY sex determining system with a very large SDR, as we have hypothesized.

Acknowledgments

We thank University of Toronto undergraduate students Victoria Marshall, Claire Ellis, Deanna Kim, and Madeline Jarvis-Cross for technical assistance, Bill Cole and Tom Gludovacz for glasshouse support, and Brechann McGoey for crossing chamber development. This research was supported by Discovery grants from the Natural Sciences and Engineering Research Council of Canada to SCHB and SIW. JLR was supported by an EEB post-doctoral fellowship.

Data accessibility statement

Raw sequence has been deposited on the Sequence Read Archive (SRA) under the accession number PRJNA692236 (embargoed until July 1, 2022 or publication). Our new genome assembly, transcriptome annotation, rDNA annotation, and TE annotation have been deposited in the CoGe comparative genomics platform at https://genomevolution.org/coge/GenomeInfo.pl?gid=62326. Scripts used in the analyses have been deposited on Github at https://github.com/joannarifkin/Rumex-sex-specific.

References

1. Sardell JM, Kirkpatrick M. 2020 Sex differences in the recombination landscape. *Am. Nat.* **195**, 361–379.
2. Campos JL, Halligan DL, Haddrill PR, Charlesworth B. 2014 The relation between recombination rate and patterns of molecular evolution and variation in *Drosophila melanogaster*. *Mol. Biol. Evol.* **31**, 1010–1028.
3. Kent TV, Uzunović J, Wright SI. 2017 Coevolution between transposable elements and recombination. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **372**, 20160458.
4. Nachman MW, Payseur BA. 2012 Recombination rate variation and speciation: theoretical predictions and empirical results from rabbits and mice. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **367**, 409–421.
5. Stapley J, Feulner PGD, Johnston SE, Santure AW, Smadja CM. 2017 Variation in recombination frequency and distribution across eukaryotes: patterns and processes. *Philos. Trans. R. Soc.*

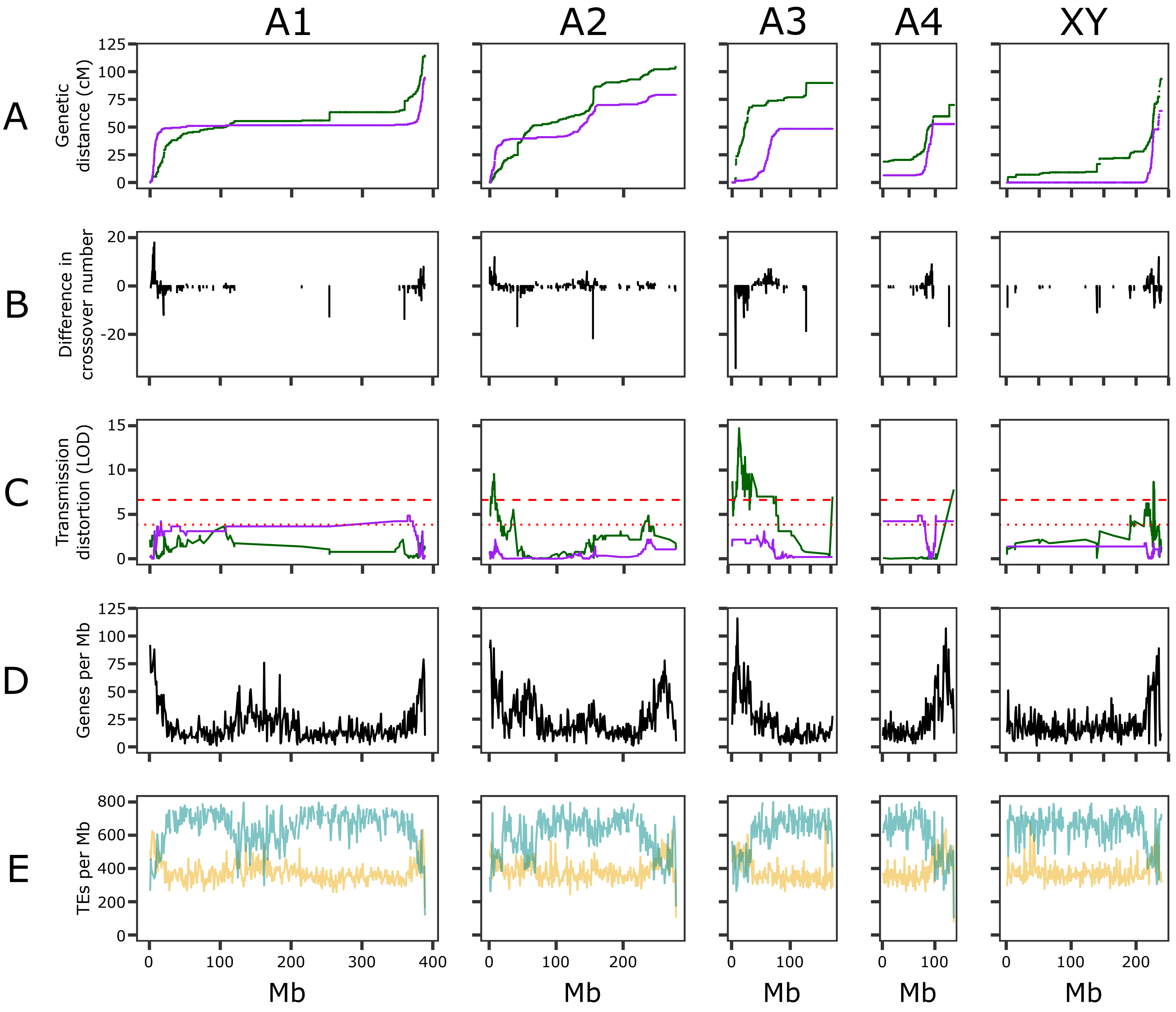
Lond. B Biol. Sci. **372**, 20160455.

