

## Gene flow mediates the role of sex chromosome meiotic drive during complex speciation

Colin D. Meiklejohn<sup>1\*</sup>, Emily L. Landeen<sup>2,4</sup>, Kathleen E. Gordon<sup>1,5</sup>, Thomas Rzatkiewicz<sup>2</sup>, Sarah B. Kingan<sup>2,6</sup>, Anthony J. Geneva<sup>2,7</sup>, Jeffrey P. Vedanayagam<sup>2,8</sup>, Christina A. Muirhead<sup>2</sup>, Daniel Garrigan<sup>2,9</sup>, David L. Stern<sup>3</sup>, and Daven C. Presgraves<sup>2\*</sup>

<sup>1</sup> School of Biological Sciences, University of Nebraska, Lincoln, Nebraska, 68588 USA

<sup>2</sup> Department of Biology, University of Rochester, Rochester, New York, 14627, USA

<sup>3</sup> Janelia Research Campus, Howard Hughes Medical Institute, 19700 Helix Drive, Ashburn, Virginia, 20147, USA

### \*Contact authors:

[cmeiklejohn2@unl.edu](mailto:cmeiklejohn2@unl.edu), [daven.presgraves@rochester.edu](mailto:daven.presgraves@rochester.edu)

### Current addresses:

<sup>4</sup> Department of Integrative Biology, University of California, Berkeley, California, 94720, USA

<sup>5</sup> Department of Molecular Biology and Genetics, Field of Genetics, Genomics, and Development, Cornell University, Ithaca, New York, 14853, USA

<sup>6</sup> Pacific Biosciences of California, Menlo Park, CA USA

<sup>7</sup> Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, 02138, USA

<sup>8</sup> Department of Developmental Biology, Sloan-Kettering Institute, New York, NY 10065, USA

<sup>9</sup> AncestryDNA, 153 Townsend Street, San Francisco, CA 94107 USA

1   **ABSTRACT**

2   During speciation, sex chromosomes often accumulate interspecific genetic  
3   incompatibilities faster than the rest of the genome. The drive theory posits that sex  
4   chromosomes are susceptible to recurrent bouts of meiotic drive and suppression, causing  
5   the evolutionary build-up of divergent cryptic sex-linked drive systems and, incidentally,  
6   genetic incompatibilities. To assess the role of drive during speciation, we combine high-  
7   resolution genetic mapping of X-linked hybrid male sterility with population genomics  
8   analyses of divergence and recent gene flow between the fruitfly species, *Drosophila*  
9   *mauritiana* and *D. simulans*. Our findings reveal a high density of genetic  
10   incompatibilities and a corresponding dearth of gene flow on the X chromosome.  
11   Surprisingly, we find that, rather than contributing to interspecific divergence, a known  
12   drive element has recently migrated between species, caused a strong reduction in local  
13   divergence, and undermined the evolution of hybrid sterility. Gene flow can therefore  
14   mediate the effects of selfish genetic elements during speciation.

15

16   **INTRODUCTION**

17   Speciation involves the evolution of reproductive incompatibilities between diverging  
18   populations, including prezygotic incompatibilities that prevent the formation of hybrids  
19   and postzygotic incompatibilities that render hybrids sterile or inviable. Two patterns  
20   characterizing speciation implicate a special role for sex chromosomes in the evolution of  
21   postzygotic incompatibilities: Haldane's rule, the observation that hybrids of the  
22   heterogametic sex preferentially suffer sterility and inviability (Haldane, 1922, Wu and  
23   Davis, 1993, Orr, 1997, Laurie, 1997, Price and Bouvier, 2002, Presgraves, 2002, Coyne  
24   and Orr, 2004); and the large X-effect, the observation that the X chromosome has a  
25   disproportionately large effect on hybrid sterility (Coyne and Orr, 1989, Coyne, 1992a,  
26   Presgraves, 2008). These patterns hold across a wide range of taxa, including female  
27   heterogametic (*ZW*) birds and Lepidoptera and male heterogametic (*XY*) plants,  
28   *Drosophila*, and mammals (Coyne and Orr, 1989, Coyne and Orr, 2004). We now know  
29   that these "two rules of speciation" (Coyne and Orr, 1989) are, in part, attributable to the  
30   rapid evolution of genetic factors that cause interspecific hybrid sterility on the X

31 chromosome relative to the autosomes (Tao and Hartl, 2003, Moehring et al., 2007,  
32 Masly and Presgraves, 2007, Presgraves, 2008, Good et al., 2008). The relatively rapid  
33 accumulation of X-linked hybrid sterility factors is associated with reduced interspecific  
34 gene flow at X-linked *versus* autosomal loci {Muirhead, 2016 #3060} {Turissini, 2017  
35 #3300} {Garrigan, 2012 #2715}. Overall, these patterns show that, for many taxa with  
36 heteromorphic sex chromosomes, the X chromosome plays a large and fundamental role  
37 in speciation.

38 Given the taxonomic breadth of Haldane's rule, the large X-effect, and reduced  
39 interspecific gene flow on the X, understanding *why* the X chromosome accumulates  
40 hybrid incompatibilities faster than the rest of the genome is imperative. At least five  
41 explanations have been proposed: faster X evolution (Charlesworth et al., 1987), gene  
42 traffic (Moyle et al., 2010), disrupted sex chromosome regulation in the germline  
43 (Lifschytz and Lindsley, 1972), the evolutionary origination of incompatibilities in  
44 parapatry (Hollinger and Hermisson, 2017), and meiotic drive (Hurst and Pomiankowski,  
45 Frank, 1991).

46 Here we focus on the potential role of meiotic drive. The drive theory posits that sex  
47 chromosomes are more susceptible than autosomes to invasion by selfish meiotic drive  
48 (*sensu lato*) elements (Hurst and Pomiankowski, 1991, Frank, 1991). Sex-linked drive  
49 compromises fertility and distorts sex ratios, which leads to evolutionary arms races  
50 between drivers, unlinked suppressors, and linked enhancers (Lindholm et al., 2016,  
51 Presgraves, 2008, Meiklejohn and Tao, 2010). These arms races can contribute to the  
52 evolution of hybrid male sterility, in at least two ways. Normally-suppressed drive  
53 elements might be aberrantly expressed in the naïve genetic backgrounds of species  
54 hybrids, causing sterility rather than sex ratio distortion (Hurst and Pomiankowski, 1991,  
55 Frank, 1991). Alternatively, recurrent bouts of invasion, spread, and coevolution among  
56 drive, suppressor, and enhancer loci might cause interspecific divergence at these loci  
57 that secondarily cause hybrid sterility and map disproportionately to sex chromosomes  
58 (Presgraves, 2008, Meiklejohn and Tao, 2010).

59 Multiple lines of evidence support the plausibility of the drive theory. First, theoretical  
60 considerations and empirical evidence suggests that both active and suppressed sex

61 chromosome meiotic drive systems are widespread in natural populations (Jaenike, 2001).  
62 Indeed, in one species, *Drosophila simulans*, at least three cryptic (normally suppressed)  
63 *sex-ratio* drive systems—Winters, Durham, and Paris—have been identified, involving  
64 distinct sets of X-linked drive loci and autosomal and/or Y-linked suppressors (Tao et al.,  
65 2001, Tao et al., 2007a, Tao et al., 2007b, Helleu et al., 2016). Second, loci involved in  
66 cryptic *sex-ratio* systems co-localize with hybrid male sterility loci in genetic mapping  
67 experiments (Tao et al., 2001, Zhang et al., 2015). Third, one of the two X-linked hybrid  
68 sterility genes identified to date also causes meiotic drive (Phadnis and Orr, 2008). These  
69 discoveries confirm that recurrent bouts of drive and suppression have occurred during  
70 the history of a single lineage and that cryptic drive genes can cause hybrid sterility.  
71 While these findings put the plausibility of the drive hypothesis beyond doubt, the  
72 question of its generality remains: what fraction of X-linked hybrid sterility factors  
73 evolved as a consequence of drive? Furthermore, the drive hypothesis (Hurst and  
74 Pomiankowski, 1991, Frank, 1991) assumes that populations evolve in strict allopatry  
75 (simple speciation) and/or that drive elements require population-specific genetic  
76 backgrounds for their activity. For populations that diverge with some level of gene flow  
77 (complex speciation), drive elements that are not population-specific can in principle  
78 migrate between species, thereby reducing divergence and undermining the evolution of  
79 hybrid sterility (Macaya-Sanz et al., 2011, Crespi and Nosil, 2013, Seehausen et al.,  
80 2014).

81 Here we investigate the special role of sex chromosomes in speciation with genetic  
82 mapping and population genomic analyses between *Drosophila mauritiana* and *D.*  
83 *simulans*. The human commensal species, *D. simulans*, originated on Madagascar,  
84 diverging from the sub-Saharan African species, *D. melanogaster*, ~3 Mya (Lachaise et  
85 al., 1988, Dean and Ballard, 2004, Baudry et al., 2006, Kopp, 2006, Ballard, 2004). The  
86 island-endemic species, *D. mauritiana*, originated on the Indian Ocean island of  
87 Mauritius, diverging from *D. simulans* ~240 Kya (Kliman et al., 2000, McDermott and  
88 Kliman, 2008, Garrigan et al., 2012). The two species are now isolated by geography—  
89 *D. simulans* has never been collected on Mauritius {David, 1989 #2349}—and by  
90 multiple incomplete reproductive incompatibilities, including asymmetric premating  
91 isolation (Coyne, 1992b), postmatting-prezygotic isolation (Price, 1997), and intrinsic

92 postzygotic isolation ( $F_1$  hybrid males are sterile,  $F_1$  hybrid females are fertile; (Lachaise  
93 et al., 1986)). Despite geographic and reproductive isolation, there is clear evidence for  
94 historical gene flow between the two species (Solignac and Monnerot, 1986, Solignac et  
95 al., 1986, Garrigan et al., 2012, Ballard, 2000a, Ballard, 2000b, Satta et al., 1988, Satta  
96 and Takahata, 1990). The X chromosome shows both an excess of factors causing hybrid  
97 male sterility (True et al., 1996b, Tao et al., 2003) and, correspondingly, a dearth of  
98 historical interspecific introgression (Garrigan et al., 2012). The rapid accumulation of  
99 X-linked hybrid male sterility factors may have contributed to reduced X-linked gene  
100 flow, limiting exchangeability at sterility factors and genetically linked loci (Muirhead  
101 and Presgraves, 2016).

102 To begin to assess the role of drive in the evolution of X-linked hybrid male sterility  
103 between these two species, we performed genetic mapping experiments using genotype-  
104 by-sequencing of advanced-generation recombinant X-linked introgressions from *D.*  
105 *mauritiana* in an otherwise pure *D. simulans* genetic background. In parallel, we  
106 performed the first population genomics analyses of speciation between *D. mauritiana*  
107 and *D. simulans* to study the chromosomal distributions of interspecific divergence and  
108 introgression. These analyses lead to two discoveries regarding the role of meiotic drive  
109 in speciation. First, we find evidence for weak X-linked segregation distortion in hybrids,  
110 supporting the hypothesis that cryptic *sex-ratio* systems are common. Second, we report  
111 a surprising discovery concerning the role of drive during complex speciation: a now-  
112 cryptic sex ratio drive system recently introgressed between species and caused large  
113 selective sweeps in both species. As a result, this large X-linked region shows greatly  
114 reduced interspecific sequence divergence and an associated lack of hybrid male sterility  
115 factors. These findings suggest that the effects of selfish genetic elements on  
116 interspecific divergence and the accumulation of incompatibilities depends on the  
117 opportunity for drive systems to migrate between species during complex speciation.

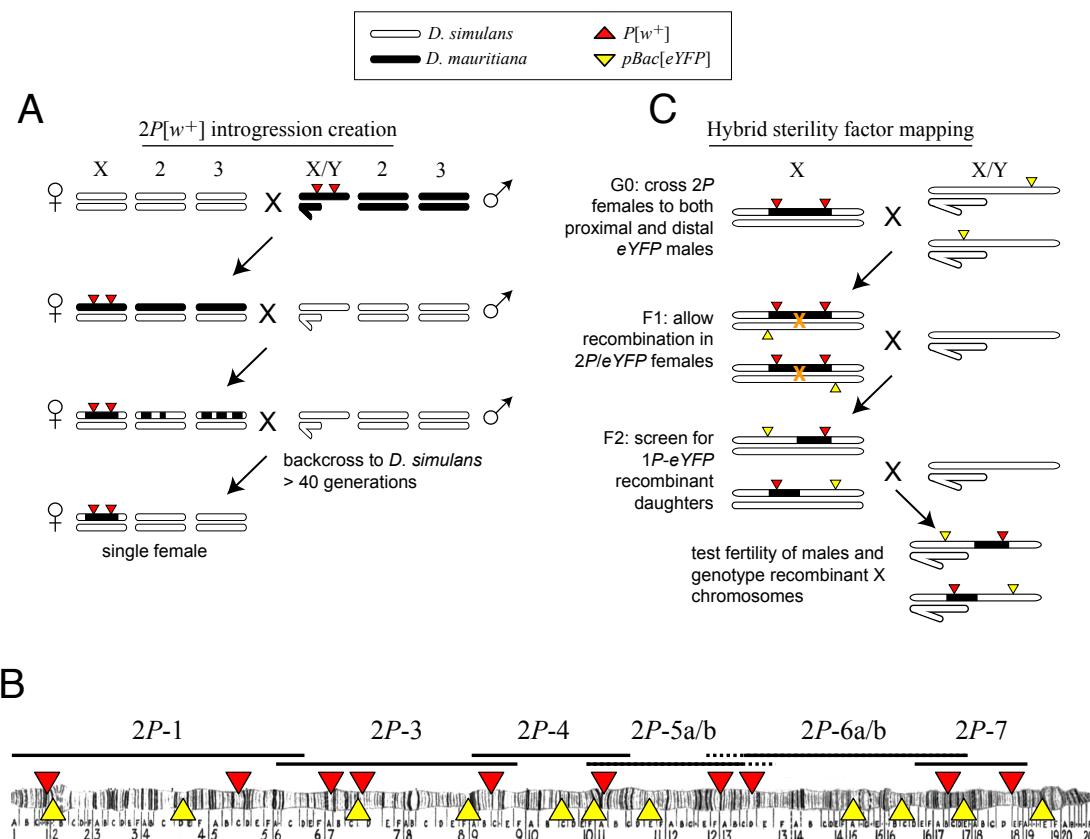
118

## 119 **RESULTS**

120 ***Mapping X-linked hybrid male sterility***

121 Multiple intervals on the X chromosome cause male sterility when introduced from *D.*  
122 *mauritiana* into *D. simulans* (True et al., 1996b, Maside et al., 1998). The number and  
123 identities of the causal factors, how they disrupt spermatogenesis, and the evolutionary  
124 forces that drove their interspecific divergence are unknown. We therefore generated a  
125 high-resolution genetic map of X-linked hybrid male sterility between the two species,  
126 with the ultimate aim of identifying a panel of sterility factors. We first introgressed  
127 eight X-linked *D. mauritiana* segments that together tile across ~85% of the euchromatic  
128 length of the X chromosome into a *D. simulans* genetic background (**Figure 1A,B; Table**  
129 **1**). Each introgressed segment was marked by two co-dominant *P[w<sup>+</sup>]* insertions (True et  
130 al., 1996a) that serve as visible genetic markers. We introgressed these “2P” segments  
131 into the *D. simulans* *w<sup>XD1</sup>* genetic background through >40 generations of repeated  
132 backcrossing (**Figure 1A**). Our ability to generate these introgression genotypes  
133 confirms that the distal 85% of the *D. mauritiana* X chromatin carries no dominant  
134 factors that cause female sterility or lethality in a *D. simulans* genetic background (True  
135 et al., 1996b, Tao et al., 2003). All eight 2P introgression genotypes are however  
136 completely male-sterile, indicating that each of the introgressed regions contains one or  
137 more hybrid male sterility factors. Two pairs of introgression genotypes carry largely  
138 overlapping introgressed *D. mauritiana* segments and were combined for further analyses  
139 (2P-5a/b and 2P-6a/b, respectively; **Figure 1B, Table 1**).

