MOLECULAR ECOLOGY

Molecular Ecology (2013) **22**, 157–169 doi: 10.1111/mec.12097

Genomic conflict drives patterns of X-linked population structure in *Drosophila neotestacea*

KELLY A. DYER, MICHAEL J. BRAY and S. JEESE LOPEZ
Department of Genetics, University of Georgia, Athens, GA 30602, USA

Abstract

Intragenomic conflict has the potential to cause widespread changes in patterns of genetic diversity and genome evolution. In this study, we investigate the consequences of *sex-ratio* (SR) drive on the population genetic patterns of the X-chromosome in *Drosophila neotestacea*. An SR X-chromosome prevents the maturation of Y-bearing sperm during male spermatogenesis and thus is transmitted to ~100% of the offspring, nearly all of which are daughters. Selection on the rest of the genome to suppress SR can be strong, and the resulting conflict over the offspring sex ratio can result in the accumulation of multiple loci on the X-chromosome that are necessary for the expression of drive. We surveyed variation at 12 random X-linked microsatellites across 16 populations of *D. neotestacea* that range in SR frequency from 0% to 30%. First, every locus was differentiated between SR and wild-type chromosomes, and this drives genetic structure at the X-chromosome. Once the association with SR is accounted for, the patterns of differentiation among populations are similar to the autosomes. Second, within wild-type chromosomes, the relative heterozygosity is reduced in populations with an increased prevalence of drive, and the heterozygosity of SR chromosomes is higher than expected based on its prevalence. The combination of the relatively high prevalence of SR drive and the structuring of polymorphism between the SR and wild-type chromosomes suggests that genetic conflict because of SR drive has had significant consequences on the patterns of X-linked polymorphism and thus also probably affects the tempo of X-chromosome evolution in *D. neotestacea*.

Keywords: demography, meiotic drive, selection, sex ratio, X-chromosome

Received 24 June 2012; revision received 11 September 2012; accepted 17 September 2012

Introduction

Genetic conflict occurs when one part of the genome interferes with the transmission of another part of the genome (Werren *et al.* 1988; Burt & Trivers 2006). One form of genetic conflict, meiotic drive, can occur anywhere in the genome but seems to be especially common on the sex chromosomes (reviewed in Jaenike 2001). When a meiotic drive factor is located on the X-chromosome, this driving X-chromosome prevents the maturation of the Y-chromosome carrying spermatids during male spermatogenesis (reviewed in Jaenike 2001). Thus, a male carrying a driving X-chromosome transfers mostly X-bearing sperm to the female and sires mostly daughters. X-chromosome drive may more easily evolve than autosomal drive because of the wide-scale genetic differences and lack of recombination between the X- and Y-chromosomes (Jaenike 2001). However, because this '*sex-ratio*' (SR) drive results in a female-biased offspring sex ratio, it may only appear to be more common because it is easier to detect.

Altering the offspring sex ratio can result in strong selection that favours the spread of autosomal and Y-linked suppressors that restore the production of sons (Hamilton 1967). In turn, the driving chromosome may evolve modifiers that restore drive. If this 'arms race' within the genome proceeds such that multiple X-linked factors become necessary for the expression of drive, inversions or other mechanisms that suppress local recombination may be favoured to maintain linkage disequilibrium (LD) among the genes necessary for

Correspondence: Kelly A. Dyer, Fax: 706-542-3910;
E-mail: kdyer@uga.edu

drive (reviewed in Jaenike 2001). This reduced recombination between SR and wild-type chromosomes can create two independent subpopulations of X-chromosomes. Thus, the genetic conflict that arises from meiotic drive has the potential to shape the amount and structure of population genetic diversity as well as the rate of molecular evolution and efficacy of selection on the X-chromosome. Over the long term, the evolutionary consequences of the genetic conflict arising from X-chromosome drive may be significant and include changes in the genomic distribution of genes expressed in the germline, changes in sex determination, the evolution of hybrid sterility and the epigenetic regulation of sex chromosomes during meiosis (reviewed in Meiklejohn & Tao 2010; Presgraves 2010; Kaiser & Bachtrog 2010).

