Submitted to Phil. Trans. R. Soc. B - Issue

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

SCHOLARONE™
Manuscripts

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

Submitted to Phil. Trans. R. Soc. B - Issue

Page 2 of 27

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

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

1

1 Summary

2   There is growing evidence across diverse taxa for sex differences in the genomic landscape of

3   recombination, but the causes and consequences of these differences remain poorly understood. Strong

4   recombination landscape dimorphism between the sexes could have important implications for the

5   dynamics of sex chromosome evolution and turnover because low recombination in the heterogametic

6   sex can help favour the spread of sexually antagonistic alleles. Here, we present a sex-specific linkage

7   map and revised genome assembly of *Rumex hastatulus*, representing the first characterization of sex

8   differences in recombination landscape in a dioecious plant. We provide evidence for strong sex

9   differences in recombination, with pericentromeric regions of highly suppressed recombination in males

10 that cover over half of the genome. These differences are found on autosomes as well as sex

11 chromosomes, suggesting that pre-existing differences in recombination may have contributed to sex

12 chromosome formation and divergence. Analysis of segregation distortion suggests that haploid

13 selection due to pollen competition occurs disproportionately in regions with low male recombination.

14 Our results are consistent with the hypothesis that sex differences in the recombination landscape

15 contributed to the formation of a large heteromorphic pair of sex chromosomes, and that pollen

16 competition is an important determinant of recombination dimorphism.

17

18 Introduction

19 The distribution of rates of recombination along chromosomes (recombination landscape [1]) shapes

20 many aspects of evolutionary genetics, including the efficacy of natural selection [2], genome structure

21 [3], and the dynamics of reproductive isolation [4]. Rates of recombination can vary between species,

22 between and within chromosomes, and between male and female meiosis in both

23 dioecious/gonochoristic and hermaphroditic species [5–7]. We refer to this phenomenon as 'sex

24 differences in the recombination landscape' [1]. Although sex differences in the rate and distribution of

25 recombination appear to be widespread and variable, the causes and consequences of this variation

26 have only recently been investigated in detail [1,6,8].

27

28 Many components of evolutionary processes depend on the sex-averaged rate of recombination.

29 Nevertheless, sex differences in recombination (heterochiasmy) in dioecious populations can have

30 important consequences for the evolution of sex chromosomes. This is because on the sex chromosome

2

31 restricted to the heterogametic sex (i.e. the Y or W chromosome), sex-specific recombination landscapes

32 entirely control the rate of recombination, and thereby influence the scale of recombination

33 suppression surrounding a sex-determining region (SDR) [1]. Recently, heterochiasmy has been

34 proposed as an important factor in maintaining sexually antagonistic (SA) variants on the sex

35 chromosomes even in the absence of recombination modifiers [9]. In particular, SA alleles can spread

36 more easily through populations because of pre-existing male-specific suppression of recombination,

37 rather than recombination suppression evolving as a secondary consequence of the segregation of SA

38 alleles [9]. Variation among species in patterns of heterochiasmy could thus be an important general

39 determinant of the evolution of sex chromosomes and their turnover [1], potentially contributing to

40 differences among lineages in the likelihood of the formation of heteromorphic sex chromosomes, the

41 maintenance of sexually antagonistic polymorphisms and the size of the SDR.

42

43 Several patterns are evident in the characteristics of sexual dimorphism in the recombination landscape.

44 First, although data are limited, many eukaryotes have recombination rates biased towards the tips of

45 chromosomes in male meiosis, whereas female recombination rates are more likely to be either

46 elevated towards the centromeres or are more uniform across the chromosome [1]. In hermaphroditic

47 plants, three of five taxa studied show this pattern [10–12], although in maize there was limited

48 evidence for large-scale differences in recombination between male and female meiosis [13], and in an

49 interspecific cross between *Solanum esculentum and S. pennellii* recombination in male gametes was

50 reduced genome-wide compared with female gametes [14]. Preliminary analysis of genetic maps in the

51 dioecious *Mercurialis annua* do not suggest major sex differences in recombination rates [15], although

52 the genomic context of these maps is still being investigated. In general, however, recombination rate

53 landscape dimorphism has not yet been investigated in dioecious plants, limiting our understanding of

54 its potential role in the evolutionary dynamics of plant sex chromosomes.

55

56 Many species have convergently evolved tip-biased recombination in male meiosis [1], but the reasons

57 for this pattern are unclear. Both non-adaptive and adaptive explanations have been proposed. If

58 recombination landscape differences are not adaptive, they may simply result from mechanistic

59 differences between the process of female and male meiosis [1]. Adaptive hypotheses include sexually

60 antagonistic selection favoring tighter linkage between sex-specific genes and regulatory elements [1],

61 selection favoring recombination near the centromere in female meiosis to remove meiotic drive alleles

62 [7], and epistatic haploid selection on male gametes and gametophytes [6]. In plants, evidence that

Submitted to Phil. Trans. R. Soc. B - Issue

Page 6 of 27

3

63   female recombination rates are elevated relative to male recombination rates in outcrossing species

64   compared with selfing species [6] is consistent with models of both female meiotic drive and male

65   gametophytic selection, as both are expected to be more intense with higher rates of outcrossing [6,7].

66   Disentangling these alternatives is challenging and will require more comparative information on sex-

67   specific recombination in both hermaphroditic and dioecious taxa.

68

69   *Rumex hastatulus* is a dioecious, wind-pollinated plant with heteromorphic sex chromosomes [16,17].

70   Recent genome sequencing combined with high marker-density linkage mapping has revealed that the

71   SDR is embedded within a very large genomic region of highly suppressed recombination [18]. Evidence

72   for similarly large non-recombining regions in the pericentromeric regions of autosomes suggested that

73   pre-existing recombination suppression may have contributed to the formation of large heteromorphic

74   sex chromosomes in *R. hastatulus* [18]. However, this study measured sex-averaged recombination

75   rates, limiting our ability to investigate the potential role of heterochiasmy in sex chromosome

76   formation and maintenance. With earlier evidence for an important role for gametophytic selection on

77   the sex ratio in this species [19,20], the influence of pollen competition in the evolution of the sex

78   chromosomes [21], and indications of frequent male and female transmission distortion in related

79   dioecious *Rumex* taxa [22], there is a strong likelihood that haploid selection in males and/or females

80   may contribute to sex-specific selection favouring sexual dimorphism in recombination landscapes in

81   this system.

82

83   Here, we explore the potential importance of heterochiasmy for the evolution of sex chromosomes and

84   test hypotheses concerning the evolutionary forces favouring sex-specific recombination rate

85   differences in *R. hastatulus*. Using a sex-specific linkage map and corrected draft genome assembly, we

86   first determine whether *R. hastatulus* shows evidence for heterochiasmy and other sex differences in

87   recombination landscape and compare this pattern between the sex chromosome and the autosomes.

88   We then examine the correlates of male and female recombination rates genome-wide and use this

89   information to explore the potential role of sexual antagonism, haploid selection, and meiotic drive as

90   evolutionary drivers of sexual dimorphism in the recombination landscape.