140 To determine the genetic basis of male sterility within each 2P interval, we generated  
141 recombinant introgressions using *D. simulans* strains carrying *pBac[eYFP]* visible  
142 markers (Stern et al., 2017) (**Figure 1C**). These crosses capture unique recombination  
143 events between *P[w<sup>+</sup>]* and *pBac[eYFP]* markers, allowing recombinant *D. mauritiana*  
144 introgressions (hereafter called 1P-YFP) to be propagated indefinitely through females  
145 without recombination via selection for the 1P-YFP genotype. From these 1P-YFP  
146 females, an unlimited number of replicate males carrying identical 1P-YFP recombinant  
147 introgressions can be generated, assayed for male fertility, and archived for genotyping  
148 (**Figure 1C**; see below). We assayed male fertility in at least ten individual males from  
149 each of 617 recombinant 1P-YFP genotypes (**Table 2**; see **Methods**). Across 1P-YFP  
150 genotypes, the mean number of offspring ranged from zero to 215 progeny; 238 (38.6%)  
151 were completely male-sterile, producing no offspring, and an additional 62 (10%)

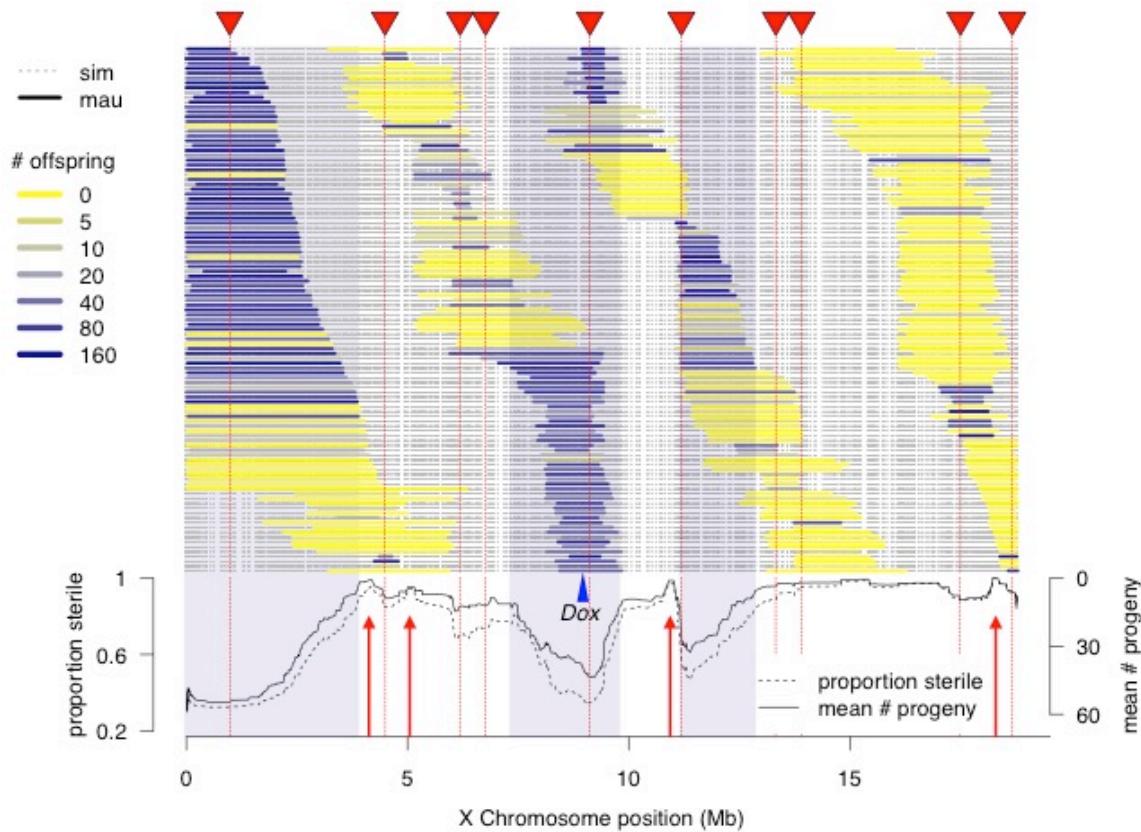


**Figure 1.** Crosses used to introgress eight regions of the *D. mauritiana* X chromosome into a *D. simulans* genome. (A) *D. mauritiana* “2P” lines were constructed by combining pairs of *P*-element insertions containing the miniwhite transgene ( $P[w^+]$ ; red triangles) distributed across the X chromosome. The  $P[w^+]$  inserts are semi-dominant visible eye-color markers that permit discrimination of individuals carrying 0, 1 or 2  $P[w^+]$ . X-linked segments from *D. mauritiana* were introgressed into a *D. simulans* genetic background by backcrossing 2P $[w^+]$  hybrid females to *D. simulans*  $w^{XD1}$  males for over 40 generations. Each introgression line was then bottlenecked through a single female to eliminate segregating variation in the recombination breakpoints flanking the 2P $[w^+]$  interval. (B) Cytological map of the *D. melanogaster* X chromosome, indicating the locations of  $P[w^+]$  and  $pBac[eYFP]$  transgene insertions. The extent of regions introgressed from *D. mauritiana* into *D. simulans* (e.g. 2P-1) are labeled above the map. Two pairs of introgression genotypes (2P-5a/b and 2P-6a/b) mostly overlap; the regions included in 2P-5b/2P-6b but not 2P-5a/2P-6a are indicated by dashed lines. (C) Meiotic mapping of sterility factors. 2P $[w^+]$  females were crossed to *D. simulans* strains carrying an X-linked  $pBac[eYFP]$  transgene (yellow triangles) that was used as an additional visible marker to score recombinant chromosomes. Recombinant X chromosomes with both  $pBac[eYFP]$  and a single  $P[w^+]$  were chosen and assayed for male fertility. Recombinant chromosomes were generated using  $pBac[eYFP]$  markers both proximal and distal to each 2P introgression.

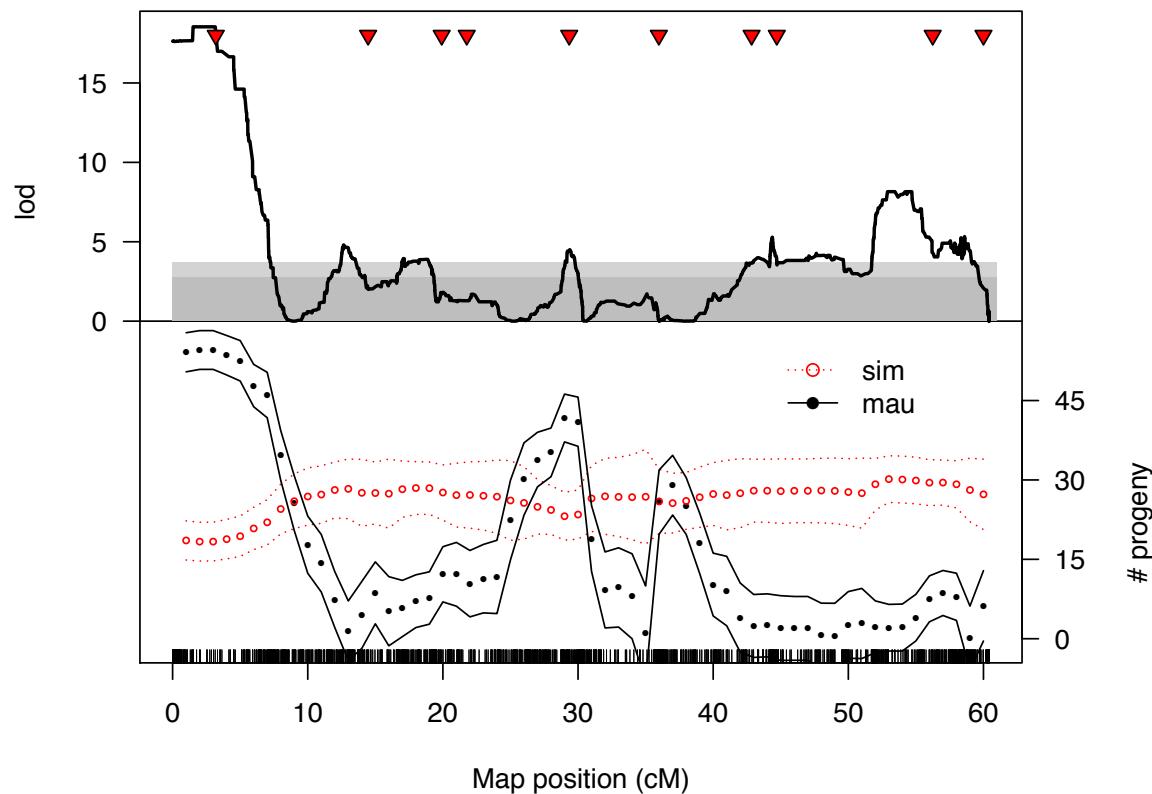
152 produced fewer than five offspring per male (**Figure S1**). Of the remaining *1P-YFP*  
153 genotypes, 231 (37.4%) had intermediate fertility, and 86 (13.9%) had fertility  
154 indistinguishable from within-species controls ( $P_{t-test} > 0.01$ ).

155 We determined high-resolution genotypes of *1P-YFP* recombinant introgressions using  
156 multiplexed whole-genome sequencing (Andolfatto et al., 2011). After quality filtering,  
157 we obtained high-confidence genome-wide genotype information for 439 *1P-YFP*  
158 recombinant introgressions. No genotype showed evidence for any autosomal *D.*  
159 *mauritiana* alleles, confirming that the introgression scheme isolated X-linked *D.*  
160 *mauritiana* segments in a pure *D. simulans* autosomal genetic background (**Figure S2**).  
161 Recombinant *1P-YFP* introgressions on the X chromosome ranged in size from 0.219 to  
162 6.32 Mbp, with a mean length of 1.97 Mb (**Table 3**). **Figure 2** shows the distribution of  
163 *D. mauritiana* introgression segments and their corresponding sterility phenotypes.  
164 Three large regions on the *D. mauritiana* X chromosome can be introgressed into *D.*  
165 *simulans* without severely negative effects on male fertility, indicating an absence of  
166 major hybrid male sterility factors in these regions (**Figure 2**). Conversely, we  
167 delineated four small regions (<700kb) that consistently and strongly reduced male  
168 fertility: 90% of replicate males with introgressions spanning these regions produce fewer  
169 than five offspring. Quantitative trait locus (QTL) analyses confirmed the existence of  
170 genetic variation among introgression genotypes that significantly affects male fertility  
171 (**Figure 3**). At least five QTL peaks are significant at  $P < 0.01$  (permutation test). Most  
172 regions containing *D. mauritiana* alleles reduce the average number of progeny to <15.  
173 Two QTL peaks (2.5cM, and 29.3cM, **Figure 3**) appear to show higher fertility  
174 associated with the *D. mauritiana* allele than the *D. simulans* allele, but this is  
175 attributable to *D. mauritiana* sterility factors located at 12.6cM and ~35cM and the  
176 negative linkage disequilibrium that is generated across a 2P interval by our meiotic  
177 mapping approach (**Figure 1C**). Similar results are obtained with QTL analyses that treat  
178 each 2P interval as a separate mapping population (**Figure S3**).

179  
180  
181



**Figure 2.** High-resolution genetic map of X-linked hybrid male sterility. Colored horizontal bars indicate the extent of introgressed *D. mauritiana* alleles for each recombinant 1P-YFP X chromosome. The color of each introgression indicates the mean fertility of ten replicate males carrying that 1P-YFP X chromosome. The three shaded areas indicate fertile regions within which *D. mauritiana* introgressions do not cause sterility, whereas the four red arrows indicate small candidate sterility regions. The blue arrowhead indicates the location of the Dox/MDox meiotic drive loci. Lines indicate the average number of offspring and average proportion of sterile males (defined as producing fewer than 5 offspring) for all 1P-YFP genotypes that carry *D. mauritiana* alleles at each genotyped SNP.