In the short term, the molecular evolutionary consequence of drive can vary depending on the evolutionary age of the drive system and the extent to which inversions have accumulated. At one extreme is the very young Paris SR drive system in *Drosophila simulans*, where the driving factor is contained within a ~40-kb segmental duplication that is not associated with an inversion, and the genomic region also exhibits the signatures of a recent selective sweep (Derome *et al.* 2004; Montchamp-Moreau *et al.* 2006; Bastide *et al.* 2011; Fouvry *et al.* 2011). At the other extreme is the much older SR chromosome in *D. recens*, which has accumulated multiple inversions and is in LD with a recessive female sterile mutation. Linkage disequilibrium between the SR and the wild-type Standard (ST) chromosome extends for over 100 cM and is because of the association of driving loci and modifiers that are tied up in inversions (Dyer *et al.* 2007). In other characterized drive systems, including the autosomal *Segregation Distorter* (*SD*) chromosome in *D. melanogaster*, SR in the stalk-eyed fly *Teleopsis dalmanni* and SR in *D. pseudoobscura*, the driving chromosome is associated with one or more inversions, and there is genetic differentiation between the driving and ST chromosomes, but it does not extend across the entire chromosome as in *D. recens* (Kovacevic & Schaeffer 2000; Presgraves *et al.* 2009; Christianson *et al.* 2011). From both comparative studies as well as careful dissection of individual drive systems, it is becoming clear that the evolutionary dynamics of this type of genetic conflict can occur rapidly and even on an ecological timescale (e.g. Presgraves *et al.* 2009; Bastide *et al.* 2011).

The consequences of SR drive for overall patterns of population genetic variation of the X-chromosome will depend not only on the age of the drive system and extent to which the SR and ST chromosomes are differentiated genetically, but also on the prevalence of SR in the population and/or species. In general, the genetic region associated with drive often has reduced polymorphism on the driving chromosome relative to nondriving chromosomes, which can often be attributed to a recent selective sweep on the drive chromosome (e.g. Derome *et al.* 2004; Presgraves *et al.* 2009). In addition, if SR is at low frequency and/or restricted to a small part of the species' range, the low effective population size can also restrict polymorphism on the SR chromosome (e.g. Dyer *et al.* 2007). However, while SR can have significant consequences on variation within this subset of chromosomes, its effect on overall patterns of genetic variation may be minimal. For example, in *D. recens*, SR is only found in about ~2% of males, so even though the SR chromosome contains very low polymorphism relative to the wild-type chromosome, because it is rare the X-chromosome still harbours an overall very high level of genetic polymorphism because the wild-type chromosome harbours a substantial amount of diversity (Dyer *et al.* 2007). One could imagine that the long-term consequences of harbouring drive for patterns of polymorphism and divergence could become significant if the driving chromosome rises to appreciable frequency but still harbours lower than expected variation, or if there is enough genetic exchange between the driving and nondriving chromosomes such that the level of variation in the wild-type chromosomes is decreased.

Here, we study the population genetic consequences of SR drive in the fly *Drosophila neotestacea*. *Drosophila neotestacea* is a mushroom-feeding fly that is common in the temperate and boreal forests of North America; it is found in western North America from Oregon north to Alaska and occurs across the continent to the east, extending south along the Appalachian Mountains to the Smokies. Populations of *D. neotestacea* vary in the prevalence of SR drive, ranging from 0% to 30% of males, with a species-wide estimate of about 15.5% (Dyer 2012). Variation in SR frequency is probably due to local selection among populations, as the frequency of SR drive is not because of the localized presence of dominant autosomal or Y-linked suppressors of drive or to generally restricted gene flow. Instead, it is correlated with local climate, with warmer locations harbouring a higher prevalence of SR drive than cooler locations. The underlying cause of this spatial variation is unknown (Dyer 2012).

To investigate the effect of SR for the evolution of the X-chromosome in *D. neotestacea*, we first ask whether loci on the X-chromosome of *D. neotestacea* show an association with SR drive. We survey patterns of diversity at 12 random X-linked microsatellite loci and find large-scale genetic differentiation between SR and Standard (ST or wild type) chromosomes. We then investigate how SR has affected patterns of X-linked

population differentiation in *D. neotestacea*, and disentangle the effects of SR from neutral demographical processes that may affect each subpopulation of X-chromosomes independently. Finally, because SR is at relatively high prevalence in many populations, we ask whether SR affects general patterns of polymorphism on the X-chromosome. In sum, our results suggest that the X-chromosome in *D. neotestacea* has been shaped to a large extent by the genetic conflict because of SR drive, with potential consequences for the efficacy of selection and the rate of evolution of the X-chromosome.