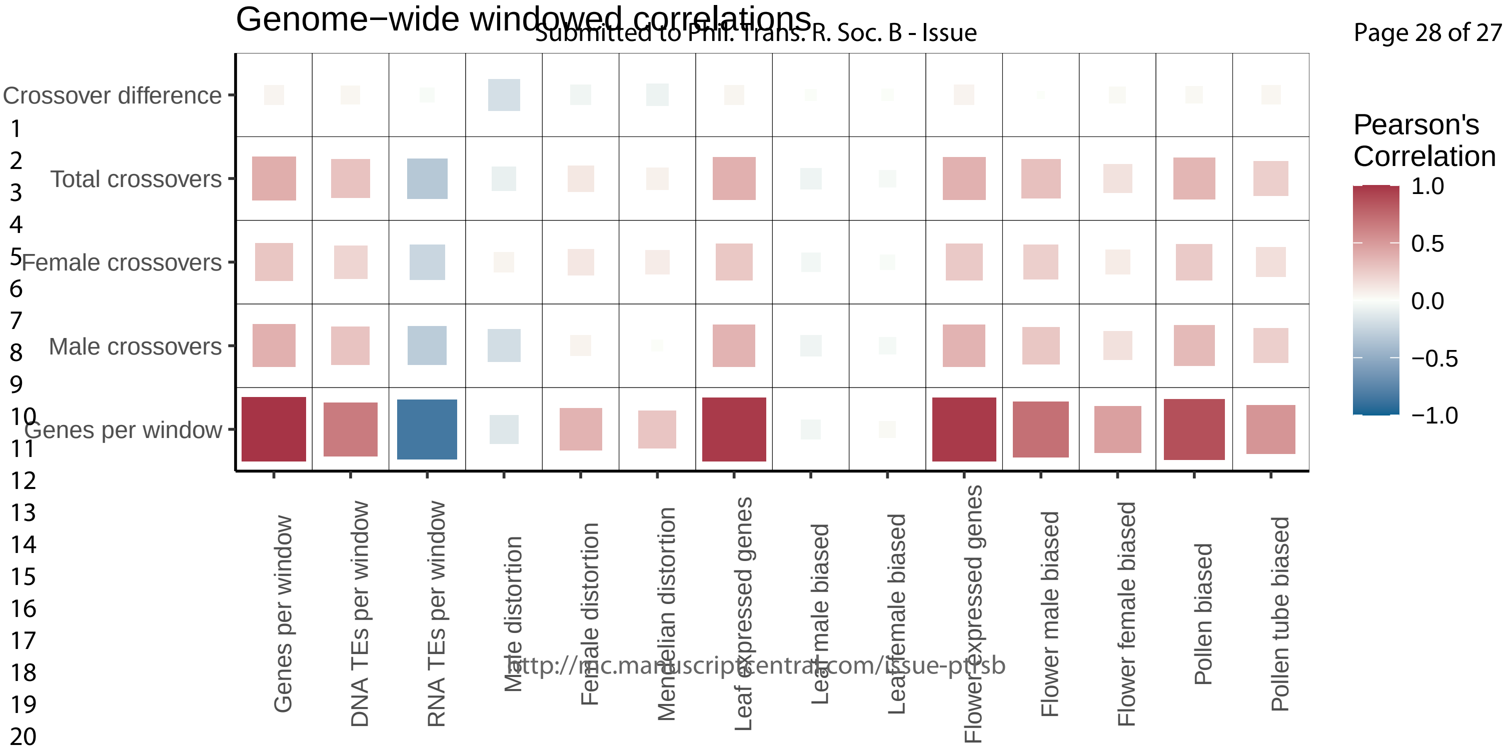
6. Lenormand T, Dutheil J. 2005 Recombination difference between sexes: a role for haploid selection. *PLoS Biol.* **3**, e63.
7. Brandvain Y, Coop G. 2012 Scrambling eggs: meiotic drive and the evolution of female recombination rates. *Genetics* **190**, 709–723.
8. Johnston SE, Huisman J, Ellis PA, Pemberton JM. 2017 A high-density linkage map reveals sexual dimorphism in recombination landscapes in Red Deer (*Cervus elaphus*). *G3* **7**, 2859–2870.
9. Bergero R, Gardner J, Bader B, Yong L, Charlesworth D. 2019 Exaggerated heterochiasmy in a fish with sex-linked male coloration polymorphisms. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 6924–6931.
10. Giraut L, Falque M, Drouaud J, Pereira L, Martin OC, Mézard C. 2011 Genome-wide crossover distribution in *Arabidopsis thaliana* meiosis reveals sex-specific patterns along chromosomes. *PLoS Genet.* **7**, e1002354.
11. Lagercrantz U, Lydiate DJ. 1995 RFLP mapping in *Brassica nigra* indicates differing recombination rates in male and female meioses. *Genome*. **38**, 255–264.
12. Phillips D *et al.* 2015 The effect of temperature on the male and female recombination landscape of barley. *New Phytol.* **208**, 421–429.
13. Kianian PMA *et al.* 2018 High-resolution crossover mapping reveals similarities and differences of male and female recombination in maize. *Nat. Commun.* **9**, 2370.
14. de Vicente MC, Tanksley SD. 1991 Genome-wide reduction in recombination of backcross progeny derived from male versus female gametes in an interspecific cross of tomato. *Theor. Appl. Genet.* **83**, 173–178.