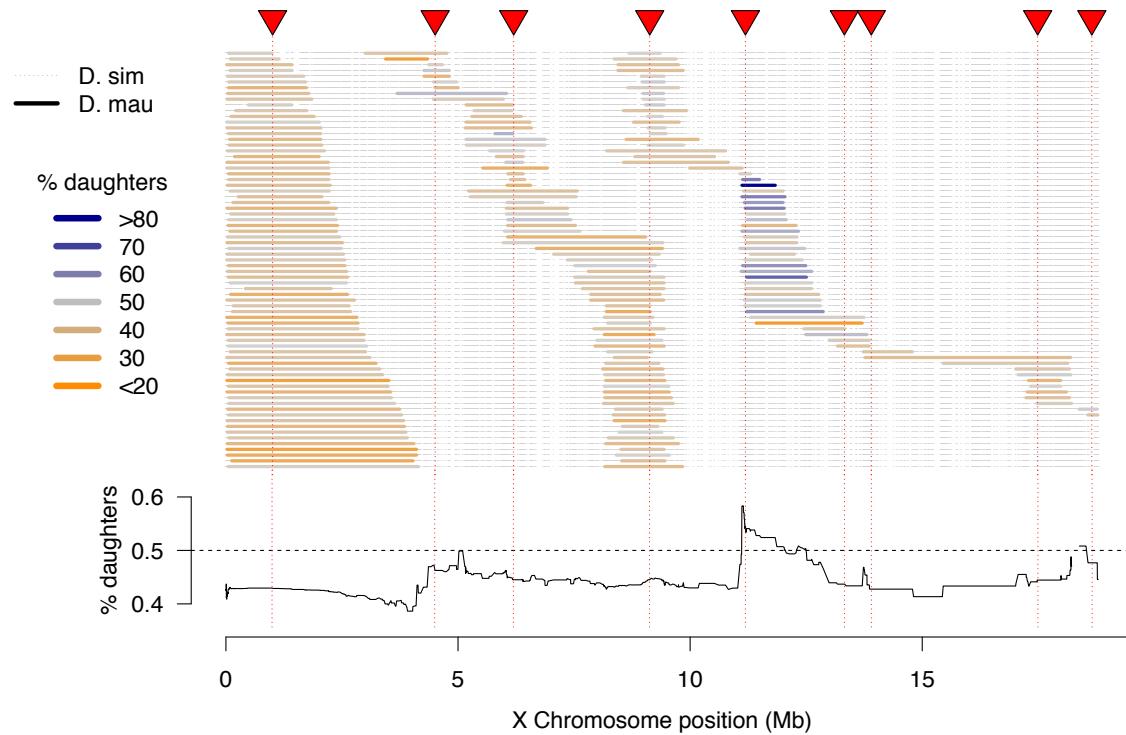


**Figure 3.** QTL analysis of male fertility. (A) Lod scores for a single-factor scan using all 1P-YFP introgression genotypes as a single mapping population. Dark and light gray regions indicate 5% and 1% significance thresholds, respectively, determined from 10,000 random permutations. (B) The estimated effects of *D. simulans* and *D. mauritiana* alleles at QTL placed every 1cM (bounding lines indicate 95% confidence intervals).

182 **Sex ratio distortion revealed through experimental introgression**

183 Among fertile 1P-YFP males, progeny sex ratios were skewed toward a slight excess of  
184 sons: the mean proportion of daughters was 0.45, and 86% of fertile 1P-YFP genotypes  
185 (260/303) produced fewer than 50% daughters (**Figure 4**). However, these skewed sex  
186 ratios are at least partially attributable to effects of the *sim w<sup>XD1</sup>* genetic background, as a  
187 similar male bias was observed among progeny of control *sim w<sup>XD1</sup>* males (mean  
188 proportion females = 0.46,  $n = 35$  sires, *t*-test vs. null hypothesis of 0.5,  $P = 0.005$ ).  
189 Introgressed *D. mauritiana* alleles may modify this modest male bias — across all fertile  
190 introgression genotypes, there is a significant negative correlation between the length of  
191 the introgressed segment and the proportion of female progeny produced by that  
192 genotype ( $r=-0.31$ ,  $P<0.0001$ ). This effect seems to be independent of a relationship  
193 between the length of introgressed segments and fertility; while there is also a significant  
194 negative correlation between introgression length and mean number of progeny ( $r=-0.22$ ,  
195  $P<0.0001$ ), there is not a significant correlation between mean number of progeny and  
196 progeny sex ratio ( $P=0.36$ ). One interpretation of these results is that the *Y* chromosome  
197 of *sim w<sup>XD1</sup>* causes weak segregation distortion, and the intensity of distortion is modified  
198 by X-linked alleles from *D. mauritiana*.

199 Although the majority of fertile 1P-YFP genotypes sired male-biased progeny,  
200 introgressions that included the distal end of the 2P-5 region sired female-biased progeny  
201 (**Figure 4**). QTL analysis of progeny sex ratio confirms a significant peak in the distal  
202 portion of 2P-5 (**Figure S.4**). The estimated effect of this QTL on progeny sex ratios is  
203 54.6% daughters for the *mauritiana* allele *versus* 42.5% daughters for the *simulans* allele.  
204 These results are consistent with the existence of a cryptic (normally-suppressed) X-  
205 linked drive allele in *D. mauritiana* that is released in a *D. simulans* genetic background,  
206 as the *D. mauritiana* *w<sup>12</sup>* strain used to generate the 2P introgressions produces slightly  
207 male-biased progeny sex-ratios using the same fertility assay (one male paired with three  
208 *D. simulans* *w<sup>XD1</sup>* females,  $n = 10$  sires, mean sex-ratio = 0.47, *t*-test vs. *D. simulans* *w<sup>XD1</sup>*  
209  $P = 0.4$ ). This region of the X chromosome does not contain any previously mapped  
210 meiotic drive loci in *D. simulans* (Montchamp-Moreau et al., 2006, Tao et al., 2007a,  
211 Helleu et al., 2016), suggesting that these experiments have uncovered a novel cryptic



**Figure 4.** High-resolution map of progeny sex ratios among fertile 1P-YFP introgression male genotypes. Colored horizontal bars indicate the extent of introgressed *D. mauritiana* alleles for each fertile recombinant 1P-YFP X chromosome. The color of each introgression indicates the sex-ratio of progeny from replicate males carrying that 1P-YFP X chromosome. The line below indicates the average progeny sex-ratio for all 1P-YFP genotypes that carry *D. mauritiana* alleles at each genotyped SNP.

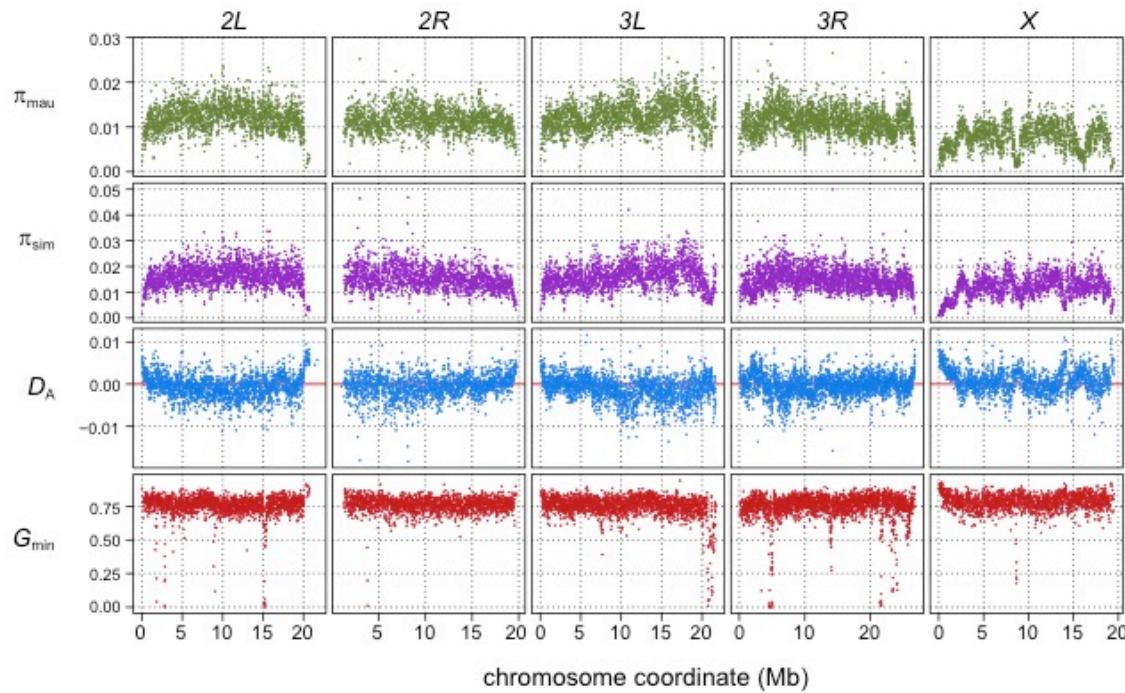
212 drive locus and provide the first evidence of cryptic X-chromosome drive in *D.*  
213 *mauritiana*.

214

215 ***Population genomics of speciation history***

216 The high density of hybrid male sterility factors, and the presence of cryptic drive  
217 systems on the X chromosome is expected to influence patterns of gene flow between *D.*  
218 *mauritiana* and *D. simulans*. We therefore analyzed whole-genome variation within and  
219 between 10 *D. mauritiana* strains (Garrigan et al., 2014) and 21 *D. simulans* strains  
220 (Rogers et al., 2014), which allows us to characterize differentiation and identify genomic  
221 regions with histories of recent interspecific introgression. These analyses complement  
222 earlier studies that characterized interspecific divergence (Garrigan et al., 2012),  
223 polymorphism within *D. mauritiana* (Garrigan et al., 2014, Nolte et al., 2013), and  
224 polymorphism within *D. simulans* (Begun et al., 2007, Rogers et al., 2014). Below we  
225 present genome-wide population genetic analyses using non-overlapping 10-kb windows  
226 (see **Methods**).

227 **Polymorphism:** Our genome-wide analyses provide multiple indicators that the island-  
228 endemic *D. mauritiana* has a smaller effective population size than *D. simulans* (**Table 4**),  
229 consistent with previous multi-locus analyses (Hey and Kliman, 1993, Kliman et al.,  
230 2000). Compared to *D. simulans*, total polymorphism (Nei and Li, 1979) in *D.*  
231 *mauritiana* is 32% lower on the X chromosome and 19% lower on the autosomes (**Figure**  
232 **5**). The X/autosome ratio of polymorphism is thus lower in *D. mauritiana* (0.656) than in  
233 *D. simulans* (0.778) and lower than the  $\frac{3}{4}$  expected for a random mating population with  
234 a 1:1 sex ratio (Garrigan et al., 2014). A substantial fraction of extant polymorphisms in  
235 both species arose in their common ancestor, reflecting the relatively recent speciation  
236 event and the fact that both species have large effective population sizes (see **Methods**).  
237 Compared to *D. simulans*, however, *D. mauritiana* has retained 74.4% as many ancestral  
238 polymorphisms and accumulated just 46.3% as many derived polymorphisms. The site  
239 frequency spectra (Tajima, 1989) in *D. mauritiana* are less skewed towards rare variants  
240 than in *D. simulans*, and average linkage disequilibrium (Kelly, 1997) is >2-fold higher.  
241 Overall, these findings show that, relative to *D. simulans*, *D. mauritiana* has lower



**Figure 5.** Population genomic scans for polymorphism, divergence, and introgression in 10-kb windows. The rows of panels show: nucleotide diversity for a sample of 10 inbred strains of *D. mauritiana* ( $\pi_{\text{mau}}$ , green dots) and 20 inbred strains of *D. simulans* ( $\pi_{\text{sim}}$ , purple dots); nucleotide divergence scaled by within-species polymorphism (blue dots); and  $G_{\min}$  (red dots), the ratio of the minimum number of nucleotide differences per site between *D. mauritiana* and *D. simulans* to the average number of differences per site, a summary statistic that is sensitive to introgression. Panels correspond to each major chromosome arm, with genome coordinates on the x-axis.

242 nucleotide diversity; retained fewer ancestral SNPs; accumulated fewer derived SNPs; a  
243 less negatively skewed site frequency spectrum; and greater linkage disequilibrium—all  
244 patterns consistent with a historically smaller effective population size in *D. mauritiana*  
245 than in *D. simulans*.

246 **Divergence and differentiation:** Net divergence levels between species are comparable  
247 to diversity levels within species. The median number of pairwise differences per site  
248 ( $D_{XY}$ ) between the two species, estimated in non-overlapping 10-kb windows, is 0.010 for  
249 the X chromosome and 0.013 for the autosomes. However, as the X chromosome has  
250 lower levels of polymorphism within species, the median net divergence ( $D_A$ ) between  
251 species is 0.0007 for the X and -0.0005 for the autosomes (a negative value of  $D_A$  on the  
252 autosomes occurs because, on average, levels of within-species polymorphism exceed  
253 levels of between-species divergence). Allele frequency differentiation is also higher for  
254 the X chromosome (median  $F_{ST}=0.378$ ) than the autosomes (median  $F_{ST}=0.279$ ,  
255  $P_{MWU}<2.2e-16$ ). These estimates imply that, for X-linked and autosomal loci, the mean  
256 within-species sample coalescence times are 1.65- and 2.58-fold deeper than the species  
257 divergence time, respectively (Slatkin, 1993). The genetic basis of hybrid  
258 incompatibilities is therefore nested within the much deeper collective genealogical  
259 history of most genetic variation in *D. simulans* and *D. mauritiana*.