Materials and methods

Drosophila samples and microsatellite genotyping

We used wild-caught *Drosophila neotestacea* from 16 populations that span most of the geographical range of this species. These populations are described in Table 1 and are the same locations used in Dyer (2012). Flies were collected by sweep-netting over baits of store-bought mushrooms or naturally occurring mushrooms. A total of 485 wild-caught flies were used in this study, which included 16–44 flies per population (Table 1). The 12 X-linked loci used in this study are listed in Table 2. Eleven of the loci were initially described in Dyer (2007) and were isolated from *D. neotestacea*. The twelfth locus, tra84, was isolated from *D. transversa* (Räisänen *et al.* 2009), and our preliminary studies revealed that it was also X-linked and polymorphic in *D. neotestacea*. Of these 12 loci, four are trinucleotide repeats and eight are dinucleotide repeats (Table 2). X-linkage was inferred because males only ever had one allele, whereas females usually had two alleles. These loci are random with respect to physical location on the X-chromosome and have not been mapped. Based on sequencing of protein-coding loci, the X-chromosome in *D. neotestacea* is the same Muller's Element as the X-chromosome in *D. melanogaster* and most other *Drosophila* species (Patterson & Stone 1952; Dyer *et al.* 2011; K. Dyer, unpublished data).

Methods for microsatellite genotyping and fragment analysis are as described previously (Dyer 2007, 2012), and the data set has been deposited in Dryad (doi:10.5061/dryad.2d315). A subset of the genotyped males were among those assayed for offspring sex ratio in Dyer (2012) and were categorized as SR if they produced at least 20 offspring, of which at least 90% were female. Males were categorized as wild type (Standard or ST), if they sired <90% female offspring, as long as they sired at least 20 offspring. Across the 216 males characterized as SR in Dyer (2012), the average offspring sex ratio was 99.3 ± 0.001% daughters, and the distribution of offspring sex ratios of all wild-caught males was highly bimodal. The rare males sired by SR males are always sterile and thus assumed to be XO

Table 1 Populations of *Drosophila neotestacea* used in this study, including the frequency of *sex-ratio* as well as sample sizes for microsatellite genotyping. Populations are listed by latitude from north to south

						Number assayed at microsatellites			
Abr.	Site	Date	Latitude	Longitude	*Sex-ratio* frequency* (±95% CI)	Total	Random sample	*Sex-ratio* males	Standard males
AB1	Winston Churchill, AB	2002	54.82	−111.98	0.01 (0, 0.05)	30	26	1	16
MB	The Pas, MB	2002	53.96	−101.10	0 (0, 0.06)	30	26	0	17
AB2	Edmonton, AB	2002	53.51	−113.54	0 (0, 0.13)	30	26	0	17
AB3	Jasper, AB	2002	52.84	−118.07	0.05 (0.01, 0.13)	31	26	4	13
BC	Vancouver, BC	2001	49.33	−122.97	0.19 (0.13, 0.27)	34	26	9	10
ON	Dogfly Lake, ON	2001	49.10	−93.12	Not surveyed	32	32	0	0
MT1	Columbia Falls, MT	2002	48.46	−113.98	0.19 (0.09, 0.33)	30	26	5	12
ND	Minot, ND	2002	48.24	−101.36	0.07 (0.01, 0.23)	30	26	2	11
ID	Coeur d'Alene ID	2001	47.61	−116.67	0.12 (0.07, 0.18)	36	26	11	12
MN	Bemidji, MN	2002	47.42	−94.70	0.25 (0.14, 0.39)	30	26	6	11
MT2	St. Regis, MT	2001	47.30	−115.10	0.18 (0.13, 0.24)	44	25	20	11
PEI	Charlottetown, PEI	2002	46.25	−63.17	0 (0, 0.37)	23	23	0	5
OR	MacKenzie Bridge, OR	2001	44.18	−122.16	0.26 (0.20, 0.33)	43	26	20	10
NY	Rochester, NY	2001	43.10	−77.65	0.23 (0.16, 0.31)	30	26	6	11
TN1	Gatlinburg, TN	2001	35.68	−83.50	0.25 (0.15, 0.37)	16	13	6	3
TN2	Clingmans Dome, TN	2001	35.60	−83.44	0.12 (0.06, 0.20)	16	14	3	7
MEAN	All populations				0.12 (0.07, 0.17)	30.3	24.6	5.8	10.4
SUM	All populations				0.16 (0.14, 0.18)	485	393	93	166