15. Veltsos P *et al.* 2019 Early sex-chromosome evolution in the diploid dioecious plant *Mercurialis annua*. *Genetics* **212**, 815–835.
16. Smith BW. 1964 The evolving karyotype of *Rumex hastatulus*. *Evolution* **18**, 93–104.
17. 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.
18. Rifkin JL, Beaudry FEG, Humphries Z, Choudhury BI, Barrett SCH, Wright SI. 2021 Widespread recombination suppression facilitates plant sex chromosome evolution. *Mol. Biol. Evol.* **38**, 1018–1030.
19. Field DL, Pickup M, Barrett SCH. 2012 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.
20. Pickup M, Barrett SCH. 2013 The influence of demography and local mating environment on sex ratios in a wind-pollinated dioecious plant. *Ecol. Evol.* **3**, 629–639.
21. Sandler G, Beaudry FEG, Barrett SCH, Wright SI. 2018 The effects of haploid selection on Y chromosome evolution in two closely related dioecious plants. *Evol Lett* **2**, 368–377.
22. Wilby AS, Parker JS. 1988 Mendelian and non-Mendelian inheritance of newly-arisen chromosome rearrangements. *Heredity* **60**, 263–268.
23. McGoey BV, Janik R, Stinchcombe JR. 2017 Individual chambers for controlling crosses in wind-pollinated plants. *Methods Ecol. Evol.* **8**, 887–891.
24. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013 STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21.
25. Dobin A, Gingeras TR. 2016 Optimizing RNA-Seq Mapping with STAR. *Methods Mol. Biol.* **1415**, 245–262.
26. Van der Auwera GA, O’Connor BD. 2020 *Genomics in the Cloud: Using Docker, GATK, and WDL in Terra*. Newton, MA: O’Reilly Media, Inc.