260 **Recent interspecific gene flow and introgression:** Gene flow between *D. mauritiana*  
261 and *D. simulans* has been rare during their speciation history, with an apparent recent  
262 increase (Garrigan et al., 2012). To identify genomic regions that have introgressed  
263 between species in the very recent past, we used the  $G_{min}$  statistic—the ratio of the  
264 minimum pairwise sequence distance between species to the average pairwise distance  
265 between species ( $\min[D_{XY}] / \bar{D}_{XY}$ ; (Geneva et al., 2015)). As populations diverge without  
266 gene flow, all loci in the genome gradually approach reciprocal monophyly, leaving just  
267 one ancestral lineage from each population available for coalescence in the ancestral  
268 population. Consequently, the minimum distance (numerator) equals the mean pairwise  
269 distance (denominator), causing  $G_{min} \rightarrow 1$  with zero variance. Conversely,  $G_{min}$  is small  
270 when the minimum distance is small relative to the mean pairwise distance.  $G_{min}$  is  
271 therefore sensitive to genealogical distortions resulting from recent gene flow,

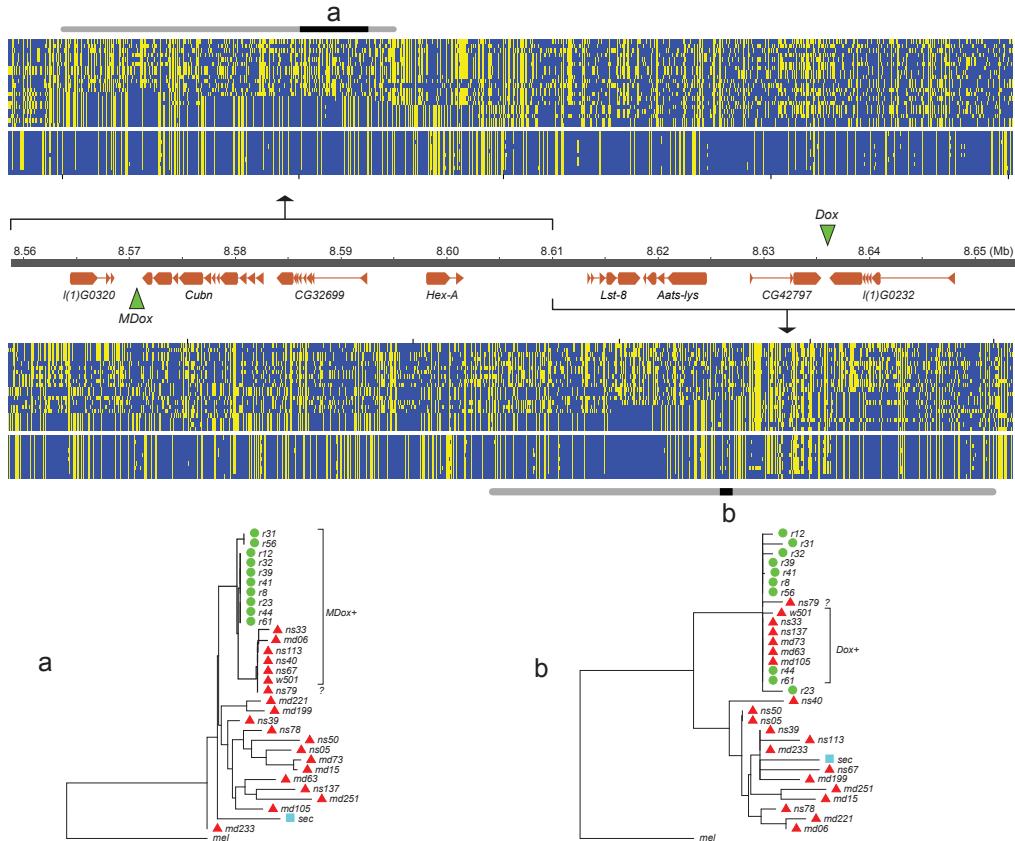
272 particularly when the introgressed sequence is at low to intermediate population  
273 frequency (Geneva et al., 2015). Between *D. mauritiana* and *D. simulans*, median  $G_{\min}$   
274 ( $\pm$  median absolute deviation) estimated for 10-kb windows across the major  
275 chromosome arms ranges from  $0.761 \pm 0.0537$  for  $3L$  to  $0.785 \pm 0.0531$  for the  $X$  (**Figure 5**;  
276 Kruskal-Wallis test,  $P < 2.2e-16$ ).

277 To identify 10-kb outlier windows that have genealogical histories inconsistent with strict  
278 allopatric divergence, we used a Monte Carlo simulation procedure that assumes a  
279 constant species divergence time across all 10-kb intervals, separately for the  $X$  and the  
280 autosomes (see **Methods**). In total, 196 of the 10,443 10-kb windows (1.9%) have a  
281 more recent common ancestry between *D. mauritiana* and *D. simulans* than expected  
282 under a strict allopatric divergence model, as indicated by significantly low values of  
283  $G_{\min}$  ( $P \leq 0.001$ ). As  $G_{\min}$  is a ratio, significantly small  $G_{\min}$  values could result from  
284 unusually small numerators (minimum  $D_{XY}$ ) or unusually large denominators ( $\bar{D}_{XY}$ ). We  
285 find that 10-kb windows with significant  $G_{\min}$  values have smaller median minimum  $D_{XY}$   
286 (0.0056 in introgression windows *versus* 0.0094 genome-wide,  $P_{MWU} < 0.0001$ ) as well  
287 as *smaller* median  $\bar{D}_{XY}$  (0.0110 in introgression windows *versus* 0.0124 genome-wide  
288  $P_{MWU} < 0.0001$ ), indicating that the significant  $G_{\min}$  values are due to unusually small  
289 minimum  $D_{XY}$  values. Introgression windows are 4.4-fold underrepresented on the  $X$   
290 chromosome: only nine of 1833 10-kb windows on the  $X$  chromosome (0.49%) have  
291 significant  $G_{\min}$  values *versus* 187 of 8414 10-kb windows on the autosomes (2.17%;  
292 Fisher's exact test  $P = 6.5e-08$ ). However, not all 10-kb introgression windows are  
293 independent: 169 of the 196 significant 10-kb windows (86.2%) can be arrayed into  
294 contiguous (or nearly contiguous) genomic regions (see **Methods**). As a result, we infer  
295 27 small (10-kb) introgressions and 21 larger introgressions ranging in size from 20 kb to  
296 280 kb (**Table S2**). Of these 48 total introgressions, only one is on the  $X$  chromosome  
297 and 47 are on autosomes ( $\chi^2$ -test,  $P = 0.0124$ ). The lengths of these introgressed  
298 haplotypes have been eroded by recombination over time in the receiving population:  
299 longer, presumably younger, introgressions have smaller minimum  $D_{xy}$  values (Spearman  
300  $\rho = -0.526$ ,  $P = 0.00012$ ). However, variation in local recombination rate has not been an  
301 important factor affecting introgression lengths (Spearman  $\rho = -0.0412$ ,  $P = 0.781$ ).

302 **Interspecific introgression of the cryptic Winters *sex ratio* drive system:** The single  
303 introgression detected on the X chromosome corresponds to a 130-kb region that  
304 comprises eight protein-coding genes plus the Winters *sex ratio* meiotic drive genes,  
305 *Distorter on the X* (*Dox*) and, its progenitor gene, *Mother of Dox* (*MDox*) (Tao et al.,  
306 2007a) (**Figure 6**). The median  $G_{\min}$  value across this 130-kb region is 0.333, a ~2.4-fold  
307 reduction relative to background  $G_{\min}$  on the X chromosome ( $P_{MWU} < 0.0001$ ). In *D.*  
308 *simulans*, when unsuppressed, *MDox* and *Dox* cause biased transmission of the X  
309 chromosome, with male carriers siring excess daughters (Tao et al., 2007b). These  
310 drivers are suppressed by an autosomal gene, *Not much yin* (*Nmy*), which is a  
311 retrotransposed copy of *Dox*, via a putative RNA-interference mechanism (Tao et al.,  
312 2007b). In non-African *D. simulans* populations, *Dox*, *MDox*, and *Nmy* are nearly fixed,  
313 although haplotypes lacking the genes segregate at low frequencies (Kingan et al., 2010).  
314 All three loci have histories consistent with selective sweeps due to the presumed  
315 transmission advantage at *MDox* and *Dox* and the associated selective advantage of  
316 suppressing drive and restoring equal sex ratios at *Nmy* (Kingan et al., 2010).

317 Maximum-likelihood phylogenetic trees for the 130-kb *MDox-Dox* region show reduced  
318 diversity within *D. mauritiana* and reduced divergence between species (**Figure 6**).  
319 Among the 10 *D. mauritiana* sequences, nucleotide diversity is just 24% ( $\pi=0.0018$ ) of  
320 background diversity levels on the X chromosome, corresponding to a massive selective  
321 sweep in the *D. mauritiana* genome ( $P_{MWU} < 0.0001$ ; see also (Nolte et al., 2013, Garrigan  
322 et al., 2014)). The distribution of variability among haplotypes in the *D. simulans*  
323 samples is consistent with a parallel, albeit incomplete, selective sweep (**Figure 6**). The  
324 most extreme 10-kb window within the 130-kb region has a minimum  $D_{XY}$  value  
325 ( $=0.00087$ ) that is 92% smaller than the X chromosome-wide  $\bar{D}_{XY}$ , implying that  
326 introgression occurred in the very recent past.

327 To determine if the *MDox* and/or *Dox* drive elements are associated with introgression  
328 between species and the selective sweeps within each species, we determined *MDox* and  
329 *Dox* presence/absence status for each line using diagnostic restriction digests (see  
330 **Methods**). Previous work shows that *MDox* and *Dox* are nearly fixed among *D. simulans*  
331 samples collected outside of Africa (Kingan et al., 2010). However, among our 19



**Figure 6.** Natural introgression of the *MDox*-*Dox* region of the X chromosome. The top half of the figure shows two DNA polymorphism tables: the top table corresponds to the *MDox* region, and the bottom corresponds to the *Dox* region. Within the tables, yellow squares denote the derived nucleotide state, and blue squares indicate the ancestral state. The top 20 rows of each table correspond to the *D. simulans* samples, and the bottom 10 rows correspond to the *D. mauritiana* samples. The genome map between the polymorphism tables shows gene models for the region (orange boxes) and the locations of the *MDox* and *Dox* genes (green triangles). The grey bars on the top and bottom of the polymorphism tables mark the sites that occur within an introgressed region, with black sections labeled “a” and “b”, which mark the location of the *MDox* and *Dox* genes, respectively. The bottom panels of the figure show two maximum likelihood phylogenetic trees labeled “a” and “b”, corresponding to the *MDox* and *Dox* regions.

332 African samples (9 Madagascar, 10 Kenya), we find that the drivers are at lower  
333 frequency: five have *MDox* (26%), five have *Dox* (26%), and only one has both genes  
334 (5%; NS33; **Table S2**). Despite these low frequencies, *MDox* and *Dox* are  
335 overrepresented among the haplotypes shared between species: 6 of the 7 shared  
336 haplotypes have *MDox* and/or *Dox* (Fisher's Exact  $P_{\text{FET}}=0.0018$ ), and 2 of the 7 possess  
337 both drivers ( $P_{\text{FET}}=0.0158$ ;  $n=19$  African samples, plus the reference strain, *D. simulans*  
338  $w^{501}$ , which has both). In *D. mauritiana*, all 10 lines have *MDox*, but only two have *Dox*  
339 (**Figure 6; Table S6**). RT-PCR confirms that *MDox* is expressed in testes from both  
340 species (see **Methods**). These findings provide support for the notion that *Dox* and  
341 (transcriptionally active) *MDox* genes mediated the introgression and parallel sweeps.

342 The large *MDox-Dox* introgression, and its associated sweep co-localize with one of the  
343 three regions of the X chromosome that, in our mapping experiments, fails to cause  
344 sterility when introgressed from *D. mauritiana* into *D. simulans* (**Figure 2**). These  
345 observations suggest that a driving haplotype moved between species and swept to high  
346 frequency in *D. simulans* or fixation in *D. mauritiana*, thereby reducing local sequence  
347 divergence. This discovery has two implications. First, the *MDox-Dox* region is the only  
348 locus on the X chromosome to have recently escaped from its linked hybrid  
349 incompatibility factors and introgressed between species. Second, by sweeping to high  
350 frequency or fixation, the *MDox-Dox* drive element region reduced local divergence  
351 between species and, incidentally, undermined the accumulation of genetic  
352 incompatibilities that might cause hybrid male sterility.

353

## 354 **DISCUSSION**

355 Our combined genetic and population genomics analysis of speciation between *D.*  
356 *mauritiana* and *D. simulans* yields three findings. First, we confirm the rapid  
357 accumulation of X-linked hybrid male sterility between these species and map four major  
358 sterility factors to small (<700 kb) intervals (**Figure 2**). Second, we find that very recent  
359 natural introgression has occurred between these species, albeit almost exclusively on the  
360 autosomes, consistent with a large X-effect on gene flow (**Table S.1**). Third, we discover  
361 new roles for meiotic drive during the history of speciation between these species. Some

362 drive seems to have contributed to functional divergence between species: one region of  
363 the *D. mauritiana* X chromosome appears to cause segregation distortion in a *D.*  
364 *simulans* genetic background. In contrast, the well-characterized X-linked Winters *sex*  
365 *ratio* distorters, *MDox* and *Dox*, have clearly migrated between species, reducing local  
366 interspecific divergence. Together, these findings suggest that genetic conflict may both  
367 promote as well as undermine the special role of sex chromosomes in speciation.

368 ***Genetic basis of X-linked hybrid male sterility***

369 Our genetic analyses were initiated by introgression of six different regions of the *D.*  
370 *mauritiana* X chromosome into a pure *D. simulans* genetic background. All six regions  
371 cause complete hybrid male sterility and therefore carry at least one, or a combination of,  
372 *D. mauritiana* allele(s) that disrupts spermatogenesis due to incompatibilities with X-  
373 linked, Y-linked, or autosomal *D. simulans* alleles. Only three large (>2Mb) regions of  
374 the *D. mauritiana* X are readily exchangeable between species, permitting male fertility  
375 in a *D. simulans* genome. Thus, after only ~250,000 years, sufficient X-linked hybrid  
376 male sterility has accumulated to render most of the *D. mauritiana* X chromosome male-  
377 sterile on a *D. simulans* genetic background (True et al., 1996b). Most of the *D.*  
378 *mauritiana* X chromosome is male-sterile in a *D. sechellia* genome as well (Masly and  
379 Presgraves, 2007).

380 We were able to define four small regions (<700kb), each sufficient to cause complete  
381 male sterility (**Figure 2**), suggesting that these may contain single, strong sterility factors.  
382 We also find a large region spanning most of 2P-6 from which we were unable to recover  
383 fertile 1P-YFP recombinants. We infer that 2P-6 contains a minimum of two strong  
384 sterility regions, one tightly linked to each of the flanking *P*-elements (**Figure 3**). While  
385 our 2P mapping scheme is designed to facilitate the identification of male sterility factors,  
386 the 2P-6 interval highlights one of its limitations: in regions like 2P-6, for which strong  
387 sterility factors are very close to both flanking *P*-elements, we cannot determine how  
388 many additional sterility factors might localize to the middle of the interval. The present  
389 experiments therefore provide only a minimum estimate of the total hybrid male sterility  
390 factors on the X chromosome. We tentatively conclude that, within the fraction of the *D.*  
391 *mauritiana* X chromosome investigated, there are at least six genetically separable

392 regions, each individually sufficient to cause virtually complete male sterility. It is worth  
393 noting these experimental approaches detect relatively large-effect sterility factors under  
394 a single set of laboratory conditions. There are likely many hybrid male sterility factors  
395 of smaller effect, generally neglected in the lab but easily detected by selection in natural  
396 populations and thus able to affect the probability of migration at linked loci.