**Sex-ratio* frequency data from Dyer (2012).

Table 2 Estimates of polymorphism and differentiation

		Random				ST males			SR males			ST vs. SR	
Locus	bp	*N*	A	H_O	H_E	*N*	A	H_E	*N*	A	H_E	F_{ST}	R_{ST}
5260	3	543	5	0.500	0.547	147	5	0.580	83	3	0.094	**0.336**	0.004
5261	2	567	12	0.360	0.392	166	8	0.193	86	5	0.393	**0.689**	**0.218**
6002	2	532	23	**0.676**	0.926	150	21	0.929	84	12	0.830	**0.032**	**0.101**
7002	2	573	18	0.739	0.780	158	13	0.747	92	8	0.787	**0.082**	**0.055**
7027	3	561	8	0.646	0.662	164	6	0.603	86	3	0.474	**0.419**	**0.290**
7029	2	575	19	**0.817**	0.899	159	14	0.904	90	12	0.845	**0.034**	−0.007
7040	2	563	19	0.779	0.860	164	16	0.878	87	4	0.333	**0.298**	−0.009
7041	2	545	26	0.865	0.885	154	22	0.870	89	17	0.882	**0.082**	**0.340**
8349	2	559	11	0.602	0.676	159	8	0.644	87	4	0.153	**0.532**	**0.831**
8377	3	568	7	0.382	0.379	165	5	0.225	87	3	0.230	**0.756**	**0.774**
8385	3	545	18	0.812	0.886	154	16	0.882	88	7	0.558	**0.486**	**0.744**
tra84	2	535	32	**0.857**	0.942	149	25	0.937	83	3	0.094	**0.198**	**0.230**
Mean	na	556	16.5	**0.670**	0.736	157	13	0.699	87	7	0.507	**0.332**	**0.367**

Polymorphism is shown for the Random data set, Standard (ST) males and *sex-ratio* (SR) males, using only males with a known phenotype based on offspring sex ratio. Shown are the number of base pairs in the microsatellite repeat motif (bp), the number of sampled alleles (*N*), different alleles (A), the observed (H_O) and expected (H_E) heterozygosities within each data set, and F_{ST} and R_{ST} between SR and ST males. H_O values in bold indicate a deficiency of observed heterozygosity at $P < 0.01$. F_{ST} and R_{ST} values in bold are significant at $P < 0.0001$, based on 1000 permutations.

and the product of nondisjunction (James & Jaenike 1990; Dyer 2012). The population-level frequency of SR drive inferred previously from each of these populations is also shown in Table 1 (data from Dyer 2012). We note that the cytological phenotype of SR is similar to other species *Drosophila* with drive, in that in SR males, about half of developing sperm decay during spermatogenesis (K. Dyer, unpublished data).

We used several different subsets of data in the analyses that are described below. First, because the data set described above is enriched for males known to be SR, while females are sampled randomly, we also sampled both males and females randomly with respect to SR status (mean = 25 flies/population); we refer to this subset as the 'Random' data set ($N = 393$). In this way, sampling randomly without respect to SR will not bias results because certain genotypes are over-represented. Second, as the results will show, we find a strong genetic association with SR, and to separate the effects of SR from other processes, we used subsets of the data that consisted only of individuals known or inferred with high confidence to be SR or ST; we refer to these as the SR and ST data sets. These data sets thus represent the SR and ST subpopulations of X-chromosomes.

Patterns of diversity and differentiation

We tested for the presence of null alleles as implemented in GENEPOP v.4.1.3 (Raymond & Rousset 1995; Rousset 2008). We tested for LD between pairs of loci, departures from Hardy–Weinberg equilibrium (HWE), and calculated allele richness, observed and expected heterozygosity, and measures of global and pairwise population differentiation (F_{ST} and R_{ST}) for individual loci and across all loci using ARLEQUIN v.3.5.1.2 (Excoffier & Lischer 2010). Significance of F_{ST} and R_{ST} was determined with 1000 permutations.

To infer the number of genetic clusters (*K*) in the data, we used the program STRUCTURE v.2.3 (Pritchard *et al.* 2000). For all STRUCTURE analyses, we ran the program five times at each value of *K* (from one to 16), with a burn-in of 150 000 steps and a run length of 200 000 steps. We determined the most likely value of *K* using the highest log-likelihood of the posterior probability of the data [$\Pr(X|K)$] across values of *K* (Pritchard *et al.* 2000) and also via ΔK, which analyses the second-order rate of change in $\ln[\Pr(X|K)]$ with respect to *K* (Evanno *et al.* 2005). All models allowed for admixture and correlated allele frequencies, and we used the sampling location as a prior when analysing the Random, ST and SR data sets to detect subtle population structure.