27. Rastas P. 2017 Lep-MAP3: robust linkage mapping even for low-coverage whole genome sequencing data. *Bioinformatics* **33**, 3726–3732.
28. Rastas P. 2020 Lep-Anchor: automated construction of linkage map anchored haploid genomes. *Bioinformatics* **36**, 2359–2364.
29. Huang S *et al.* 2012 HaploMerger: reconstructing allelic relationships for polymorphic diploid genome assemblies. *Genome Res.* **22**, 1581–1588.
30. Morgulis A, Gertz EM, Schäffer AA, Agarwala R. 2006 WindowMasker: window-based masker for sequenced genomes. *Bioinformatics* **22**, 134–141.
31. Li H. 2018 Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100.
32. Plotly Technologies Inc. Collaborative data science. Montréal, QC, 2015. https://plot.ly.
33. Haldane JBS. 1919 The combination of linkage values and the calculation of distances. *J. Genet.* **8**, 299-309.
34. Cantarel BL, Korf I, Robb SMC, Parra G, Ross E, Moore B, Holt C, Sánchez Alvarado A, Yandell M. 2008 MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* **18**, 188–196.
35. Peng Y, Leung HCM, Yiu S-M, Lv M-J, Zhu X-G, Chin FYL. 2013 IDBA-tran: a more robust de novo de Bruijn graph assembler for transcriptomes with uneven expression levels. *Bioinformatics* **29**, i326–34.
36. Zhang L *et al.* 2017 The Tartary Buckwheat genome provides insights into rutin biosynthesis and abiotic stress tolerance. *Mol. Plant* **10**, 1224–1237.
37. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990 Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410.
38. Jones P *et al.* 2014 InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240.
39. Lagesen K, Hallin P, Rødland EA, Staerfeldt H-H, Rognes T, Ussery DW. 2007 RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* **35**, 3100–3108.
40. Ou S *et al.* 2019 Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. *Genome Biol.* **20**, 275.
41. Ou S, Jiang N. 2019 LTR_FINDER_parallel: parallelization of LTR_FINDER enabling rapid identification of long terminal repeat retrotransposons. *Mob. DNA* **10**, 48.
42. Ellinghaus D, Kurtz S, Willhoeft U. 2008 LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. *BMC Bioinformatics* **9**, 18.
43. Ou S, Jiang N. 2018 LTR_retriever: A highly accurate and sensitive program for identification of long terminal repeat retrotransposons. *Plant Physiology*. **176**, 1410–1422.
44. Su W, Gu X, Peterson T. 2019 TIR-Learner, a new ensemble method for TIR transposable element annotation, provides evidence for abundant new transposable elements in the Maize henome. *Mol. Plant* **12**, 447–460.
45. Xiong W, He L, Lai J, Dooner HK, Du C. 2014 HelitronScanner uncovers a large overlooked cache of Helitron transposons in many plant genomes. *Proc. Natl Acad. Sci. USA* **111**, 10263–10268. (doi:10.1073/pnas.1410068111).
46. Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, Smit AF. 2020 RepeatModeler2 for automated genomic discovery of transposable element families. *Proc. Natl Acad. Sci. U. S. A.* **117**, 9451–9457.
47. Smit, AFA, Hubley, R & Green, P. 2013-2015 Repeat-Masker Open-4.0. *http://www.repeatmasker.org*. See https://ci.nii.ac.jp/naid/10029514778/ (accessed on 15 September 2021).
48. Quinlan AR, Hall IM. 2010 BEDTools: a flexible suite of utilities for comparing genomic features.