397

### 398 ***Genomic signatures of complex speciation with gene flow***

399 The two species studied here are allopatric: *D. simulans* has never been reported on  
400 Mauritius, and *D. mauritiana* has never been found anywhere other than Mauritius  
401 (David et al., 1989, Legrand et al., 2011). *D. mauritiana* appears to have originated from  
402 a *D. simulans*-like ancestor, probably from Madagascar, that migrated and established a  
403 population on Mauritius (Hey and Kliman, 1993, Kliman et al., 2000). Our  
404 characterization of genome-wide variation within and between *D. mauritiana* and *D.*  
405 *simulans* confirms a coalescent history that reaches considerably deeper into the past than  
406 the inferred species split time of ~250,000 years (Hey and Kliman, 1993, Kliman et al.,  
407 2000). Nested within this largely shared coalescent history, many functional differences  
408 have evolved between the two species, including extreme ones that mediate large-effect  
409 hybrid incompatibilities. The signatures of gene flow found in the genomes of these  
410 species imply recurrent bouts of migration and interbreeding. To introgress between  
411 species, immigrating foreign haplotypes must escape their locally disfavored  
412 chromosomal backgrounds by recombination before being eliminated by selection against  
413 linked incompatibilities and locally maladaptive alleles (Petry, 1983, Bengtsson, 1985,  
414 Barton and Bengtsson, 1986). Conditional on escape, the lengths of foreign haplotypes  
415 will be subject to gradual erosion by recombination with the resident genetic background.

416 Here, and in previous work (Garrigan et al., 2012), we detect evidence consistent with  
417 weak migration: 2-5% of the genome shows evidence of introgression between *D.*  
418 *simulans* and *D. mauritiana* during their recent history. Our population genomic analysis  
419 of polymorphism and divergence using the  $G_{\min}$  statistic identified 48 segregating foreign  
420 haplotypes. We find evidence that the genomic locations and lengths of introgressed  
421 foreign haplotypes have been shaped by selection and recombination in the receiving

422 population, respectively. First, selection has likely affected the genomic distribution of  
423 foreign haplotypes: only one of the 48 introgressions occurs on the X chromosome. The  
424 opportunity for foreign haplotypes to escape the X chromosome via recombination is  
425 more constrained than on the autosomes, as the X has a higher density of incompatible  
426 alleles, and hemizygous selection eliminates foreign X-linked haplotypes more quickly  
427 (Muirhead and Presgraves, 2016). Second, after escaping locally deleterious  
428 chromosomal backgrounds, recombination has eroded the lengths of foreign haplotypes  
429 over time: recently introgressed, and hence less diverged, haplotypes tend to be longer. It  
430 is worth noting here that the 10-kb windows used for our  $G_{\min}$  scan for foreign haplotypes  
431 almost certainly fails to identify smaller and/or older introgressions.

432 ***Meiotic drive and complex speciation***

433 The drive theory posits that hybrid incompatibilities accumulate as incidental by-products  
434 of recurrent bouts of meiotic drive and suppression (Hurst and Pomiankowski, 1991,  
435 Frank, 1991). Our mapping experiments provide no direct evidence in support of this  
436 theory in *D. mauritiana* and *D. simulans*, as no hybrid male sterility loci co-localized with  
437 sex-ratio loci. Our genetic mapping experiments have however provided further evidence  
438 for the accumulation of cryptic *sex-ratio* drive systems. We mapped a small region of the  
439 *D. mauritiana* X that, when introgressed into a naïve *D. simulans* genetic background,  
440 causes weak segregation distortion resulting in female-biased progeny sex ratios (**Figure**  
441 **4**). As the *D. mauritiana* X-drive locus does not map to the location of any of the three  
442 cryptic drivers known from *D. simulans*, we infer that it may be a new, previously  
443 undiscovered drive system in *D. mauritiana*.

444 Across *D. simulans* and *D. mauritiana*, four cryptic drive systems have been identified so  
445 far: two X-drive systems in *D. simulans* (Paris and Durham); one X-drive system in *D.*  
446 *mauritiana* (see above); and one X-drive system found in both species (Winters; see  
447 below). We regard this as a minimum for several reasons. First, weak segregation  
448 distortion that may be significant in natural populations can go undetected in laboratory  
449 experiments. Second, cryptic drive systems may not be fixed within species, and we have  
450 only surveyed genotypes derived from one strain each of *D. mauritiana* and *D. simulans*.  
451 Third, no study has yet comprehensively assayed *D. simulans* material introgressed into a

452 *D. mauritiana* genetic background. Finally, some cryptic drive systems might go to  
453 fixation and then simply die because once fixed (or suppressed), a driver is in a race:  
454 either suffer mutational decay or acquire a mutation that confers a new bout of drive.  
455 These considerations—and the discovery of multiple alternative cryptic drive systems in  
456 closely related species—imply that sex chromosome drive is not infrequent during the  
457 history of species divergence (Jaenike, 2001).

458 We have found that the Winters sex-ratio drivers, *MDox* and *Dox*, have moved between  
459 these two species. This discovery highlights an implicit assumption of the drive theory of  
460 the large X-effect—namely that species evolve in strict allopatry. With gene flow, drive  
461 elements (and other selfish genes) have the opportunity to jump species boundaries and  
462 undermine divergence in a process analogous to adaptive introgression (Seehausen et al.,  
463 2014, Crespi and Nosil, 2013). The *t*-haplotype has, for instance, introgressed between  
464 sub-species of house mouse, *Mus musculus* (Macaya-Sanz et al., 2011). Between *D.*  
465 *mauritiana* and *D. simulans*, the  $G_{\min}$  statistic and the genealogies associated with the  
466 *MDox-Dox* introgressed haplotype (**Figure 6**) are agnostic on the direction of  
467 introgression. Nonetheless, the finding that a drive element crossed a species boundary  
468 has important implications for the drive theory explanation of Haldane’s rule and the  
469 large X-effect. For *MDox* and *Dox* to introgress between species, three things must be  
470 true: (1) neither *MDox* nor *Dox* alleles from the donor species caused male sterility in the  
471 recipient species; (2) no X-linked hybrid male sterility factors are so tightly linked to  
472 *MDox* and *Dox* as to prevent their eventual escape by recombination into the recipient  
473 species genetic background; and (3) any sterility factors located within the introgressed  
474 region of the recipient X will have been replaced by foreign alleles. Together, these  
475 inferences suggest that a selfish drive system was able to invade a new species by *not*  
476 causing male sterility and, for one X-linked region, may have impeded or undone the  
477 evolution of hybrid male sterility.

478

## 479 METHODS

480 **Drosophila husbandry and genetics:** All *Drosophila* crosses and phenotyping were  
481 done in parallel in two locations, using standard cornmeal media (Rochester, NY) or

482 minimal cornmeal media (Bloomington, IN) at room temperature (23-25C). We  
483 constructed *D. mauritiana* “2P” lines that carry pairs of X-linked *P*-element insertions  
484 that contain the mini-white transgene (*P[w<sup>+</sup>]*) (True et al., 1996a) which serve as semi-  
485 dominant visible genetic eye-color markers and allow us to distinguish individuals  
486 carrying 0, 1 or 2 *P[w<sup>+</sup>]*. These “2P” regions were then introgressed into the *D. simulans*  
487 *w<sup>XD1</sup>* genetic background through more than 40 generations of repeated backcrossing  
488 while following the two *P[w<sup>+</sup>]* insertions (**Figure 1A**). Each 2P introgression line was  
489 then bottlenecked through a single female to eliminate segregating variation in the  
490 recombination breakpoints flanking the 2P[*w<sup>+</sup>*] interval.

491 We performed meiotic mapping to ascertain the genetic basis of male sterility within each  
492 2P introgression by generating recombinant 1P introgression genotypes (**Figure 1B**).  
493 2P[*w<sup>+</sup>*] females were crossed to *D. simulans* strains carrying an X-linked *pBac[eYFP]*  
494 transgene (Stern et al., 2017) that served as an additional visible marker. Progeny from  
495 this cross were scored for recombinant X chromosomes carrying both *pBac[eYFP]* and a  
496 single *P[w<sup>+</sup>]* (1P-YFP). Recombinant 1P-YFP chromosomes were generated using  
497 *pBac[eYFP]* markers both proximal and distal to each 2P introgression. Virgin 1P-YFP  
498 females were individually crossed to *D. simulans* *w<sup>XD1</sup>* males to initiate 1P-YFP strains.  
499 Each 1P-YFP X chromosome was then assayed for male fertility. At least 10 individual  
500 1P-YFP males of each genotype were collected 1-2 days post-eclosion and aged 3-5 days,  
501 then placed singly in a vial with three virgin *D. simulans* *w<sup>XD1</sup>* females. After seven days,  
502 both the male and females were discarded, and all offspring emerging from the vial were  
503 counted. Additional 1P-YFP males were archived for DNA extraction.

504 Progeny sex ratios were calculated as the number of female offspring/total number of  
505 offspring (% female). Males that sired fewer than five offspring were excluded from sex  
506 ratio analyses, as were genotypes with fewer than three males that sired more than four  
507 offspring. This resulted in 2538 males and 303 recombinant 1P-YFP chromosomes that  
508 were used to estimate progeny sex ratios; 210 recombinant 1P-YFP genotypes had both  
509 progeny sex ratio and sequence data.

510 **Genotyping recombinant chromosomes by sequencing:** We determined the fine-scale  
511 genetic architecture of hybrid male sterility within each introgressed region by

512 genotyping recombinant 1P-YFP X chromosomes using multiplexed whole-genome  
513 sequencing. DNA extraction and library construction followed published methods for  
514 high-throughput sequence analysis of a large number of recombinant genotypes  
515 (Andolfatto et al., 2011, Peluffo et al., 2015). Sequence reads were mapped to the  
516 reference genome sequence of the *D. mauritiana* stock used for mapping (*mau w*<sup>12</sup>)  
517 (Garrigan et al., 2012), our unpublished genome sequence of *sim w*<sup>XD1</sup>, and the *D.*  
518 *simulans pBac[eYFP]* strains (Stern et al., 2017). Ancestry from each parent species was  
519 determined by a Hidden Markov Model (HMM) (Andolfatto et al., 2011).

520 Across the 439 genotypes with sufficiently high-quality sequence data for ancestry  
521 assignment, we recovered 64,373 X-linked markers. A subset of 2,835 non-redundant  
522 markers were retained that delimit the extent of each 1P-YFP *D. mauritiana* segment. No  
523 genotype showed evidence for any autosomal *D. mauritiana* alleles (see **Figure S2** for  
524 exemplars), confirming that our introgression scheme isolated X-linked *D. mauritiana*  
525 segments in a pure *D. simulans* autosomal genome.

526 **Samples and short read alignment:** We used genome sequence data from 10 lines of *D.*  
527 *mauritiana*, including nine inbred wild isolates and the genome reference strain, *mau w*<sup>12</sup>  
528 (14021-0241.60); 20 lines of *D. simulans*, including 10 inbred wild isolates from Kenya  
529 (14021-0251.302-311), 9 wild isolates from Madagascar (14021-0251.293-301), and the  
530 reference strain, *sim w*<sup>501</sup> (14021-0251.011); and the reference strain of *D. melanogaster*.  
531 The *D. mauritiana* and *D. simulans* sequence data were reported previously (Garrigan et  
532 al., 2012, Garrigan et al., 2014, Rogers et al., 2014). We performed short read alignment  
533 against the *D. mauritiana* genome assembly (version 2) using the “aln/sampe” functions  
534 of the BWA short read aligner and default settings (Li and Durbin, 2009). Reads  
535 flanking indels were realigned using the SAMTOOLS software (Li et al., 2009).  
536 Individual BAM files were merged and sorted with SAMTOOLS.

537 **Polymorphism and divergence analyses:** Both within- and between-population  
538 summary statistics were estimated in 10-kb windows using the software package  
539 POPBAM (Garrigan, 2013). The within population summary statistics include: unbiased  
540 nucleotide diversity  $\pi$  (Nei, 1987); the summary of the folded site frequency spectrum  
541 Tajima’s  $D$  (Tajima, 1989); and the unweighted average pairwise value of the  $r^2$  measure

542 of linkage disequilibrium,  $Z_{nS}$ , excluding singletons (Kelly, 1997). The between  
543 population summary statistics include: two measures of nucleotide divergence between  
544 populations,  $D_{XY}$ , and net divergence,  $D_A$  (Nei, 1987); the ratio of the minimum between-  
545 population nucleotide distance to the average,  $G_{\min}$  (Geneva et al., 2015); and the fixation  
546 index,  $F_{ST}$  (Wright, 1951). In the analysis of absolute numbers of ancestral *versus*  
547 derived polymorphisms, we restricted the *D. simulans* sample to the 10 from Madagascar  
548 to enforce equal sample sizes with *D. mauritiana* ( $n=10$ ). From a total of 11083 scanned  
549 10-kb windows, we only analyzed windows for which at least 50% of aligned sites  
550 passed the default quality filters in POPBAM, which resulted in a final alignment for  
551 10443 scanned 10-kb windows. POPBAM output was formatted for use in the R  
552 statistical computing environment using the package, POPBAMTools  
553 (<https://github.com/geneva/POPBAMTools>). All statistics and data visualization were  
554 done in R (Team, 2013).