To investigate the relationship between geographical distance and population differentiation, we used Mantel tests (Mantel 1967). The geographical distance between pairs of populations was determined using the latitude/longitude population coordinates, and F_{ST} was adjusted to Slatkin's linearization of $F_{ST}/(1-F_{ST})$ (Slatkin 1995). We also used Mantel tests to compare genetic differentiation to the difference in population structure based on SR frequency, where SR frequency was treated as a single locus, and estimates of pairwise F_{ST} between populations were obtained using the observed

proportions weighted by sample size, as described in Dyer (2012). Mantel tests were calculated as implemented in ARLEQUIN v3.5.1.2, with significance calculated using 10 000 matrix randomizations (Excoffier & Lischer 2010).

To compare the patterns of genetic diversity and differentiation at the X-chromosome vs. the autosomes, we used data from 468 of these same 485 individuals at five autosomal microsatellite loci that were previously shown to be in HWE and in linkage equilibrium with each other (Dyer 2012). We used Mantel tests as described above and compared the Random, ST and SR data sets to the autosomes. We adjusted F_{ST} to Slatkin's linearization of $F_{ST}/(1-F_{ST})$ and calculated significance using 10 000 matrix randomizations (Slatkin 1995; Excoffier & Lischer 2010).

To investigate an effect of SR on the overall level of genetic diversity, we asked whether the prevalence of SR is correlated with the level of X-linked genetic diversity. We also compared the X-chromosome to autosome (X/A) ratio of genetic diversity across populations; under neutral equilibrium conditions and an equal population sex ratio, the expected X/A ratio is 0.75 (reviewed in Vicoso & Charlesworth 2006). For both of these analyses, we used the expected heterozygosity averaged across loci and compared the Random, ST and SR data sets. Correlations with the population-level prevalence of SR were weighted by the number of X-chromosomes surveyed to determine the prevalence in each population (Table 1). Statistics were completed using JMP version 9.0.2 (SAS Institute, Cary, NC, USA).

Finally, we investigated the demographical patterns of the X-chromosome, asking whether there is a signature of a population expansion or contraction specific to the SR chromosome. An observation of an excess of heterozygosity relative to that expected based on the allele number may indicate a recent population bottleneck, and a deficiency of heterozygosity may indicate a growing population (Nei *et al.* 1975). We tested for a departure from mutation–drift equilibrium using the program BOTTLENECK (Cornuet & Luikart 1996; Piry *et al.* 1999), implementing an infinite alleles model, a pure stepwise mutation model (SMM) and a two-phase model (TPM), allowing for 5%, 10% and 20% multistep changes. The TPM model is most realistic for the mutational process of microsatellites (Di Rienzo *et al.* 1994; Estoup *et al.* 2002). We analysed the autosomes, Random, ST, and SR data sets separately and tested for significance using a two-tailed Wilcoxon signed rank test with 1000 coalescent simulations for each model. Because this test is sensitive to population size, for the ST data set, we binned TN1 and TN2 together for a total of 15 populations; for the SR data set, the OR, BC and MT2 were each treated separately, but we binned the MT1 and ID, TN1 and TN2, AB1, AB3, and ND, and MN, NY, and ON populations for a total of seven populations. Populations were binned together based on geographical proximity and genetic similarity.

Results

Association of X-linked markers with sex-ratio

We first asked whether there was an association of loci on the X-chromosome with *sex-ratio*. Combining individuals across populations and only using males with a known SR phenotype (SR or ST) based on offspring sex ratio, we find that at every locus there is a highly significant F_{ST} between ST and SR males ($P < 0.0001$), and this difference is also found at all but three loci at R_{ST} (Table 2). The global F_{ST} between SR and ST males is 0.332 ($P < 0.0001$) and $R_{ST} = 0.368$ ($P < 0.0001$). No locus is monomorphic within SR, only in a few instances is an allele found exclusively in SR, and these are always at low frequency (Table 2; Fig. S1, Supporting information). The results of the clustering analysis in the STRUCTURE analysis on the entire data set ($N = 485$), but which is enriched for SR males, indicate that the most likely number of clusters is two, with no clear geographical pattern of population composition within these clusters (Tables S1 and S2, Fig. S2a, Supporting information). Combining across populations, we asked whether there was an association of SR phenotype and genetic cluster. As can be seen in Fig. 1a, every ST male has a proportion of ancestry in Cluster 1 >0.6, with all but three >0.8 ($N = 166$; mean = 0.967 ± 0.004). Likewise, there is an equally strong association of the SR males belonging to the second cluster (Cluster 2; $N = 93$; mean = 0.967 ± 0.005). Using a logistic regression, the association of SR phenotype and membership in Cluster 1 is highly significant ($r^2 = 0.99$; $F_{1,257} = 22{,}312$; $P < 0.0001$). These results together suggest that there is widespread restriction of gene flow between the SR and ST chromosomal types.