91

4

Methods

Linkage mapping and genome assembly

94 We generated a mapping population from a cross between a male and female derived from a single

95 population collected in Rosebud, TX [20]. Seeds from the field collection were grown in the glasshouse

96 and at onset of flowering one male and one female individual were randomly paired for a controlled

97 cross to develop the F_1 generation. Paired plants were immediately moved into miniature crossing

98 chambers [23] to avoid pollen contamination from other plants growing in the same glasshouse, and F_1

99 seeds were harvested after maturation. To obtain tissue from F_1 plants, we sterilized seeds using 5%

100 (V/V) bleach and germinated them on filter paper in refrigerated petri dishes. After germination, we

101 transplanted seedlings into six-inch plastic pots containing a 3:1 ratio of Promix soil and sand and a slow-

102 release fertilizer (Nutricote, 14:13:13, 300mL per 60lbs) and placed them in a glasshouse at the

103 University of Toronto, St. George campus. We watered plants every other day, and their positions on

104 benches were randomized weekly. On day 43 or 44 after transplant, between 10:00 and 12:00, we

105 collected and flash-froze 30mg of leaf tissue for RNA extraction using liquid nitrogen. When plants

106 flowered, we phenotyped for sex. We used Spectrum Plant Total RNA Kits (Sigma Aldrich) for RNA

107 extraction. The sequenced library included 188 individuals: 102 female offspring, 84 male offspring, and

108 three replicate samples of each parent.

109

110 For library preparation and sequencing, we sent our RNA samples to the Centre d'expertise et de

111 services Génome Québec (CES, McGill University, Montréal, QC, Canada). CES prepared libraries using

112 NEBNext library prep kits and sequenced them on a NovaSeq6000 S4 PE100. Sequencing resulted in a

113 total of 3.1 billion reads (3,060,570,370), with between 10 and 49 million reads per sample (mean

114 15,940,471, median 14,548,490). Raw sequence has been deposited on the Sequence Read Archive

115 (SRA) under the accession number PRJNA692236 (embargoed until July 1, 2022 or publication).

116

117 We aligned our raw sequencing reads to the *R. hastatulus* Dovetail draft genome assembly [18] using

118 Star 2-pass version 2.7.6 [24,25]. We processed the aligned files to sort, mark PCR duplicates, and

119 (splitNCigar reads) using PicardTools (http://broadinstitute.github.io/picard/) and the Genome Analysis

120 Tool Kit [26].

5

121   We initially generated a linkage map using Lep-Map3 [27] from reads aligned to the original *R.*

122   *hastatulus* draft assembly. The markers could be split into five linkage groups using a LOD score limit of

123   30 for the initial split followed by a LOD score limit of 34 for the resulting largest linkage group

124   (SeparateChromosomes2). Most of the remaining single markers were put into these groups using a LOD

125   score limit of 25 (JoinSingles2All), totaling about 120,000 markers. We then calculated the marker order

126   for each linkage group with OrderMarkers2 (with default settings).

127

128   To improve our linkage map, we reduced redundancy in our genome assembly and constructed a new

129   pseudo-chromosome assembly using the Lep-Anchor [28] software. To obtain reliable linkage maps, we

130   removed the 13 most-recombining individuals from the maps and constructed three independent

131   linkage maps (Lep-MAP3: OrderMarkers2), using only male informative markers (parameter

132   informativeMask=1), only female informative markers (informativeMask=2) and all markers. These maps

133   were used by the Lep-Anchor software.

134

135   To reduce redundancy in the genome assembly, we first split the existing Dovetail assembly into contigs

136   based on assembly gaps. Due to a high number of contigs (>44,000), we removed all contigs of < 500bp,

137   full length haplotypes and joined partial haplotypes in windows of five adjacent contigs (link strength

138   was 6 - |distance| - |difference in orientations|, where distance between contig i and j is |i-j| and

139   difference is 0-2 based on how the orientations differ: same=0, one different=1, both different=2). This

140   was done by Lep-Anchor giving it only the alignment chain computed by Haplomerger2 [29] on the

141   WindowMasker [30] soft-masked (contig-split) genome. This allowed us to reduce the number of contigs

142   to about 33,000. All data were lifted to the new contig assembly coordinates using the liftover script and

143   LiftoverHaplotypes module in Lep-Anchor.

144

145   We then ran Lep-Anchor (lepanchor_wrapper2.sh) on the final contig assembly using the three linkage

146   maps, new alignment chains (HaploMerger2) and alignments of raw Pacbio sequence aligned by

147   minimap2 [31]. The resulting pseudo-chromosome assembly was used to calculate physical order of

148   linkage map markers and the maps were evaluated (OrderMarkers2 parameter evaluateOrder) in this

149   order to obtain the final linkage maps. After assembly improvement, we compared contig orders

150   between our previous [18] and new maps using custom R scripts incorporating Plotly [32] interactive

151   plotting.

6

Recombination rates and transmission distortion

153 We quantified recombination rates in two ways. First, we described recombination using map lengths in

154 centimorgans (cM) from the maps produced by Lep-Map3. Based on the scale of recombination

155 observed in previous work [18] and the current map, we performed all downstream analyses using 1Mb

156 windows. We also calculated recombination rates as the sum of crossover events per 1Mb window. We

157 first calculated the number of crossovers per site from cM differences using the inverse of the Haldane

158 mapping function [33], then summed crossovers in 1Mb windows. To describe the extent of

159 recombination suppression, we identified the total number of consecutive windows with zero

160 crossovers. We estimated transmission ratio distortion as a likelihood ratio of the deviation from 1:1

161 transmission of haplotypes from the male and female parent using custom scripts by PR.

162

Gene and TE content

164 We also developed a new annotation using MAKER version 3.01.03 [34]. For our MAKER annotation, we

165 used the soft-masked [30] genome integrated with previously published floral transcriptomes from six

166 individuals [21] and leaf transcriptomes from six populations [17]. Transcripts were assembled with

167 IDBA-tran version 1.2.0 [35] and annotated in four rounds, using the transcripts and the Tartary

168 buckwheat annotation version FtChromosomeV2.IGDBv2 [36] as the evidences for MAKER. We

169 functionally annotated the final annotation based on homology using BLAST version 2.2.28+ [37] and

170 InterProScan 5.52-86.0 [38]. This annotation resulted in 59,121 genes. We also annotated the locations

171 of rDNA repeats using rnammer-1.2 [39]. The parameters used, '-S euk' and '-m tsu,ssu,lsu', indicate that

172 the input reference is a eukaryote, and that we are annotating 5/8s, 16/18s, and 23/28s rDNA.

173

174 We produced the TE annotation using the EDTA (Extensive de-novo TE Annotator) version 1.9.7 pipeline

175 [40]. This pipeline combines the best-performing structure- and homology-based TE finding programs

176 (LTR_FINDER_parallel [41], LTR_HARVEST_parallel [42], LTR_retriever [43], TIR-Learner2.5 [44],

177 HelitronScannerv1.1[45], Repeatmodeler-2.0.1 [46] and RepeatMasker-4.1.1 [47] and filters their results

178 to produce a comprehensive and non-redundant TE library [40]. The optional parameters '--sensitive 1'

179 and '--anno 1' were used to identify remaining unidentified TEs with RepeatModeler and to produce an

180 annotation. The 'EDTA.TEanno.split.gff3' output file was used as our non-overlapping TE annotation. This

181 file is produced by EDTA by removing overlaps according to the following priorities: structure-based

182   annotation > homology-based annotation, longer TE > shorter TE > nested inner TE > nested outer TE

183   [40].

184

185   For all gene content analyses, we used a stringently filtered set of genes to remove gene annotations

186   associated with transposable elements. We first used BEDtools [48] to remove any exons that

187   overlapped a TE, although genes containing both exons that overlapped TEs and exons that did not

188   overlap TEs were retained. We then removed any gene functionally annotated with 'transpos*'

189   (transposon, transposase, etc.), 'ribonuclease H,' 'pol poly,' 'mitochondri*,' 'chloroplast,' or 'retrovirus.'