Bioinformatics **26**, 841–842.

49. Love MI, Huber W, Anders S. 2014 Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550.
50. Liao Y, Smyth GK, Shi W. 2014 featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930.
51. Adhikari KN, Campbell CG. 1998 In vitro germination and viability of buckwheat (*Fagopyrum esculentum* Moench) pollen. *Euphytica* **102**, 87–92.
52. Garrison E, Marth G. 2012 Haplotype-based variant detection from short-read sequencing. *arXiv preprint arXiv:1207.3907.
53. R Development Core Team. 2021 R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. See https://www.R-project.org/. (accessed on 2021).
54. R Team. 2020 RStudio: Integrated Development for R. RStudio, PBC, Boston, MA, 2020.
55. Hadley Wickham RF, Henry L, Müller K. 2017 dplyr: A Grammar of Data Manipulation. R package version 0.7. 4.
56. Wickham H. 2019 stringr: Simple, consistent wrappers for common string operations. R package version 1.4.0. See https://CRAN.R-project.org/package=stringr (accessed on 2019).
57. Kim, S. (2015) ppcor: An R Package for a fast calculation to semi-partial correlation coefficients. *Commun. Stat. Appl. Methods*, **22(6)**, 665-674.
58. Paape T, Zhou P, Branca A, Briskine R, Young N, Tiffin P. 2012 Fine-scale population recombination rates, hotspots, and correlates of recombination in the *Medicago truncatula* genome. *Genome Biol. Evol.* **4**, 726–737.
59. Brooks ME, Kristensen K, Van Benthem KJ, Magnusson A, Berg CW, Nielsen A, Skaug HJ, Machler M, Bolker BM. 2017 glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *R J.* **9**, 378–400.
60. Hartig F. 2019 DHARMa: residual diagnostics for hierarchical (multi-level/mixed) regression models. R package version 0.1. 0.
61. Kasjaniuk M, Grabowska-Joachimiak A, Joachimiak AJ. 2019 Testing the translocation hypothesis and Haldane's rule in *Rumex hastatulus*. *Protoplasma* **256**, 237–247.
62. Otto SP. 2019 Evolutionary potential for genomic islands of sexual divergence on recombining sex chromosomes. *New Phytol.* **224**, 1241–1251.
63. Charlesworth, D. 2019 Young sex chromosomes in plants and animals. *New Phyt.* **224**, 1095-1107.
64. Renner SS, Müller NA. 2021 Plant sex chromosomes defy evolutionary models of expanding recombination suppression and genetic degeneration. *Nat Plants* **7**, 392–402.
65. Gschwend, AR, Weingartner, LA, Moore, RC, and R Ming. 2012 The sex-specific region of sex chromosomes in animals and plants. *Chromosome Res*. **20**, 57–69.
66. Fishman L, McIntosh M. 2019 Standard Deviations: The biological bases of transmission ratio distortion. *Annu. Rev. Genet.* **53**, 347–372.
67. Seymour DK, Chae E, Arioz BI, Koenig D, Weigel D. 2019 Transmission ratio distortion is frequent in *Arabidopsis thaliana* controlled crosses. *Heredity* **122**, 294–304.
68. Cooney CR, Mank JE, Wright AE. 2021 Constraint and divergence in the evolution of male and female recombination rates in fishes. *Evolution* https://doi.org/10.1111/evo.14357


A1
A2
A3
A4
XY
A
Genetic distance (cM)
125
100
75
50
25
0
B
Difference in crossover number
20
0
-20
C
Transmission distortion (LOD)
15
10
5
0
D
Genes per Mb
125
100
75
50
25
0
E
TEs per Mb
800
600
400
200
0
0
100
200
300
400
Mb


Genome−wide windowed correlations
Crossover difference
Total crossovers
Female crossovers
Male crossovers
Genes per window
Genes per window
DNA TEs per window
RNA TEs per window
Male distortion
Female distortion
Mendelian distortion
Leaf expressed genes
Leaf male biased
Leaf female biased
Flower expressed genes
Flower male biased
Flower female biased
Pollen biased
Pollen tube biased
Pearson's Correlation
1.0
0.5
0.0
−0.5
−1.0