Using the genetic association identified in males, we can also infer the SR genotypes of the females. In contrast to the bimodal distribution seen in the males, the females show a distinctly trimodal distribution of the genetic identity in Cluster 1, suggestive of the three different SR carrier statuses that can be found in females (Fig. 1b). While <1% of males have a probability of ancestry in Cluster 1 between 0.25 and 0.75, 20.2% females have a probability between 0.25 and 0.75 of ancestry in Cluster 1; most of these females are presumably heterozygous for *sex-ratio* (SR/ST). Likewise, the females with a probability of <0.25 in Cluster 1 are probably SR/SR, and those with >0.75 probability of ancestry in Cluster 1 are probably ST/ST. The

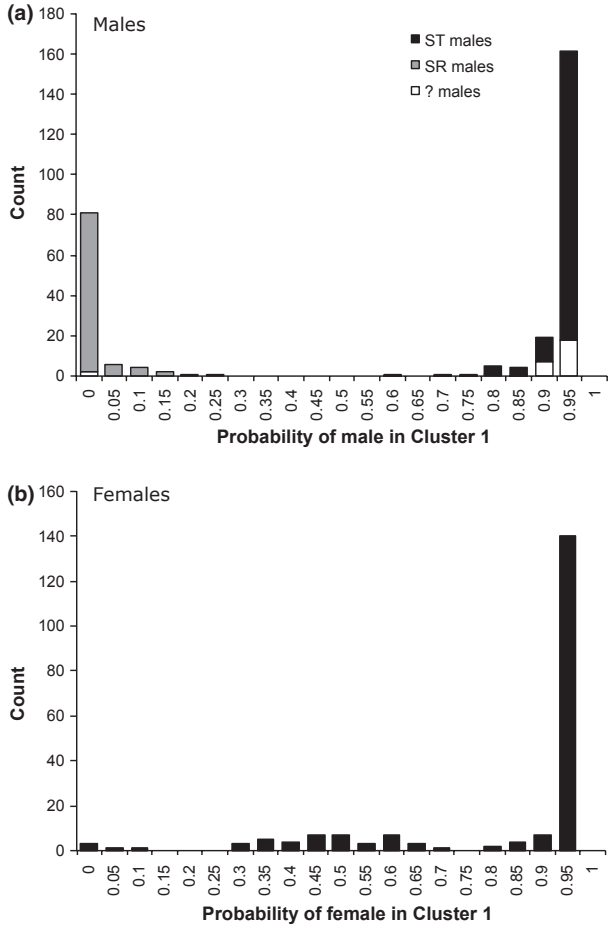

Fig. 1 The proportion of ancestry of each individual in genetic Cluster 1 from the STRUCTURE analysis. Individuals are separated into (a) males and (b) females. Males are distinguished by *sex-ratio* (SR) phenotype, based on their offspring sex ratio, with standard (ST) in black, SR in grey and males with an unknown SR phenotype in white. Among the males, there is a highly significant association of SR phenotype and percentage of ancestry in Cluster 1 ($r^2 = 0.99$; $F_{1,257} = 22\,312$; $P < 0.0001$). This association suggests that the trimodal distribution of the females is reflective of the three possible carrier types (SR/SR, SR/ST and ST/ST).

population-averaged expectation under HWE is to find SR/ST females at an overall frequency of 20.4% ($=2\bar{p}\bar{q}$) and SR/SR at a frequency of 2.6% ($=\bar{q}^2$). We observe SR/ST females at a frequency of 20.2% and SR/SR females at a frequency of 2.5%, almost exactly the expected values ($\chi^2 = 0.011$, $P = 0.92$).

Patterns of X-linked population diversity and differentiation

We investigated the patterns of X-linked diversity and population differentiation and attempted to disentangle the effects of an association with *sex-ratio* from those of neutral population demographical processes that may affect each subpopulation of X-chromosomes independently. First, to capture a random sample of genetic diversity, we analysed the Random data set ($N = 393$ across 16 populations), which as described earlier is not sampled with respect to male SR status. Second, we analysed data sets that consisted only of ST ($N = 345$ across 16 populations) and SR ($N = 100$ across 13 populations) individuals. To construct these data sets, the males with unknown SR status as well as all of the females were assigned an SR carrier status based on their membership in Cluster 1 of the complete data set. Males and homozygous SR/SR and ST/ST females were combined with the SR and ST males, respectively, to create the SR and ST data sets.