190   This filtered annotation contained 30,641 genes.

191

192   Differential expression and SNP calling

193   We performed differential expression analyses using DESeq2 (v. 1.28.1) [49] and our new annotation.

194   For DESeq2 analyses, we aligned reads to the new genome pseudomolecules using STAR version 2.7.6a

195   [49] and generated readcounts using featureCounts (2.02) [50]. Cutoffs for differential expression were

196   as follows: adjusted p-value <0.1, absolute Log2Fold change >1. We identified genes that were

197   differentially expressed between male and female leaf tissues using published leaf RNA sequence data

198   from population samples of the XY cytotype [17], and between male and female floral tissue using

199   published RNA sequence data from the XY cytotype [21]. Genes with fewer than 20 reads across all

200   samples were removed from these analyses. We also identified sequences that were differentially

201   expressed in pollen tissue compared to male leaf tissue, using published sequence data [21]. Finally, we

202   identified sequences differentially expressed in pollen tubes compared to pollen, using pollen from the

203   individuals in the mapping population described [18]. We collected pollen using a kief box (Wacky

204   Willy's, Victoria, BC, Canada), germinated and grew it in 100μL of media [51] for 24 hours, and flash

205   froze it in LN2. After removing media, we lysed cells and extracted total RNA using Spectrum Plant Total

206   RNA Kits (Sigma Aldrich) for RNA extraction. To identify sex-linked SNPs and fixed differences between

207   the X and Y chromosome (all females homozygous reference or non-reference, all males heterozygous)

208   for our new assembly, we used FreeBayes v0.9.10-3-g47a713e [52] to call SNPs from the population

209   transcriptome data from the XY cytotype (six males and six females) [17] and the crossing transcriptome

210   data from the same cytotype (six male and six female offspring, plus parents) [17]. We filtered the SNPs

211   to exclude any with a SNP quality score of lower than 60, any sites with missing data, and fixed

212   heterozygous SNPs across all samples that likely reflected paralogous mapping.

8

213

Linear modelling predictors of recombination rate

215 We combined our linkage map data with our annotation, TE annotation, differential expression data and

216 summed and averaged variables in 1-Mb windows to perform analyses of recombination landscape,

217 gene content, and differential expression in R version 4.1.0 [53] in RStudio version 1.4.1717 [54] using

218 the packages dplyr version 1.0.7 [55] and stringr version 1.4 [56]. We performed correlations using R's

219 built-in cor function, and estimated partial correlations using the package ppcor version 1.1 [57].

220

221 To identify factors associated with genome structure that predicted recombination rates and

222 recombination rate differences [58], we created linear models with the following response variables:

223 female crossovers per window, male crossovers per window, sex-averaged crossovers per window,

224 crossover number sex difference per window, and female vs. male biased recombination across window.

225 We fit all responses using generalized linear models with either negative binomial or Tweedie

226 distributions except for female versus male biased recombination, for which we used logistic regression.

227 We performed linear models in glmmtmb version 1.12 [59], evaluated fit using DHARMa version 0.4.3

228 [60], and compared models using ANOVAs. We performed separate models for each response variable

229 on each chromosome. Scripts are available at https://github.com/joannarifkin/Rumex-sex-specific.

230

Results

231

Linkage mapping and genome assembly improvement

233 We identified five linkage groups, consistent with both the karyotype of the XY cytotype of this species

234 [16, 61] and our previous sex-averaged linkage mapping results [18] (table 1). We again identified two

235 apparently metacentric linkage groups (A1 and A2) and three apparently submetacentric linkage groups

236 (A3, A4, XY) based on the patterns of recombination across the chromosomes (figure 1A) and the

237 identities of the scaffolds constituting the linkage groups. We have retained the same autosomal labels

238 across both maps, and they continue to reflect chromosome sizes from largest (A1) to smallest (A4).

239

Submitted to Phil. Trans. R. Soc. B - Issue

Page 12 of 27

9

240 Our larger genetic mapping population and improved genome assembly led to considerable

241 improvement in higher-order chromosome-scale scaffolding of the genome of *R. hastatulus*. Our

242 improved genome assembly contained 1.45Gb, a reduction of 0.2Gb from our previous 1.65Gb assembly

243 [18] due to the collapsing of redundant haplotypes (see Methods). For this assembly, 1.212Gb (84%) is

244 now grouped in the five linkage groups (previously 1.08GB, 65% of the previous primary assembly), with

245 the remaining 0.23Gb in smaller contigs. These corrections have substantially increased the size of the

246 assembled sex chromosome, with an additional 88.6 MB of sequence assembled on the sex

247 chromosome, the largest increase of any of the chromosomal scaffolds (table 1). Consistent with this

248 increase, analysis of our past genome assembly showed that only 52% of sex-linked SNPs mapped to the

249 assembled sex chromosome; our new assembly integrated with transcriptomes from independent

250 crossing data [17] now shows that 94% of SNPs showing X-Y segregation patterns map to the sex

251 chromosome.

Recombination rates

253 As in our previous study, we found that recombination was unevenly distributed across the genome,

254 with very large non-recombining regions on all chromosomes (figure 1). We identified clear evidence of

255 heterochiasmy (table 1, figure 1A). Male map lengths were shorter than female map lengths: across

256 chromosomes, female map length was 1.4x male map length (table 1, figure 1A), and the sex

257 chromosome was not an obvious outlier for this metric. Males also had longer blocks of non-

258 recombining windows across all chromosomes. To summarize this pattern, we identified the longest

259 stretches of markers on each chromosome with zero crossovers. On the autosomes, males had runs of

260 non-recombining windows approximately twice as long as those of females, with male-specific non-

261 recombining regions as large as 238MB (table 1). By this measure, the sex chromosomes were an

262 exception: the largest run of male-specific non-recombining windows on the sex chromosome was four

263 times the length of the longest run of female-specific non-recombining sequence (table 1). Thus,

264 although all chromosomes exhibited reduced male recombination rates, the XY chromosome showed

265 the largest region of differentially suppressed recombination between the sexes despite not being the

266 largest chromosome. Overall, male and female recombination rates in *R. hastatulus* were only weakly

267 correlated (*r*=0.333, correlation of male and female crossover number across 1Mb windows across all

268 chromosomes; figure 2, table S1).

269

Page 13 of 27

Submitted to Phil. Trans. R. Soc. B - Issue

10

270 The extent of sex differences in recombination varied both along and between chromosomes (figure

271 1B). Chromosomes A1, A2, and the sex chromosome conformed to the common pattern of more tip-

272 focused recombination in males, but the submetacentric chromosome A3 showed female-biased

273 recombination in the more highly recombining end, and A4 appeared to have low-recombination

274 regions at both ends. This differs from the previous linkage map, likely because of the difficulty of

275 positioning low-recombination regions. In general, pericentromeric regions showed reduced

276 recombination in both sexes, but female map lengths were larger in these regions and showed apparent

277 hotspots of recombination with large jumps in centimorgan position (figure 1). All five chromosomes

278 had regions of female-biased and male-biased recombination, as well as shared recombining and non-

279 recombining regions (figure 1B). This complex pattern creates a highly heterogeneous recombination

280 landscape.

281

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

282 *Values in brackets indicate the lengths from the previous assembly

283

284

11

Gene and TE content

285

286 As in our previous study, we found that genes were generally concentrated in high-recombination

287 regions (figure 1D). However, A1 contained one gene-dense yet low-recombination region (~100Mb-

288 200Mb). RNA (class 1) TEs were concentrated in low-recombination regions, whereas DNA (class 2) TEs

289 were concentrated in high-recombination regions (figure 1E). Despite the reduced gene density in low-

290 recombination regions, the extent of high recombination-suppressed regions means that a large fraction

291 of genes in the genome (approximately 37%) are in these large regions with no male recombination.