555 **Identification of introgressed regions:** We used the  $G_{\min}$  statistic (Geneva et al., 2015)  
556 to scan the genome for haplotypes that have recent common ancestry between *D.*  
557 *simulans* and *D. mauritiana*.  $G_{\min}$  is defined as the ratio of the minimum number of  
558 nucleotide differences per aligned site between sequences from different populations to  
559 the average number of nucleotide differences per aligned site between populations. The  
560  $G_{\min}$  statistic was calculated in 10-kb intervals across each major chromosome arm using  
561 the same quality filtering criteria used for all other summary statistics. From these values,  
562 we estimated the probability of the observed  $G_{\min}$  under a model of allopatric divergence,  
563 conditioned on the divergence time. For each 10-kb interval, the significance of the  
564 observed  $G_{\min}$  value was tested via Monte Carlo coalescent simulations of two  
565 populations diverging in allopatry with all mutations assumed to be neutral. We assumed  
566 a population divergence time of  $1.21 \times 2N_{\text{sim}}$  generations before the present, in which  
567  $N_{\text{sim}}$  is the current estimated effective population size of *D. simulans* (Garrigan et al.,  
568 2012). In the simulations, the observed local value of  $D_{XY}$  was used to determine the  
569 neutral population mutation rate for that 10-kb interval. To account for uncertainty in  
570 local population recombination rate, for each simulated replicate, a rate was drawn from a  
571 normally distributed prior (truncated at zero) with the mean estimated from genetically

572 determined crossover frequencies (True et al., 1996a). The empirical crossover rate  
573 estimates were converted from cM to  $\rho$  (the population crossover rate,  $4N_{\text{sim}}c$ , by  
574 assuming  $N_{\text{sim}} \approx 10^6$ ). The effective population sizes of both species were assumed to be  
575 equal and constant. While the assumption of constant size is not realistic, it serves to  
576 make the test more conservative as any factor that decreases the effective population size  
577 also increases the rate of coalescence within populations, which has the net effect of  
578 increasing  $G_{\min}$  and decreasing its variance, since fewer ancestral lineages will be  
579 available for inter-specific coalescence in the ancestral population (Geneva et al., 2015).  
580 For each 10-kb interval,  $10^5$  simulated replicates were generated and the probability of  
581 the observed  $G_{\min}$  value was estimated from the simulated cumulative density. To  
582 identify putatively introgressed haplotypes, we used a significance threshold of  $P \leq 0.001$   
583 from the simulations, which yields a proportion of null tests of 0.982 and a false  
584 discovery rate of 5%. To infer the full length of any putative introgressions  $\geq 10$  kb, we  
585 identified runs of contiguous (or semi-contiguous) 10-kb windows with significant  $G_{\min}$   
586 values ( $P \leq 0.001$ ). Finally, we estimated maximum likelihood phylogenies for each of  
587 the putative introgression intervals using RAxML v. 8.1.1 (Stamatakis, 2014).

588 **Genotyping the Winters sex ratio genes:** We extracted genomic DNA from single male  
589 flies using the Qiagen DNeasy Blood and Tissue Kit. The meiotic drive genes of the  
590 Winters *sex ratio* system (Tao et al., 2007a), *Dox* and *MDox*, were PCR-amplified as  
591 previously described (Kingan et al., 2010). To assay the presence or absence of the *Dox*  
592 and *MDox* gene insertions, the amplicons for the *Dox* and *MDox* regions were digested  
593 with the *StyI* and *StuI* restriction enzymes (NEB), respectively. The digests were run on a  
594 1% agarose gel stained with EtBr and the band size was estimated using the GeneRuler 1  
595 kb plus ladder (Thermo Scientific). For both genes, only haplotypes containing the gene  
596 insertions have restriction sites as confirmed by samples with known genotypes (Kingan  
597 et al., 2010).

598 **Quantitative PCR for *Dox/MDox* expression in fly testes:** We assayed expression of  
599 the *Dox* and *MDox* genes in testes from *D. simulans* strain MD63 and *D. mauritiana*  
600 strain *mau w*<sup>12</sup> using quantitative PCR. Total RNA was extracted from the dissected  
601 testes of 5-10 day old flies using the Nucleospin RNA XS kit (Macherey-Nagel,

602 Germany), and cDNA was synthesized with poly dT oligos and random hexamers using  
603 Superscript III RT cDNA synthesis kit (Invitrogen, CA). qPCR assays were performed  
604 on a BioRad Real-time PCR machine using the cycling conditions: 95° C for 3 mins.; 40  
605 cycles of 95° C for 10s, 58° C for 30s, and 72° C for 30s. The primer sequences used for  
606 qPCR are provided in **Table S1**.

607

608 **ACKNOWLEDGEMENTS**

609 The authors thank Brian Calvi for generously providing space and Shelby Biel, Ally  
610 Shambaugh, and Amanda Meiklejohn for assistance with fertility assays.

611

612 **COMPETING INTERESTS**

613 The authors declare no competing interests.

614

615

616

617

618

619

**Table 1. Locations and lengths of 2P intervals.**

2P interval	left P[w <sup>+</sup> ] <sup>1</sup>	right P[w <sup>+</sup> ] <sup>1</sup>	Length (Mbp)
2P-1	993419	4498520	3.51
2P-3	6192555	9126133	2.93
2P-4	9126133	11189873	2.06
2P-5a	11189873	13324017	2.13
2P-5b	11189873	13903934	2.71
2P-6a	13903934	17492084	3.59
2P-6b	13324017	17492084	4.17
2P-7	17492084	18660037	1.17

<sup>1</sup>coordinate position in the assembled *D. simulans* w<sup>501</sup> genome

**Table 2. Fertility and sex ratio phenotypes for 1P-YFP recombinant genotypes.**

2P interval	n tested	n sterile <sup>0</sup>	n fertile <sup>1</sup>	mean fertility <sup>2</sup>	% fertile <sup>1</sup>	mean SR <sup>2</sup>
2P-1	171	48	108	73.52	0.63	0.43
2P-3	97	12	64	67.38	0.66	0.45
2P-4	77	17	51	71.94	0.66	0.45
2P-5a/b	92	23	53	68.15	0.58	0.51
2P-6a/b	97	69	18	73.83	0.19	0.44
2P-7	83	69	9	133.00	0.11	0.47
all 1P-YFP genotypes	617	238	303	81.30	0.47	0.45

<sup>0</sup>genotypes where no male produced any offspring

<sup>1</sup>genotypes where at least one male produced at least five offspring

<sup>2</sup>among males producing more than four offspring

**Table 3. Distribution of 1P-YFP recombinant introgression lengths.**

2P interval	sequenced	min size	mean size	max size
2P-1	129	295,225	2,617,833	6,322,871
2P-3	73	306,052	1,636,944	3,818,569
2P-4	55	226,018	1,482,659	2,917,578
2P-5	61	365,004	1,627,632	3,276,930
2P-6	55	692,350	2,400,499	4,764,204
2P-7	66	218,722	1,412,108	2,502,552
total	439			

**Table 4. Population genomics summary statistics.**

Inference	Statistic <sup>a</sup>	<i>D. simulans</i>	<i>D. mauritiana</i>	P-value
Polymorphism	median $\pi_X$	0.0119	0.0076	< 0.0001 <sup>c</sup>
	median $\pi_A$	0.0152	0.0116	< 0.0001 <sup>c</sup>
	SNPs with inferred ancestry <sup>b</sup>	4,324,740	2,181,959	< 0.0001 <sup>d</sup>
	% ancestral SNPs	14.6	21.6	< 0.0001 <sup>e</sup>
	% derived SNPs	85.3	78.3	
Site frequency spectra	median Tajima's $D_X$	-1.218	-0.536	< 0.0001 <sup>c</sup>
	median Tajima's $D_A$	-1.127	-0.359	< 0.0001 <sup>c</sup>
Linkage disequilibrium	median $Z_{ns, X}$	0.056	0.122	< 0.0001 <sup>c</sup>
	median $Z_{ns, A}$	0.058	0.129	< 0.0001 <sup>c</sup>

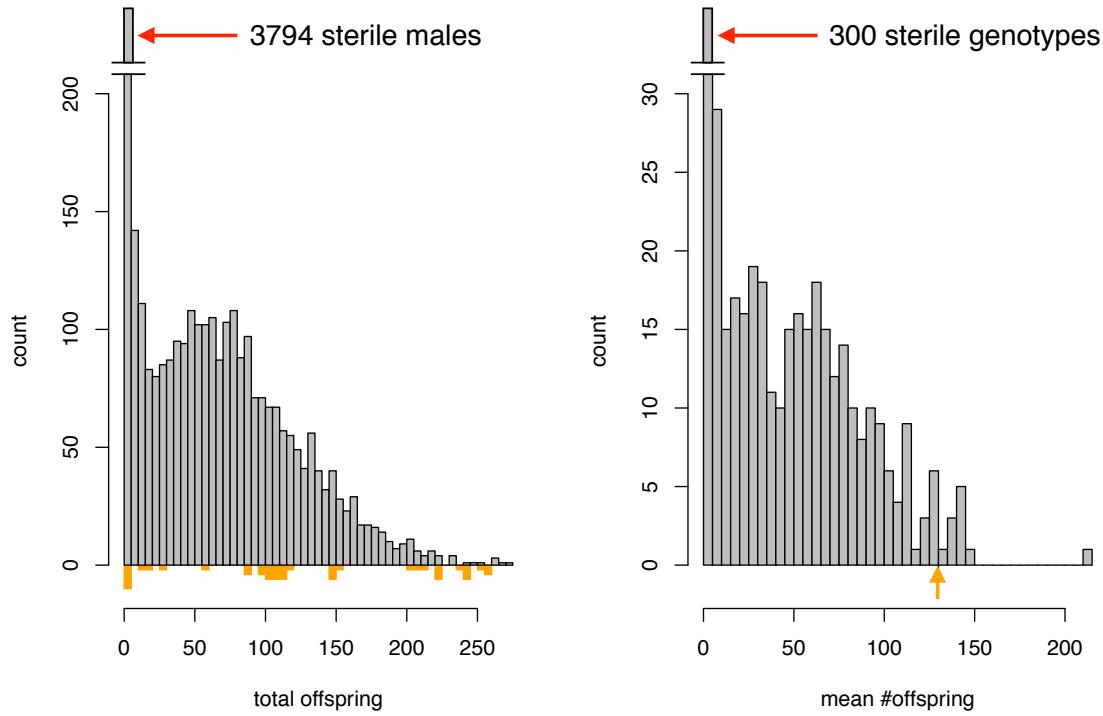
<sup>a</sup> Summary statistics estimated from 10-kb non-overlapping windows.

<sup>b</sup> SNP were inferred as ancestral or derived using parsimony, with *D. melanogaster* as an outgroup (see **Methods**).

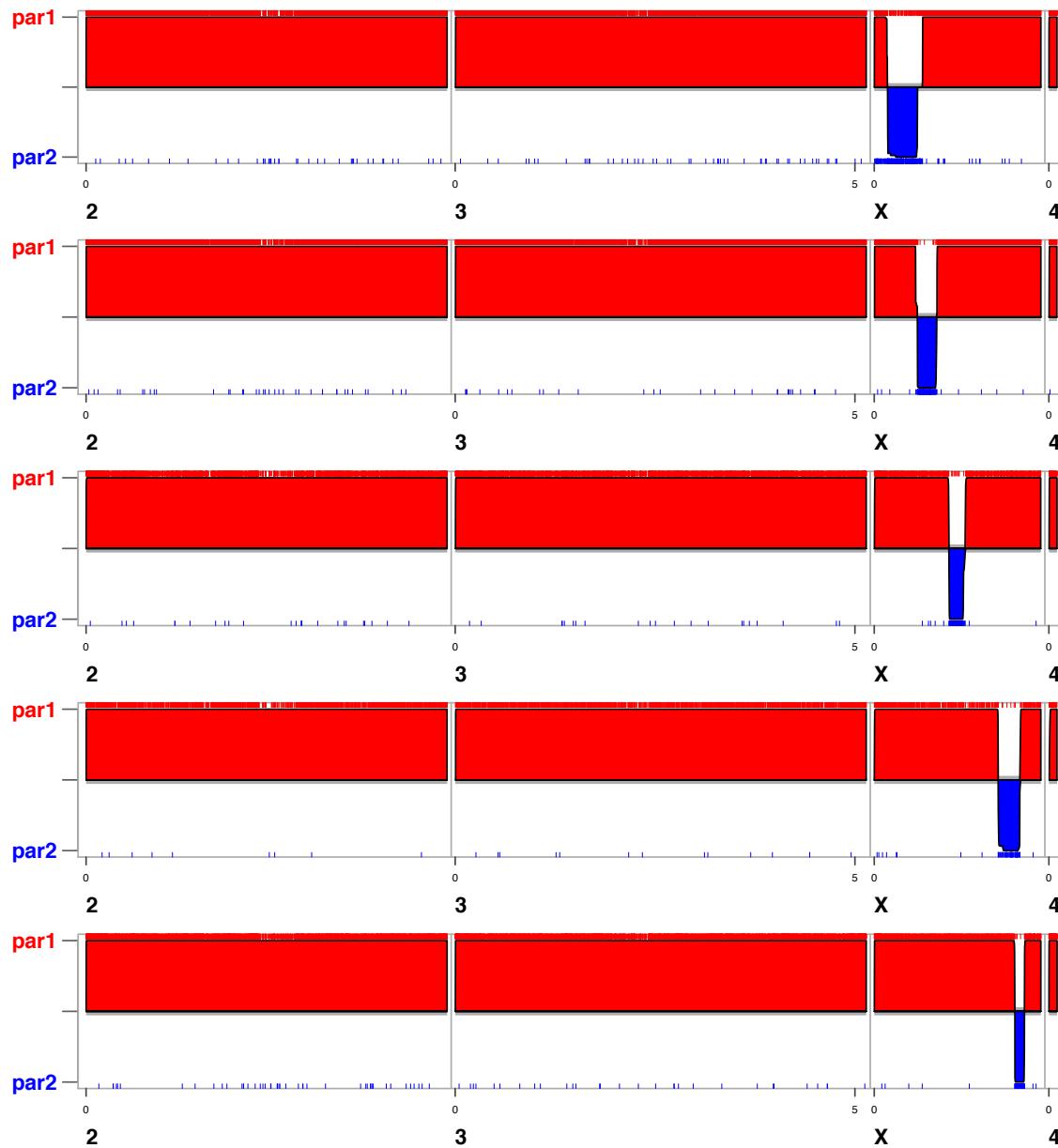
<sup>c</sup> P-value for Mann-Whitney U-test.

<sup>d</sup> P-value for  $\chi^2$ -test.