Within each of these three data sets, we did not identify any consistent evidence of null alleles at any locus, when all individuals were considered either together or separated by population. Patterns of genetic diversity by data set are summarized by locus in Table 2, and patterns of diversity for each locus and population are in Table S3 (Supporting information). Across the data sets, there is no consistent signature of a departure from HWE for any locus (Table 2, Table S3, Supporting information). In the Random data set, four pairs of loci are in LD within four or more populations at the $P < 0.01$ level (7040 and 7029, 5261 and 8377, 7029 and 8377, 7029 and 8385). However, within the ST or SR data sets, there is no LD between any pair of loci that is significant at the $P < 0.01$ level within more than two populations. This suggests that gene flow is reduced between but not within chromosomal types and that the loci we sampled are probably not physically clustered.

For the Random data set, which is sampled without respect to SR, using STRUCTURE yielded essentially the same results as we found for the complete data set that was enriched for SR males. The most likely number of clusters is $K = 2$ based on ΔK and $K = 3$ based on highest averaged lnL (Table S1, Supporting information). For the $K = 2$ results, because every population is most likely to be mostly in Cluster 1 (Fig. 2a, Table S2, Supporting information), we asked whether there was an association between SR genotype and cluster, as we found for the complete data set. Indeed, we find a strong association between male SR phenotype based on offspring sex ratio and proportion of ancestry in Cluster 1 ($r^2 = 0.94$; $F_{1,169} = 2922$; $P < 0.0001$), and we also find a strong association between the inferred genotype of females and membership in Cluster 1 ($r^2 = 0.95$; $F_{2,194} = 1836$; $P < 0.0001$), with the putative SR/ST individuals having a roughly equal ancestry in each cluster (mean Cluster 1 = 0.55 ± 0.016). Furthermore, we find a significant positive correlation between

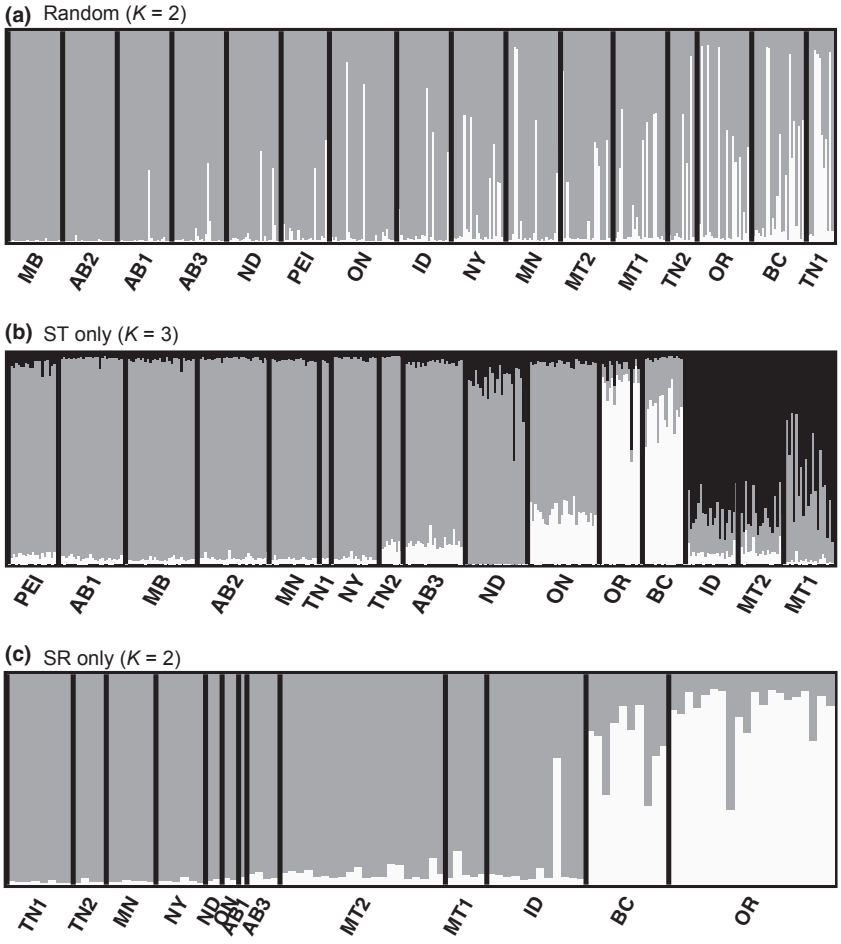

Fig. 2 Results of the STRUCTURE analyses. Individuals are grouped by population, with proportional ancestry of each individual from each cluster shown in white, grey or dark grey. Populations are sorted by genetic cluster composition, using the abbreviations in Table 1. Shown here are the results for (A) the random data set that is sampled without respect to *sex-ratio* (SR), with $K = 2$, (B) ST individuals only, with $K = 3$, and (C) SR individuals only, with $K = 2$. There is no geographical pattern found using the random data set; instead, clustering is by SR status. In contrast, within each chromosome type, the geographical structure is on an east–west gradient across the species range. See text for details.