292 Ribosomal genes were concentrated mostly on A3 (62 rDNA features annotated) and A4 (28 rDNA

293 features annotated). rDNA genes occurred in the first 50Mb of A3 (5S subunit sequence) with additional

294 rDNA sequence located around 130Mb of A3 (18S and 28S subunit sequence) and in the first 3 Mb of A4

295 (18S and 28S subunit sequence). These rDNA locations are consistent with past cytological findings [61],

296 further confirming our identification of the A3 and A4 chromosomes.

297

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

298

299 With female and male recombination separated, it is clear that the male-specific non-recombining

300 region on the sex chromosome in *R. hastatulus*, i.e. the SDR, is extensive. In particular, our linkage

301 mapping suggests that the SDR is as large as 209MB (14% of the total assembly: table 1, figure 1),

302 including as many as 3595 genes (12% of the total filtered annotated genes). The gene-dense

303 recombining pseudoautosomal region of the sex chromosome is similarly the smallest male recombining

304 segment of any chromosome, representing only approximately 13% of the physical size of the

305 chromosome. Note that the 'true' SDR may be narrower and the pseudoautosomal region larger, since

306 rare male recombination may have gone unobserved in our cross. However, BLAST [37] searches of our

307 sex-linked transcripts that have at least one fixed difference between the X and Y chromosomes from a

308 population sample [17] are found across most of the length of this nonrecombining region (from 1.6 MB

309 to 208.5 MB), and fixed differences mapped onto the chromosome are common across the first 210 MB

310 (figure S1) suggesting that most of this region is nonrecombining and linked to the SDR.

Transmission ratio distortion

311

312 We identified loci with biased haplotype transmission through either paternal or maternal inheritance

313 based on the haplotype reconstructions from Lep-Map 3. Transmission ratio distortion varied between

12

314   chromosomes (figure 1C, figure S2, table S2). More sites experienced biased transmission through

315   maternal (276) than paternal (48) inheritance, across a larger fraction of the genome. On A1, 18 sites

316   were significantly distorted in transmission from males using a 0.05 cutoff in a chi-squared distribution

317   (LOD > 3.841) and on A4, 30 sites were significantly distorted in transmission from males with a 0.05

318   cutoff. Although deviations from 1:1 male haplotype transmission occurred on other chromosomes,

319   there were no significant male-distorted sites on A2, A3, or the sex chromosome. In contrast, female

320   haplotype distortion was extensive on A2 (91 sites in females at 0.05, 38 at 0.01 cutoff of LOD > 6.635),

321   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

322   negligible on A1 (0 sites) and A4 (1 site at 0.01).
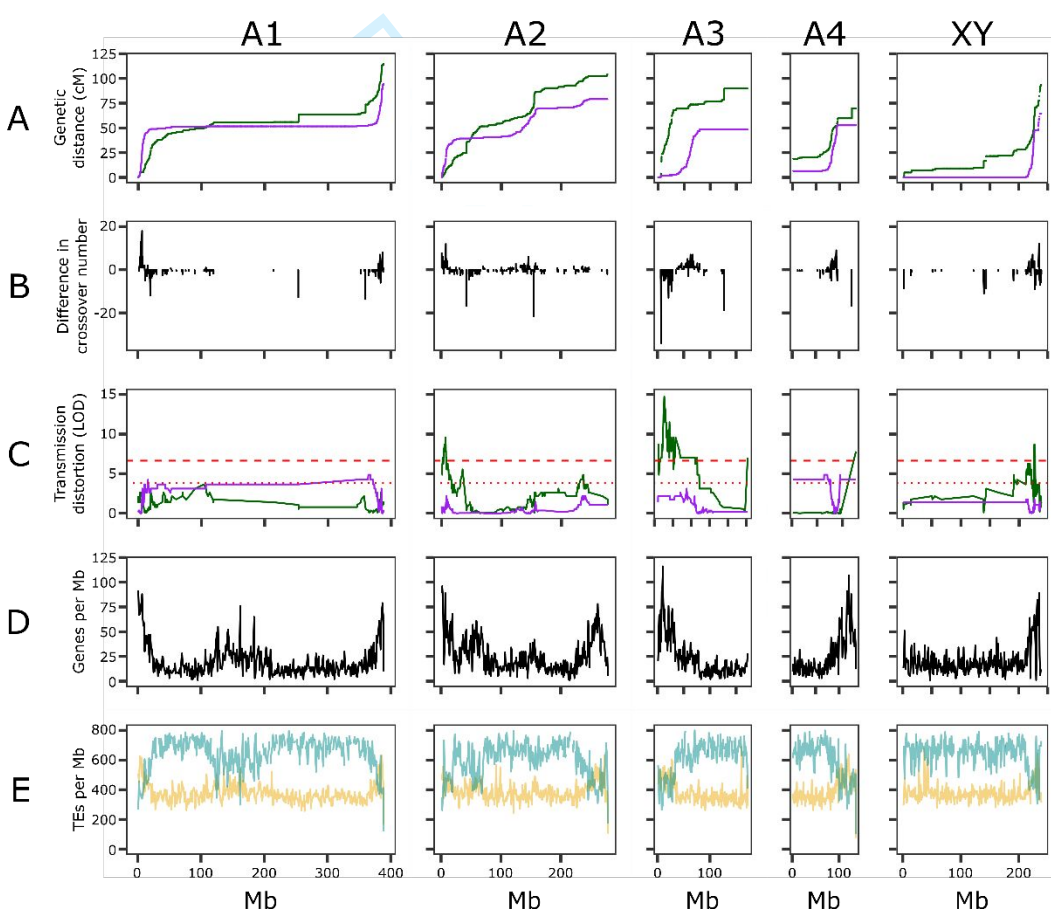

323

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

330

13

Correlates of recombination rate differences

331

332 Across the genome, the number of genes, leaf- and flower-expressed genes, and DNA (Class 2) TE

333 density were all positively correlated with recombination rate, and RNA (Class 1) TE density was

334 negatively correlated with recombination rate (figure 2, table S1, figure S3-S6). Male crossover number

335 and female crossover number were both positively correlated with gene density, but this correlation

336 was stronger for male crossovers (male Pearson's r=0.390, female Pearson's r=0.278). However, the

337 correlations between transmission ratio distortion and crossover number were in opposite directions: in

338 females, more distorted regions were also more recombining (r=0.111) whereas in males distortion was

339 negatively correlated with crossover number (r=-0.197), reflecting the fact that male transmission

340 distortion signals were enriched in the large non-recombining regions of male meiosis (figure 1).

341 Correlations varied in strength between chromosomes, with notable differences in correlates of

342 transmission ratio distortion and recombination rate difference, both of which varied in magnitude,

343 position, and direction along and between chromosomes (figure 2, figure S7, table S3). The difference in

344 crossover number between the sexes was most strongly correlated with signals of male distortion

345 (figure 2), where regions of particularly low male crossover number represented regions with larger

346 signals of male distortion. We also estimated partial correlation coefficients controlling for gene density,

347 which was consistently correlated with many genomic variables (figure S8, figure S9, tables S4-S5).

348 Correlations between distortion and crossover number in both males and females persisted after

349 controlling for gene density, and varied in direction across chromosomes (figure S9). However, all

350 chromosomes except A2 showed a consistent negative correlation between male-biased recombination

351 and male transmission distortion, even when gene density was controlled.