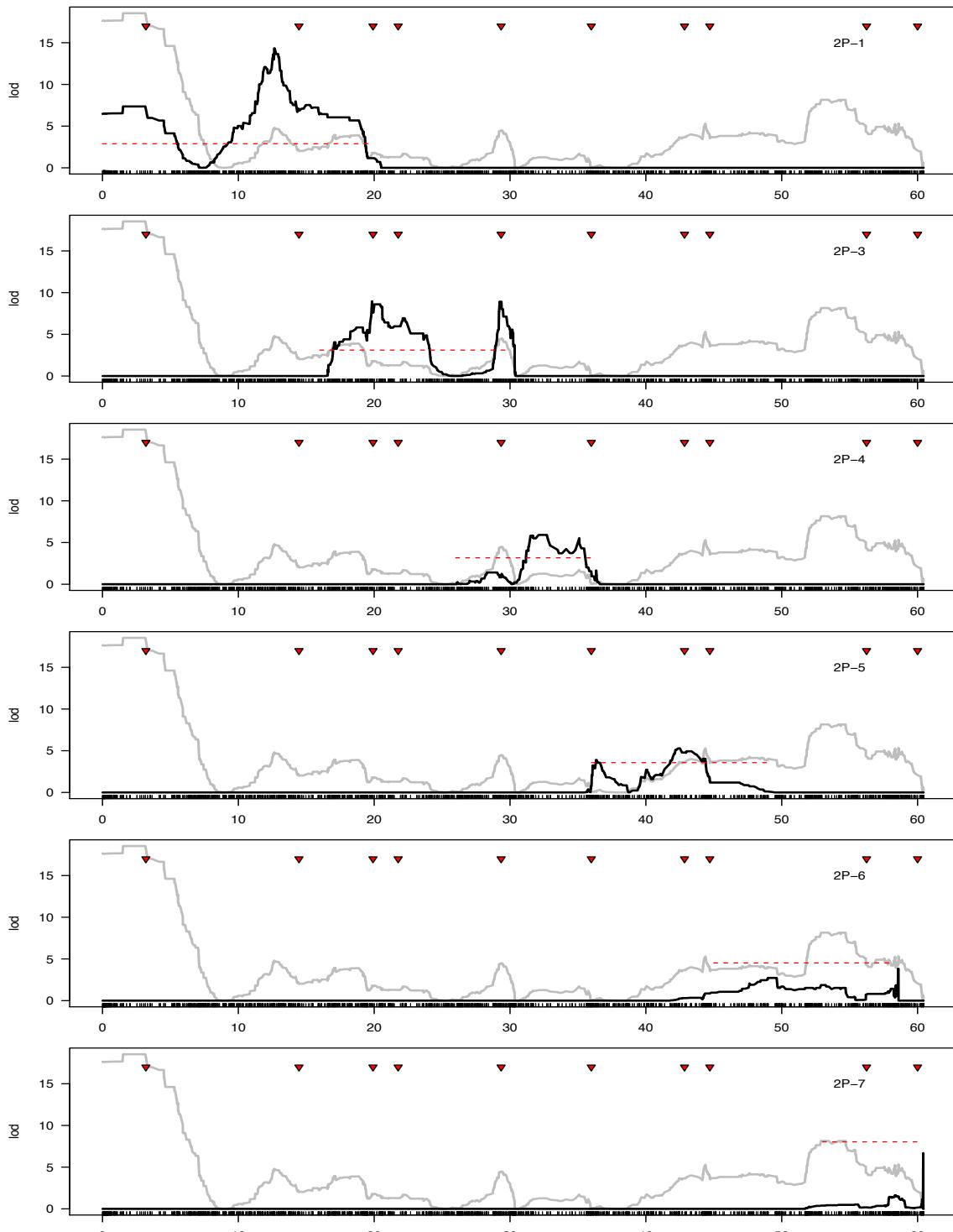
<sup>e</sup> P-value from Fisher's Exact Test.



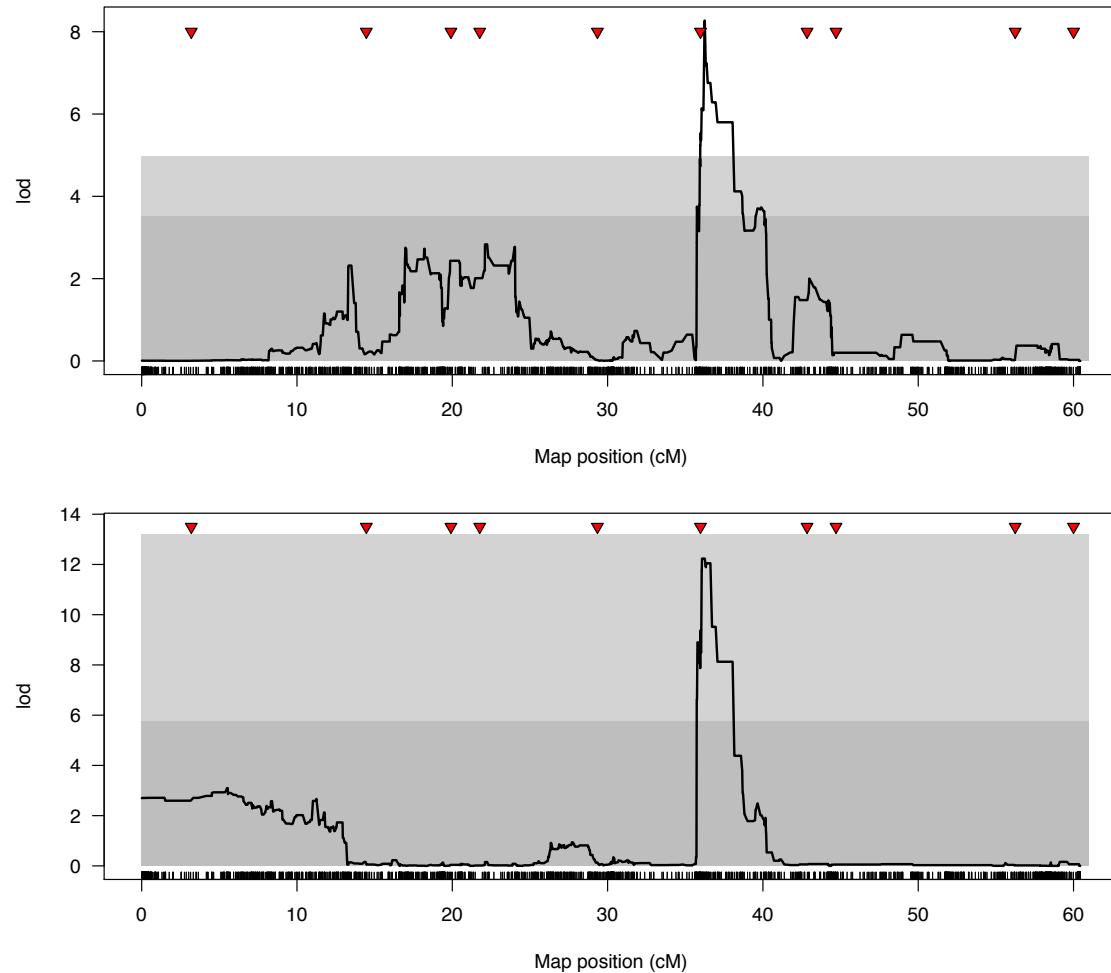
**Figure S1.** Distribution of fertility (number of progeny) among all males carrying recombinant 1P-YFP X chromosomes, and average number of progeny among all 1P-YFP genotypes. Colored bars and arrow below indicate individual male and mean fertility for *D. simulans*  $w^{XD1}$ , respectively.



**Figure S2.** SNP locations and inferred ancestry for five recombinant 1P-YFP genotypes. Red ticks indicate *D. simulans* alleles (par1), blue ticks indicate *D. mauritiana* alleles (par2), and the red (blue) shaded regions indicate the location of inferred *D. simulans* (*D. mauritiana*) ancestry.



**Figure S3.** QTL analysis of male fertility. Recombinant 1P-YFP genotypes from each 2P introgression were used as separate mapping populations. Light grey lines in the background show LOD scores derived from using all 1P-YFP genotypes as a single mapping population (see Figure 3), red dashed line indicates significance threshold ( $P < 0.01$ ) determined from 10,000 permutations.



**Figure S4.** QTL analysis of progeny sex ratio associated with introgression genotypes. Dark grey and light grey regions indicate 5% and 1% significance thresholds determined from 10,000 random permutations. Top panel includes all males that produced any offspring; bottom panel includes only males that sired more than four offspring and genotypes with at least three males that sired more than four offspring.

620 LITERATURE CITED

- 621 ANDOLFATTO, P., DAVISON, D., EREZYILMAZ, D., HU, T. T., MAST, J.,  
622 SUNAYAMA-MORITA, T. & STERN, D. L. 2011. Multiplexed shotgun  
623 genotyping for rapid and efficient genetic mapping. *Genome Research*, 21, 610-  
624 617.
- 625 BALLARD, J. W. 2004. Sequential evolution of a symbiont inferred from the host:  
626 Wolbachia and *Drosophila simulans*. *Mol Biol Evol*, 21, 428-42.
- 627 BALLARD, J. W. O. 2000a. Comparative genomics of mitochondrial DNA in members  
628 of the *Drosophila melanogaster* subgroup. *Journal of Molecular Evolution*, 51,  
629 48-63.
- 630 BALLARD, J. W. O. 2000b. When one is not enough: introgression of mitochondrial  
631 DNA in *Drosophila*. *Molecular Biology and Evolution*, 17, 1126-1130.
- 632 BARTON, N. & BENGTSSON, B. O. 1986. The barrier to genetic exchange between  
633 hybridising populations. *Heredity (Edinb)*, 57 ( Pt 3), 357-76.
- 634 BAUDRY, E., DEROME, N., HUET, M. & VEUILLE, M. 2006. Contrasted  
635 polymorphism patterns in a large sample of populations from the evolutionary  
636 genetics model *Drosophila simulans*. *Genetics*, 173, 759-67.
- 637 BEGUN, D. J., HOLLOWAY, A. K., K., S., HILLIER, L. W., POH, Y. P., HAHN, M.  
638 W., NISTA, P. M., JONES, C. D., KERN, A. D., DEWEY, C. N., PACHTER, L.,  
639 MYERS, E. & LANGLEY, C. H. 2007. Population genomics: Whole-genome  
640 analysis of polymorphism and divergence in *Drosophila simulans* *Public Library*  
641 *of Science Biology*, 5, 2534-2559
- 642 BENGTSSON, B. O. 1985. The flow of genes through a genetic barrier. In:  
643 GREENWOOD, P. J., HARVEY, P. H. & SLATKIN, M. (eds.) *Evolution: essays*  
644 *in honour of John Maynard Smith*. Cambridge: Cambridge University Press.
- 645 CHARLESWORTH, B., COYNE, J. A. & BARTON, N. H. 1987. The Relative Rates of  
646 Evolution of Sex-Chromosomes and Autosomes. *American Naturalist*, 130, 113-  
647 146.
- 648 COYNE, J. A. 1992a. Genetics and speciation. *Nature*, 355, 511-515.
- 649 COYNE, J. A. 1992b. Genetics of sexual isolation in females of the *Drosophila simulans*  
650 species complex. *Genet. Res. Cambridge*, 60, 25-31.
- 651 COYNE, J. A. & ORR, H. A. 1989. Two rules of speciation. In: OTTE, D. & ENDLER,  
652 J. (eds.) *Speciation and Its Consequences*. Sunderland, MA: Sinauer Associates.
- 653 COYNE, J. A. & ORR, H. A. 2004. *Speciation*, Sunderland, Massachusetts, Sinauer.
- 654 CRESPI, B. & NOSIL, P. 2013. Conflictual speciation: species formation via genomic  
655 conflict. *Trends Ecol Evol*, 28, 48-57.
- 656 DAVID, J., MCEVEY, S. F., SOLIGNAC, M. & TSACAS, L. 1989. *Drosophila*  
657 communities on Mauritius and ecological niche of *D. mauritiana* (Diptera,  
658 Drosophilidae). *Journal of African Zoology*, 103, 107-116.
- 659 DEAN, M. D. & BALLARD, J. W. O. 2004. Linking phylogenetics with population  
660 genetics to reconstruct the geographic origin of a species. *Molecular*  
661 *Phylogenetics and Evolution*, 32, 998-1009.
- 662 FRANK, S. H. 1991. Divergence of meiotic drive-suppressors as an explanation for sex-  
663 biased hybrid sterility and inviability. *Evolution*, 45, 262-267.
- 664 GARRIGAN, D. 2013. POPBAM: Tools for Evolutionary Analysis of Short Read  
665 Sequence Alignments. *Evol Bioinform Online*, 9, 343-53.

- 666 GARRIGAN, D., KINGAN, S. B., GENEVA, A. J., ANDOLFATTO, P., CLARK, A. G.,  
667 THORNTON, K. R. & PRESGRAVES, D. C. 2012. Genome sequencing reveals  
668 complex speciation in the *Drosophila simulans* clade. *Genome Res.*, 22, 1499-511.  
669 GARRIGAN, D., KINGAN, S. B., GENEVA, A. J., VEDANAYAGAM, J. P. &  
670 PRESGRAVES, D. C. 2014. Genome diversity and divergence in *Drosophila*  
671 *mauritiana*: multiple signatures of faster X evolution. *Genome Biol Evol*, 6, 2444-  
672 58.  
673 GENEVA, A. J., MUIRHEAD, C. A., KINGAN, S. B. & GARRIGAN, D. 2015. A new  
674 method to scan genomes for introgression in a secondary contact model. *PLoS*  
675 *One*, 10, e0118621.  
676 GOOD, J. M., DEAN, M. D. & NACHMAN, M. W. 2008. A complex genetic basis to X-  
677 Linked hybrid male sterility between two species of house mice. *Genetics*, 179,  
678 2213-2228.  
679 HALDANE, J. B. S. 1922. Sex ratio and unisexual sterility in animal hybrids. *Journal of*  
680 *Genetics*, 12, 101-109.  
681 HELLEU, Q., GERARD, P. R., DUBRUILLE, R., OGEREAU, D., PRUD'HOMME, B.,  
682 LOPPIN, B. & MONTCHAMP-MOREAU, C. 2016. Rapid evolution of a Y-  
683 chromosome heterochromatin protein underlies sex chromosome meiotic drive.  
684 *Proc Natl Acad Sci U S A*, 113, 4110-5.  
685 HEY, J. & KLIMAN, R. M. 1993. Population genetics and phylogenetics of DNA  
686 sequence variation at multiple loci within the *Drosophila melanogaster* species  
687 complex. *Molecular Biology and Evolution*, 10, 804-822.  
688 HOLLINGER, I. & HERMISSON, J. 2017. Bounds to parapatric speciation: A  
689 Dobzhansky-Muller incompatibility model involving autosomes, X chromosomes,  
690 and mitochondria. *Evolution*, 71, 1366-1380.  
691 HURST, L. D. & POMIANKOWSKI, A. 1991. Causes of sex ratio bias may account for  
692 unisexual sterility in hybrids: a new explanation of Haldane's rule and related  
693 phenomena. *Genetics*, 128, 841-858.  
694 JAENIKE, J. 2001. Sex chromosome meiotic drive. *Annual Review of Ecology and*  
695 *Systematics*, 32, 25-49.  
696 KELLY, J. K. 1997. A test of neutrality based on interlocus associations. *Genetics*, 146,  
697 1197-1206.  
698 KINGAN, S. B., GARRIGAN, D. & HARTL, D. L. 2010. Recurrent selection on the  
699 Winters sex-ratio genes in *Drosophila simulans*. *Genetics*, 184, 253-65.  
700 KLIMAN, R. M., ANDOLFATTO, P., COYNE, J. A., DEPAULIS, F., KREITMAN, M.,  
701 BERRY, A. J., MCCARTER, J., WAKELEY, J. & HEY, J. 2000. The population  
702 genetics of the origin and divergence of the *Drosophila simulans* complex species.  
703 *Genetics*, 156, 1913-1931.  
704 KOPP, A. 2006. Basal relationships in the *Drosophila melanogaster* species group. *Mol*  
705 *Phylogenet Evol*, 39, 787-98.  
706 LACHAISE, D., CARIOU, M.-L., DAVID, J. R., LEMEUNIER, F., TSACAS, L. &  
707 ASHBURNER, M. 1988. Historical biogeography of the *Drosophila*  
708 *melanogaster* species subgroup. *Evolutionary Biology*, 22, 159-225.  
709 LACHAISE, D., DAVID, J. R., LEMEUNIER, F., TSACAS, L. & ASHBURNER, M.  
710 1986. The reproductive relationships of *Drosophila sechellia* with *D. mauritiana*,