352

353

14
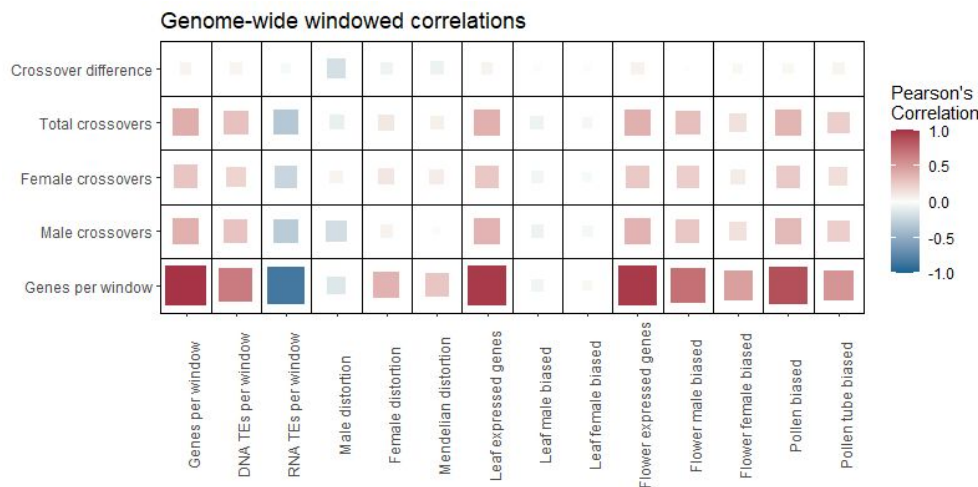
354 Figure 2: Correlations among genome window characteristics across the whole genome of *Rumex hastatulus*.
355 Colours and sizes correspond to the strength and direction of the correlations.
356

Linear models of recombination rate differences

358 We used generalized linear models to identify predictors of sex-specific and sex-averaged recombination
359 rates, whether recombination was male- or female-biased, and the magnitude of the recombination rate
360 difference in *R. hastatulus*. Full modeling results are available in table S6 and table S7.
361

362 In our linear models, the variables that significantly predicted male recombination rate varied between
363 chromosomes (table 2A). Number of genes predicted increased male recombination rate on three
364 chromosomes (as well as a fourth in some models), and position along the genome emerged as an
365 important predictor on two submetacentric chromosomes, suggesting that distance from centromere
366 predicts male recombination (figure 1). RNA TE count also appeared to play a role, but in inconsistent
367 directions, predicting increased male recombination on two chromosomes and decreased male
368 recombination on a third; this pattern also appears in the partial correlations with gene density removed
369 (figure S9). Transmission ratio distortion predicted reduced male recombination rate on three
370 chromosomes (A2, A3, and A4), and predicted increased male recombination on A1. Either pollen bias or
371 pollen tube bias predicted male recombination rate on three chromosomes (negatively on A4 and the
372 sex chromosome, positively on A3).
373

374 In contrast, the predictors of female recombination are more consistent across chromosomes (table 2B).
375 Number of genes per window positively predicted female recombination rate on all five chromosomes,
376 and position relative to centromere affected female recombination rate on four chromosomes. Only one
377 other variable, pollen tube-biased expression, predicted female recombination rate on more than one
378 other chromosome. Female transmission ratio distortion did not emerge as a significant predictor of
379 recombination rate on any chromosome.
380

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

15

386

387 Finally, we used logistic regression to identify variables that predicted whether windows exhibited

388 female or male recombination bias (table S6C). Our logistic regressions also suggested considerable

389 variation in the factors that predicted sex differences in recombination. Number of genes per window

390 was an important predictor for four out of five chromosomes, but in variable directions, predicting both

391 male bias (A1, XY) and female bias (A2, A3) in recombination.

392

393

394 **Table 2**

395 Variables identified as significant predictors of recombination in linear models of windows of the *Rumex*

396 *hastatulus* genome.

397

398 **Male recombination rate predictors**

399 '+' - variable predicts increased male recombination. '-' - variable predicts decreased male

400 recombination. * - 95% confidence interval for slope estimate includes zero, chiefly because term is

401 significant in interaction term.

402

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

403

404 Interaction effects in best model:

405 A1: n RNA TEs:male distortion (-)

406 A2: n genes:male distortion (+ ns)

407 A3: male distortion:position window (+)

408 A3 (more complex model): position window:n genes (- ns), male distortion:position window (+), n

409 genes:male distortion, n genes:position window:male distortion (+ns)

410 A4: none

411 XY: position window:n genes (-)

412

413

414 **Female recombination rate predictors**

415 '+' - variable predicts increased female recombination. '-' - variable predicts decreased female

416 recombination. * - 95% confidence interval for slope estimate includes zero, chiefly because term is

417 significant in interaction term.

16

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

418

419 Interaction effects in best model:
420 A1: n RNA TEs:pollen tube bias (+)
421 A2: n genes: position window (-ns)
422 A3: n genes: position window (-), leaf expression: position window (+)
423 A4: n genes: position window (- ns)
424 XY: position window:female flower bias (-)
425 XY (more complex model): position window:female flower bias (-), position window:pollen tube bias (-
426 ns), position window:pollen tube bias:flower female bias (+)
427

Discussion

429 The main findings of this study are consistent with the general pattern of extensive recombination

430 suppression that we previously inferred based on sex-averaged recombination in *R. hastatulus* [18].

431 However, our sex-specific maps show that recombination suppression is not evenly distributed between

432 males and females, and that the observed very large pericentromeric regions of suppressed

433 recombination are particularly influenced by highly suppressed male recombination. Across all

434 chromosomes, females recombine more frequently than males and males have much larger non-

435 recombining blocks than females. Our results are in line with several studies of hermaphroditic plants, as

436 well as other eukaryotes [1], with the very large male-specific non-recombining regions that we report

437 making this an extreme case.

438

439 However, *R. hastatulus* does not strictly follow the common eukaryotic pattern of tip-focused male

440 recombination [1] across all chromosomes. The recombination landscapes of both the metacentric and

441 submetacentric chromosomes suggest greater variation in the distribution of recombination than simply

442 less male recombination in the centres and more at the chromosome ends, with highly recombining

443 regions scattered along chromosomes (particularly A2, A3, and A4). Thus, male recombination is more

444 concentrated, but not always at the tips of the chromosome. These differences among chromosomes

Submitted to Phil. Trans. R. Soc. B - Issue
Page 20 of 27

17

445   may reflect an ongoing history of chromosomal rearrangements in the genus and additional patterns of

446   chromosome structure such as the locations of centromeres and rDNA clusters [61].

447

448   The pattern of larger non-recombining regions in males is consistent with an evolutionary bias toward

449   the evolution of male heterogamety and XY sex chromosomes in *Rumex*. In particular, suppressed

450   recombination can facilitate the maintenance and invasion of sexually antagonistic variants linked to

451   sex-determining regions [62], and sex-determining regions that evolve in large pericentromeric non-

452   recombining regions, especially sex-specific ones, may contribute to the evolution of sex chromosomes

453   [63]. The existence of male-specific non-recombining regions may thus facilitate the evolution of XY

454   rather than ZW sex chromosomes [1]. Given our observation of no crossovers in male meiosis over very

455   large fractions of each chromosome, it is possible that recombination suppression was ancestral to the

456   evolution of dioecy in the genus, and that subsequent recombination modifiers did not evolve following

457   the origin of the SDR. The observed size of the SDR, which is over 200 MB using population-validated

458   sex-linked genes and includes over 14% of the assembled genome and over 3500 genes, is much larger

459   than those recently reviewed to date in plants [64], although larger population samples of *R. hastatulus*

460   should be used to test for very rare recombination between the X and Y, and the sex-linked region of *S.*

461   *latifolia* may be even larger [65]. Although we cannot rule out a role for subsequent recombination

462   modifiers, particularly because the sex chromosomes show the most extreme size dimorphism of the

463   nonrecombining region (table 1), our results do suggest that sex differences in heterochiasmy may have

464   played an important role in determining the large size of the SDR facilitating the evolution of large

465   heteromorphic sex chromosomes in this system. Comparative studies of heterochiasmy in both

466   hermaphroditic and other dioecious species in *Rumex* will be important to further assess the extent to

467   which ancestral heterochiasmy and derived changes in recombination rates have contributed to sex

468   chromosome evolution in this lineage.

469

470   Models for the evolution of heterochiasmy due to male haploid selection and female meiotic drive both

471   predict lower overall male recombination rates and higher female recombination near centromeres

472   [1,6,7], as we observed in *R. hastatulus*. Our mapping population provided evidence for both male and

473   female transmission ratio distortion, which varied within and among chromosomes and between the

474   sexes. Overall, more sites displayed significant distortion in female than in male transmission, but

475   significant regions of both types of distortion were observed across the genome. Transmission ratio

476   distortion through female inheritance is generally thought to be consistent with female meiotic drive. In

18

477   contrast, transmission ratio distortion through male inheritance may result from haploid (pollen)

478   competition [66].

479

480   Nevertheless, zygotic distortion (i.e., differential seed germination or seedling survival) could also lead

481   to transmission ratio distortion, and may result from alleles inherited from either parent [66].

482   In our study, we genotyped reproductive adults rather than pollen or seeds, which conflates several

483   opportunities for natural and sexual selection causing biased transmission. Zygotic selection may be

484   particularly likely to explain our observed female distortion, since these regions were not focused on

485   low-recombination centromeric regions, where meiotic drive is expected to act [7]. In contrast, signals

486   of male transmission distortion are particularly enriched in regions of low male recombination and high

487   sex bias in recombination (figure 1 and figure 2, table 2), suggesting that haploid selection in males may

488   be an important selective pressure for sex differences in recombination. Distorted regions can vary

489   widely between populations of the same species [67], so distortion in a single cross should be

490   interpreted with some caution. Furthermore, patterns of biased pollen or pollen tube expression do not

491   show similarly consistent enrichment in regions of low male recombination except for the sex

492   chromosome (table 2, figure S6), although direct observation of transmission distortion likely provides a

493   stronger indicator of loci involved in pollen competition. Similarly, although we did not identify evidence

494   consistent with ongoing recombination increases to counter meiotic drive, the overall pattern of

495   increased and more centromere-biased recombination is still consistent with a history of selection

496   eliminating meiotic drive alleles. With those caveats, our results do provide some evidence in accord

497   with the hypothesis that pollen competition may play an important role in sex differences in

498   recombination.

499

500   We used both pairwise correlations and regression models to investigate various genomic correlates of

501   male and female recombination and sex bias in recombination, in order to further explore other possible

502   factors favouring sex differences in recombination. On a genome-wide scale, both male and female

503   recombination rates in *R. hastatulus* are consistent with widely observed patterns that genes and Class 2

504   DNA TEs concentrate in high-recombination regions and Class 1 RNA TEs concentrate in low-

505   recombination regions (figure 2; [3], [56]). These patterns are consistent with recombination rates in

506   plants occurring preferentially upstream of genes, and with epigenetic silencing of retrotransposable

507   elements causing a reduction of recombination, although they could also be explained if transposable

508   elements preferentially accumulate in regions of low sex-averaged recombination [3].

19

509

510 Because the recombination landscape varied widely between the chromosomes of *R. hastatulus*, we

511 also identified chromosome-specific predictors of recombination rates and recombination rate bias

512 using linear models and partial correlations controlling for gene density. At this more granular scale, a

513 different and more complex picture emerges (table 2, figure S8, S9). In particular, our logistic regression

514 models of recombination bias direction support the possibility that different mechanisms may be

515 contributing to variation in recombination rates on different chromosomes. Both number of genes per

516 window and position along the genome predicted female-biased recombination on some chromosomes

517 and male-biased recombination on others. Aside from physical position and gene density, different

518 diverse factors predicted both male and female bias on different chromosomes, including RNA TEs,

519 number of genes expressed in different tissues, sex-biased floral expression, pollen-biased expression,

520 and transmission ratio distortion (table 2C). This finding suggests a general picture in which gene density

521 and proximity to centromere shape recombination on a 'global scale', but variation in gene content and

522 haploid selection may lead to region-specific selective forces acting on both male and female

523 recombination rates. Consistent with this, a recent comparative study in fish [68] found that sex

524 differences in recombination are labile at the species level, but do not present clear trends consistent

525 with adaptive hypotheses across species.

526

527

Conclusions

529 Our study has provided some of the first evidence of sex differences in recombination and identified one

530 of the largest known SDRs in a dioecious plant species, or in fact in any eukaryote [65]. We identified

531 both genome-wide and chromosome-specific factors predicting sex differences in recombination, and

532 also found evidence consistent with a role for male gametophytic selection in driving these differences.

533 Future work in this system will allow more precise dissection of the genetic and evolutionary

534 mechanisms favouring sex differences in recombination landscapes. In particular, exploring both sex-

535 specific eQTL positions and further study of transmission ratios in pollen and seeds will allow us to

536 further differentiate between sexually antagonistic cis epistasis in diploids and epistasis in haploids [1].

537 Finally, characterizing sex-specific recombination landscapes of hermaphroditic *Rumex* species should

538 make it possible to determine whether these sex differences in recombination landscape did indeed

Submitted to Phil Trans R Soc B - Issue

20

539 precede and promote the evolution of a heterogametic XY sex determining system with a very large

540 SDR, as we have hypothesized.

541

Acknowledgments

Data accessibility statement

549 Raw sequence has been deposited on the Sequence Read Archive (SRA) under the accession number

550 PRJNA692236 (embargoed until July 1, 2022 or publication). Our new genome assembly, transcriptome

551 annotation, rDNA annotation, and TE annotation have been deposited in the CoGe comparative

552 genomics platform at https://genomevolution.org/coge/GenomeInfo.pl?gid=62326. Scripts used in the

553 analyses have been deposited on Github at https://github.com/joannarifkin/Rumex-sex-specific.