- 711 D. simulans, and D. melanogaster from the Afrotropical region. *Evolution*, 1986,  
712 262-271.
- 713 LARSON, E. L., WHITE, T. A., ROSS, C. L. & HARRISON, R. G. 2014. Gene flow and  
714 the maintenance of species boundaries. *Mol Ecol*, 23, 1668-78.
- 715 LAURIE, C. C. 1997. The weaker sex is heterogametic: 75 years of Haldane's rule.  
716 *Genetics*, 147, 937-951.
- 717 LEGRAND, D., CHENEL, T., CAMPAGNE, C., LACHAISE, D. & CARIOU, M. L.  
718 2011. Inter-island divergence within *Drosophila mauritiana*, a species of the *D.*  
719 *simulans* complex: Past history and/or speciation in progress? *Mol Ecol*, 20, 2787-  
720 804.
- 721 LI, H. & DURBIN, R. 2009. Fast and accurate short read alignment with Burrows-  
722 Wheeler transform. *Bioinformatics*, 25, 1754-60.
- 723 LI, H., HANDSAKER, B., WYSOKER, A., FENNELL, T., RUAN, J., HOMER, N.,  
724 MARTH, G., ABECASIS, G., DURBIN, R. & GENOME PROJECT DATA  
725 PROCESSING, S. 2009. The Sequence Alignment/Map format and SAMtools.  
726 *Bioinformatics*, 25, 2078-9.
- 727 LIFSCHYTZ, E. & LINDSLEY, D. L. 1972. The role of the X-chromosome inactivation  
728 during spermatogenesis. *Proceedings of the National Academy of Sciences*, 69,  
729 182-186.
- 730 LINDHOLM, A. K., DYER, K. A., FIRMAN, R. C., FISHMAN, L., FORSTMEIER, W.,  
731 HOLMAN, L., JOHANNESSEN, H., KNIEF, U., KOKKO, H.,  
732 LARRACUENTE, A. M., MANSER, A., MONTCHAMP-MOREAU, C.,  
733 PETROSYAN, V. G., POMIANKOWSKI, A., PRESGRAVES, D. C.,  
734 SAFRONOVA, L. D., SUTTER, A., UNCKLESS, R. L., VERSPOOR, R. L.,  
735 WEDELL, N., WILKINSON, G. S. & PRICE, T. A. 2016. The Ecology and  
736 Evolutionary Dynamics of Meiotic Drive. *Trends Ecol Evol*.
- 737 MACAYA-SANZ, D., SUTER, L., JOSEPH, J., BARBARA, T., ALBA, N.,  
738 GONZALEZ-MARTINEZ, S. C., WIDMER, A. & LEXER, C. 2011. Genetic  
739 analysis of post-mating reproductive barriers in hybridizing European *Populus*  
740 species. *Heredity (Edinb)*, 107, 478-86.
- 741 MASIDE, X. R., BARRAL, J. P. & NAVERA, H. F. 1998. Hidden effects of X  
742 chromosome introgressions on spermatogenesis in *Drosophila simulans* x *D.*  
743 *mauritiana* hybrids unveiled by interactions among minor genetic factors.  
744 *Genetics*, 150, 745-54.
- 745 MASLY, J. P. & PRESGRAVES, D. C. 2007. High-resolution genome-wide dissection  
746 of the two rules of speciation in *Drosophila*. *Public Library of Science Biology*, 5,  
747 1890-1898.
- 748 MCDERMOTT, S. R. & KLIMAN, R. M. 2008. Estimation of isolation times of the  
749 island species in the *Drosophila simulans* complex from multilocus sequence data.  
750 *Plos ONE*, 3, e2442.
- 751 MEIKLEJOHN, C. D. & TAO, Y. 2010. Genetic conflict and sex chromosome evolution.  
752 *Trends Ecol Evol*, 25, 215-23.
- 753 MOEHRING, A. J., TEETER, K. C. & NOOR, M. A. 2007. Genome-wide patterns of  
754 expression in *Drosophila* pure species and hybrid males. II. Examination of  
755 multiple-species hybridizations, platforms, and life cycle stages. *Mol Biol Evol*,  
756 24, 137-45.

- 757 MONTCHAMP-MOREAU, C., OGEREAU, D., CHAMINADE, N., COLARD, A. &  
758 AULARD, S. 2006. Organization of the *sex-ratio* meiotic drive region in  
759 *Drosophila simulans*. *Genetics*, 166, 1357-1366.  
760 MOYLE, L. C., MUIR, C. D., HAN, M. V. & HAHN, M. W. 2010. The contribution of  
761 gene movement to the "two rules of speciation". *Evolution*, 64, 1541-57.  
762 MUIRHEAD, C. A. & PRESGRAVES, D. C. 2016. Hybrid Incompatibilities, Local  
763 Adaptation, and the Genomic Distribution of Natural Introgression between  
764 Species. *Am Nat*, 187, 249-61.  
765 NEI, M. 1987. *Molecular Evolutionary Genetics*, New York, Columbia University Press.  
766 NEI, M. & LI, W. H. 1979. Mathematical model for studying genetic variation in terms  
767 of restriction endonucleases. *Proc Natl Acad Sci U S A*, 76, 5269-73.  
768 NOLTE, V., PANDEY, R. V., KOFLER, R. & SCHLÖTTERER, C. 2013. Genome-wide  
769 patterns of natural variation reveal strong selective sweeps and ongoing genomic  
770 conflict in *Drosophila mauritiana*. *Genome Res*, 23, 99-110.  
771 ORR, H. A. 1997. Haldane's rule. *Annual Review of Ecology and Systematics*, 28, 195-  
772 218.  
773 PELUFFO, A. E., NUEZ, I., DEBAT, V., SAVISAAR, R., STERN, D. L. &  
774 ORGOGOZO, V. 2015. A Major Locus Controls a Genital Shape Difference  
775 Involved in Reproductive Isolation Between *Drosophila yakuba* and *Drosophila*  
776 *santomea*. *G3 (Bethesda)*, 5, 2893-901.  
777 PETRY, D. 1983. The effect on neutral gene flow of selection at a linked locus. *Theor  
778 Popul Biol*, 23, 300-13.  
779 PHADNIS, N. & ORR, H. A. 2008. A single gene causes both male sterility and  
780 segregation distortion in *Drosophila* hybrids. *Science*, 323, 376-379.  
781 PRESGRAVES, D. C. 2002. Patterns of postzygotic isolation in Lepidoptera. *Evolution*,  
782 56, 1168-1183.  
783 PRESGRAVES, D. C. 2008. Sex chromosomes and speciation in *Drosophila*. *Trends in  
784 Genetics*, 24, 336-343.  
785 PRICE, C. S. C. 1997. Conspecific sperm precedence in *Drosophila*. *Nature*, 388, 663-  
786 666.  
787 PRICE, T. D. & BOUVIER, M. M. 2002. The evolution of F1 postzygotic  
788 incompatibilities in birds. *Evolution*, 56, 2083-2089.  
789 ROGERS, R. L., CRIDLAND, J. M., SHAO, L., HU, T. T., ANDOLFATTO, P. &  
790 THORNTON, K. R. 2014. Landscape of standing variation for tandem  
791 duplications in *Drosophila yakuba* and *Drosophila simulans*. *Mol Biol Evol*, 31,  
792 1750-66.  
793 SATTA, Y. & TAKAHATA, N. 1990. Evolution of *Drosophila* mitochondrial DNA and  
794 the history of the melanogaster subgroup. *Proc Natl Acad Sci U S A*, 87, 9558-62.  
795 SATTA, Y., TOYOHARA, N., OHTAKA, C., TATSUNO, Y., WATANABE, T. K.,  
796 MATSUURA, E. T., CHIGUSA, S. I. & TAKAHATA, N. 1988. Dubious  
797 Maternal Inheritance of Mitochondrial-DNA in *Drosophila-Simulans* and  
798 Evolution of D-Mauritiana. *Genetical Research*, 52, 1-6.  
799 SEEHAUSEN, O., BUTLIN, R. K., KELLER, I., WAGNER, C. E., BOUGHMAN, J. W.,  
800 HOHENLOHE, P. A., PEICHEL, C. L., SAETRE, G. P., BANK, C.,  
801 BRANNSTROM, A., BRELSFORD, A., CLARKSON, C. S.,  
802 EROUKHMANOFF, F., FEDER, J. L., FISCHER, M. C., FOOTE, A. D.,

- 803 FRANCHINI, P., JIGGINS, C. D., JONES, F. C., LINDHOLM, A. K., LUCEK,  
804 K., MAAN, M. E., MARQUES, D. A., MARTIN, S. H., MATTHEWS, B.,  
805 MEIER, J. I., MOST, M., NACHMAN, M. W., NONAKA, E., RENNISON, D. J.,  
806 SCHWARZER, J., WATSON, E. T., WESTRAM, A. M. & WIDMER, A. 2014.  
807 Genomics and the origin of species. *Nat Rev Genet*, 15, 176-92.
- 808 SLATKIN, M. 1993. Isolation by Distance in Equilibrium and Non-Equilibrium  
809 Populations. *Evolution*, 47, 264-279.
- 810 SOLIGNAC, M. & MONNEROT, M. 1986. Race formation, speciation, and  
811 introgression within *Drosophila simulans*, *D. mauritiana*, and *D. sechellia*  
812 inferred from mitochondrial DNA analysis. *Evolution*, 40, 531-539.
- 813 SOLIGNAC, M., MONNEROT, M. & MOUNOLOU, J.-C. 1986. Mitochondrial DNA  
814 evolution in the melanogaster species subgroup of *Drosophila*. *Journal of*  
815 *Molecular Evolution*, 23, 31-40.
- 816 STAMATAKIS, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-  
817 analysis of large phylogenies. *Bioinformatics*, 30, 1312-3.
- 818 STERN, D. L., CROCKER, J., DING, Y., FRANKEL, N., KAPPES, G., KIM, E.,  
819 KUZMICKAS, R., LEMIRE, A., MAST, J. D. & PICARD, S. 2017. Genetic and  
820 Transgenic Reagents for *Drosophila simulans*, *D. mauritiana*, *D. yakuba*, *D.*  
821 *santomea*, and *D. virilis*. *G3 (Bethesda)*, 7, 1339-1347.
- 822 TAJIMA, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA  
823 polymorphism. *Genetics*, 123, 585-595.
- 824 TAO, Y., ARARIPE, L., KINGAN, S. B., KE, Y., XIAO, H. & HARTL, D. L. 2007a. A  
825 sex-ratio meiotic drive system in *Drosophila simulans*. II: An X-linked distorter.  
826 *Public Library of Science Biology*, 5, e293.
- 827 TAO, Y., CHEN, S., HARTL, D. L. & LAURIE, C. C. 2003. Genetic dissection of  
828 hybrid incompatibilities between *Drosophila simulans* and *D. mauritiana*. I.  
829 Differential accumulation of hybrid male sterility effects on the X and autosomes.  
830 *Genetics*, 164, 1383-1397.
- 831 TAO, Y. & HARTL, D. L. 2003. Genetic dissection of hybrid incompatibilities between  
832 *Drosophila simulans* and *D. mauritiana*. III. Heterogeneous accumulation of  
833 hybrid incompatibilities, degree of dominance, and implications for Haldane's rule.  
834 *Evolution*, 57, 2580-2589.
- 835 TAO, Y., HARTL, D. L. & LAURIE, C. C. 2001. Sex-ratio segregation distortion  
836 associated with reproductive isolation in *Drosophila*. *Proceedings of the National*  
837 *Academy of Sciences*, 98, 13183-13188.
- 838 TAO, Y., MASLY, J. P., ARARIPE, L., KE, Y. & HARTL, D. L. 2007b. A sex-ratio  
839 meiotic drive system in *Drosophila simulans*. I: An autosomal suppressor. *Public*  
840 *Library of Science Biology*, 5, e292.
- 841 TEAM, R. D. C. 2013. R: A Language and Environment for Statistical Computing.  
842 Vienna, Austria: R Foundation for Statistical Computing.
- 843 TRUE, J. R., MERCER, J. M. & LAURIE, C. C. 1996a. Differences in crossover  
844 frequency and distribution among three sibling species of *Drosophila*. *Genetics*,  
845 142, 507-523.
- 846 TRUE, J. R., WEIR, B. S. & LAURIE, C. C. 1996b. A genome-wide survey of hybrid  
847 incompatibility factors by the introgression of marked segments of *Drosophila*  
848 *mauritiana* chromosomes into *Drosophila simulans*. *Genetics*, 142, 819-837.

- 849 WRIGHT, S. 1951. The genetical structure of populations. *Ann Eugen*, 15, 323-54.  
850 WU, C.-I. & DAVIS, A. W. 1993. Evolution of postmating reproductive isolation: the  
851 composite nature of Haldane's rule and its genetic bases. *American Naturalist*,  
852 142, 187-212.  
853 ZHANG, L., SUN, T., WOLDESELLASSIE, F., XIAO, H. & TAO, Y. 2015. Sex ratio  
854 meiotic drive as a plausible evolutionary mechanism for hybrid male sterility.  
855 *PLoS Genet*, 11, e1005073.

856

857