554
555

556

References

558

559 1. Sardell JM, Kirkpatrick M. 2020 Sex differences in the recombination landscape. *Am. Nat.* **195**,
560   361–379.
561 2. Campos JL, Halligan DL, Haddrill PR, Charlesworth B. 2014 The relation between recombination
562   rate and patterns of molecular evolution and variation in *Drosophila melanogaster*. *Mol. Biol.*
563   *Evol.* **31**, 1010–1028.
564 3. Kent TV, Uzunović J, Wright SI. 2017 Coevolution between transposable elements and
565   recombination. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **372**, 20160458.
566 4. Nachman MW, Payseur BA. 2012 Recombination rate variation and speciation: theoretical
567   predictions and empirical results from rabbits and mice. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*
568   **367**, 409–421.
569 5. Stapley J, Feulner PGD, Johnston SE, Santure AW, Smadja CM. 2017 Variation in recombination
570   frequency and distribution across eukaryotes: patterns and processes. *Philos. Trans. R. Soc.*
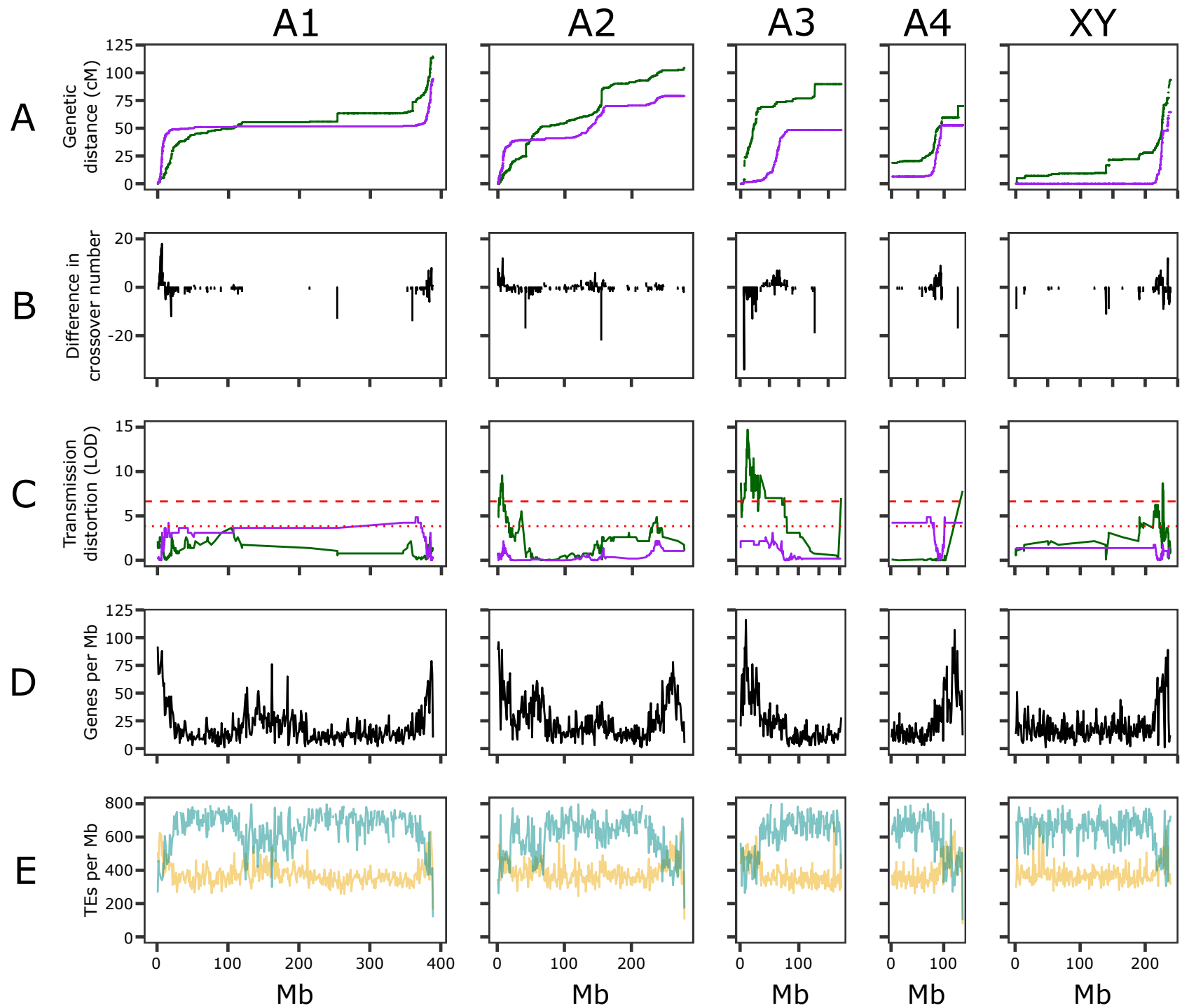
21

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

572 6. Lenormand T, Dutheil J. 2005 Recombination difference between sexes: a role for haploid
573   selection. *PLoS Biol.* **3**, e63.

574 7. Brandvain Y, Coop G. 2012 Scrambling eggs: meiotic drive and the evolution of female
575   recombination rates. *Genetics* **190**, 709–723.

576 8. Johnston SE, Huisman J, Ellis PA, Pemberton JM. 2017 A high-density linkage map reveals sexual
577   dimorphism in recombination landscapes in Red Deer (*Cervus elaphus*). *G3* **7**, 2859–2870.

578 9. Bergero R, Gardner J, Bader B, Yong L, Charlesworth D. 2019 Exaggerated heterochiasmy in a
579   fish with sex-linked male coloration polymorphisms. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 6924–
580   6931.

581 10. Giraut L, Falque M, Drouaud J, Pereira L, Martin OC, Mézard C. 2011 Genome-wide crossover
582   distribution in *Arabidopsis thaliana* meiosis reveals sex-specific patterns along chromosomes.
583   *PLoS Genet.* **7**, e1002354.

584 11. Lagercrantz U, Lydiate DJ. 1995 RFLP mapping in *Brassica nigra* indicates differing recombination
585   rates in male and female meioses. *Genome*. **38**, 255–264.

586 12. Phillips D *et al.* 2015 The effect of temperature on the male and female recombination
587   landscape of barley. *New Phytol.* **208**, 421–429.

588 13. Kianian PMA *et al.* 2018 High-resolution crossover mapping reveals similarities and differences
589   of male and female recombination in maize. *Nat. Commun.* **9**, 2370.

590 14. de Vicente MC, Tanksley SD. 1991 Genome-wide reduction in recombination of backcross
591   progeny derived from male versus female gametes in an interspecific cross of tomato. *Theor.
592   Appl. Genet.* **83**, 173–178.

593 15. Veltsos P *et al.* 2019 Early sex-chromosome evolution in the diploid dioecious plant *Mercurialis
594   annua*. *Genetics* **212**, 815–835.

595 16. Smith BW. 1964 The evolving karyotype of *Rumex hastatulus*. *Evolution* **18**, 93–104.

596 17. Hough J, Hollister JD, Wang W, Barrett SCH, Wright SI. 2014 Genetic degeneration of old and
597   young Y chromosomes in the flowering plant *Rumex hastatulus*. *Proc. Natl Acad. Sci. USA* **111**,
598   7713–7718.

599 18. Rifkin JL, Beaudry FEG, Humphries Z, Choudhury BI, Barrett SCH, Wright SI. 2021 Widespread
600   recombination suppression facilitates plant sex chromosome evolution. *Mol. Biol. Evol.* **38**,
601   1018–1030.

602 19. Field DL, Pickup M, Barrett SCH. 2012 The influence of pollination intensity on fertilization
603   success, progeny sex ratio, and fitness in a wind-pollinated, dioecious plant. *Int. J. Plant Sci.* **173**,
604   184–191.

605 20. Pickup M, Barrett SCH. 2013 The influence of demography and local mating environment on sex
606   ratios in a wind-pollinated dioecious plant. *Ecol. Evol.* **3**, 629–639.

607 21. Sandler G, Beaudry FEG, Barrett SCH, Wright SI. 2018 The effects of haploid selection on Y
608   chromosome evolution in two closely related dioecious plants. *Evol Lett* **2**, 368–377.

609 22. Wilby AS, Parker JS. 1988 Mendelian and non-Mendelian inheritance of newly-arisen
610   chromosome rearrangements. *Heredity* **60**, 263–268.

611 23. McGoey BV, Janik R, Stinchcombe JR. 2017 Individual chambers for controlling crosses in
612   wind-pollinated plants. *Methods Ecol. Evol.* **8**, 887–891.

613 24. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR.
614   2013 STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21.

615 25. Dobin A, Gingeras TR. 2016 Optimizing RNA-Seq Mapping with STAR. *Methods Mol. Biol.* **1415**,
616   245–262.

617 26. Van der Auwera GA, O'Connor BD. 2020 *Genomics in the Cloud: Using Docker, GATK, and WDL in
618   Terra*. Newton, MA: O'Reilly Media, Inc.

22

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.

23

667   *Bioinformatics* **26**, 841–842.

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

670 50. Liao Y, Smyth GK, Shi W. 2014 featureCounts: an efficient general purpose program for assigning
671   sequence reads to genomic features. *Bioinformatics* **30**, 923–930.

672 51. Adhikari KN, Campbell CG. 1998 In vitro germination and viability of buckwheat (*Fagopyrum*
673   *esculentum* Moench) pollen. *Euphytica* **102**, 87–92.

674 52. Garrison E, Marth G. 2012 Haplotype-based variant detection from short-read sequencing.
675   *arXiv preprint arXiv:1207.3907.

676 53. R Development Core Team. 2021 R: A language and environment for statistical computing. R
677   Foundation for Statistical Computing, Vienna, Austria. See https://www.R-project.org/.
678   (accessed on 2021).

679 54. R Team. 2020 RStudio: Integrated Development for R. RStudio, PBC, Boston, MA, 2020.

680 55. Hadley Wickham RF, Henry L, Müller K. 2017 dplyr: A Grammar of Data Manipulation. R package
681   version 0.7. 4.

682 56. Wickham H. 2019 stringr: Simple, consistent wrappers for common string operations. R package
683   version 1.4.0. See https://CRAN.R-project.org/package=stringr (accessed on 2019).

684 57. Kim, S. (2015) ppcor: An R Package for a fast calculation to semi-partial correlation coefficients.
685   *Commun. Stat. Appl. Methods*, **22(6)**, 665-674.

686 58. Paape T, Zhou P, Branca A, Briskine R, Young N, Tiffin P. 2012 Fine-scale population
687   recombination rates, hotspots, and correlates of recombination in the *Medicago truncatula*
688   genome. *Genome Biol. Evol.* **4**, 726–737.

689 59. Brooks ME, Kristensen K, Van Benthem KJ, Magnusson A, Berg CW, Nielsen A, Skaug HJ, Machler
690   M, Bolker BM. 2017 glmmTMB balances speed and flexibility among packages for zero-inflated
691   generalized linear mixed modeling. *R J.* **9**, 378–400.

692 60. Hartig F. 2019 DHARMa: residual diagnostics for hierarchical (multi-level/mixed) regression
693   models. R package version 0.1. 0.

694 61. Kasjaniuk M, Grabowska-Joachimiak A, Joachimiak AJ. 2019 Testing the translocation hypothesis
695   and Haldane's rule in *Rumex hastatulus*. *Protoplasma* **256**, 237–247.

696 62. Otto SP. 2019 Evolutionary potential for genomic islands of sexual divergence on recombining
697   sex chromosomes. *New Phytol.* **224**, 1241–1251.

698 63. Charlesworth, D. 2019 Young sex chromosomes in plants and animals. *New Phyt.* **224**, 1095-
699   1107.

700 64. Renner SS, Müller NA. 2021 Plant sex chromosomes defy evolutionary models of expanding
701   recombination suppression and genetic degeneration. *Nat Plants* **7**, 392–402.

702 65. Gschwend, AR, Weingartner, LA, Moore, RC, and R Ming. 2012 The sex-specific region of sex
703   chromosomes in animals and plants. *Chromosome Res.* **20**, 57–69.

704 66. Fishman L, McIntosh M. 2019 Standard Deviations: The biological bases of transmission ratio
705   distortion. *Annu. Rev. Genet.* **53**, 347–372.

706 67. Seymour DK, Chae E, Arioz BI, Koenig D, Weigel D. 2019 Transmission ratio distortion is frequent
707   in *Arabidopsis thaliana* controlled crosses. *Heredity* **122**, 294–304.

708 68. Cooney CR, Mank JE, Wright AE. 2021 Constraint and divergence in the evolution of male and
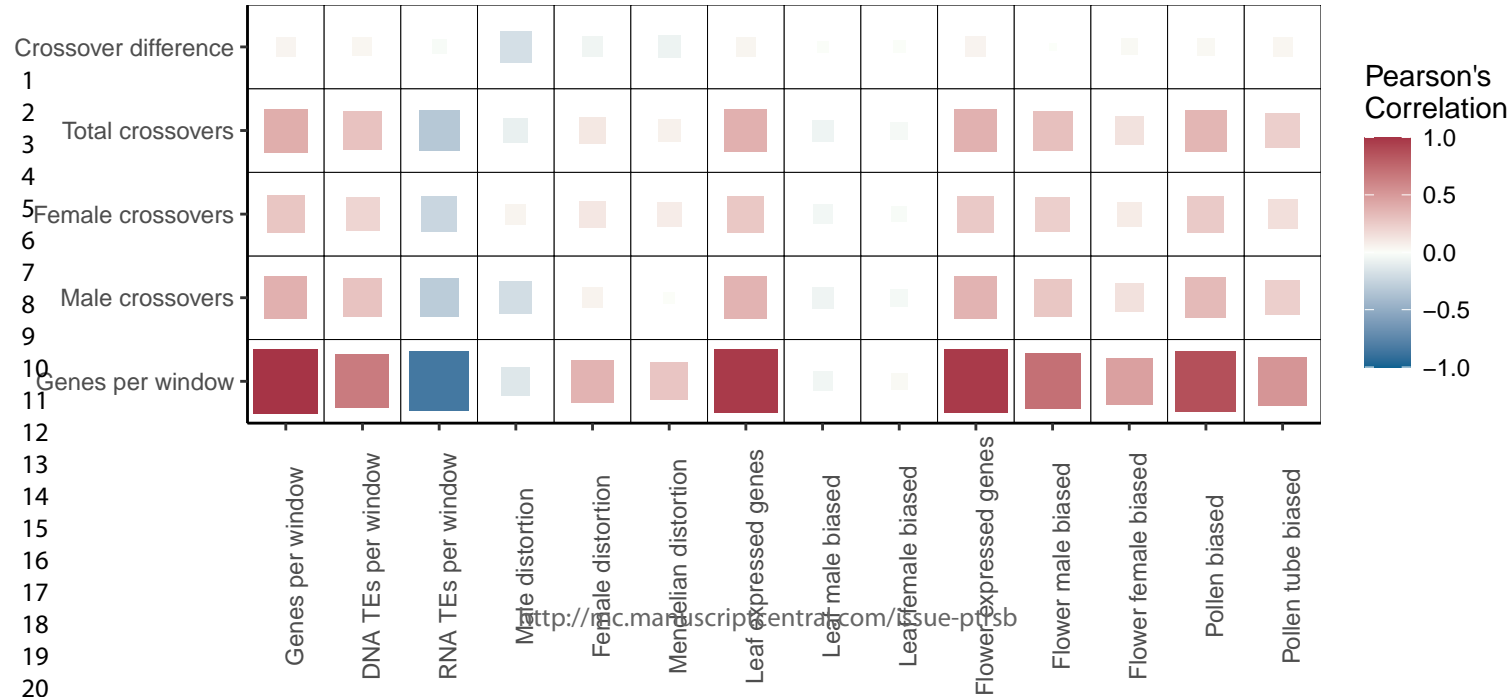709   female recombination rates in fishes. *Evolution* https://doi.org/10.1111/evo.14357

710

711

Page 27 of 27

Submitted to Phil. Trans. R. Soc. B - Issue

Genome−wide windowed correlations