the mean ancestry in Cluster 2 (the SR cluster) across individuals and the local prevalence of SR ($r^2 = 0.50$; $F_{1,13} = 13.3$; $P = 0.003$; Fig. 3). This means that the effect of SR on population structure is so strong that the local prevalence of SR can be estimated using genetic methods, even when phenotypic estimates using offspring sex ratios are not available. The difference between the $K = 2$ and $K = 3$ STRUCTURE results is that in $K = 3$, the ID, MT1 and MT2 populations form a separate though weakly supported cluster (Fig. S2b, Supporting information). Thus, from these analyses, we can conclude that the signal on patterns of polymorphism because of an association with SR is stronger than a neutral genetic signal because of limited gene flow among populations. Using the Random data set with standard indices of population differentiation, pairwise values of F_{ST} and R_{ST} that show the most differentiation are between the western and southern populations (OR, BC, MT, TN) and the main part of the range (Table S4, Supporting information). Using a Mantel test, we find a significant association between geographical and genetic distance ($r = 0.34$, $P = 0.021$) but not between SR frequency and genetic distance ($r = -0.15$, $P = 0.16$).

We next investigated the patterns of geographical differentiation within the ST and SR data sets, asking

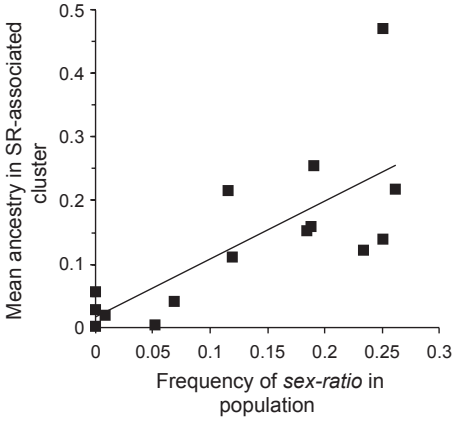

Fig. 3 Correlation between the frequency of *sex-ratio* (SR) in each population based on scoring male offspring sex ratio and the mean fraction of ancestry in the SR-associated Cluster 2 in the STRUCTURE analysis of the random data set. The significant positive correlation ($r^2 = 0.50$; $F_{1,13} = 13.3$; $P = 0.003$) suggests that genetic data can be used to estimate SR prevalence in natural populations.

whether patterns of geographical differentiation emerge once we remove the effect of SR. For the ST data set, the STRUCTURE analysis resulted in a most likely value of $K = 3$ (Tables S1 and S2, Supporting information, Fig. 2b). Overall, the western part of the range shows more differentiation than the eastern part of the range: one cluster formed the main part of the species range in eastern and central North America, a second cluster encompasses Idaho and Montana (MT1, ID, and MT2), and a third cluster contains the west coast (OR and BC). Patterns of pairwise population differentiation using F_{ST} and R_{ST} are similar (Table S5, Supporting information). Using a Mantel test, we found a significant association between geographical and genetic distance ($r = 0.31$, $P = 0.026$) but not between SR frequency and genetic distance ($r = -0.09$, $P = 0.98$). For the SR data set, the most likely number of genetic clusters was two (Table S1, Supporting information), where the two clusters separate the west coast populations (OR and BC) from the rest of the geographical range (Fig. 2c). Pairwise estimates of population differentiation using F_{ST} and R_{ST} are highest between OR, BC and the rest of the range, similar to the pattern from the STRUCTURE analysis (Table S6, Supporting information). Using a Mantel test, we do not find a significant association between geographical and genetic distance ($r = 0.14$, $P = 0.17$) or between the SR frequency and genetic distance ($r = -0.34$, $P = 0.98$).

Contrasting patterns of X-linked and autosomal population differentiation

We asked whether patterns of geographical differentiation at the X-chromosome are similar to those of the autosomes. Previous research using five autosomal loci found generally little genetic differentiation among populations of *D. neotestacea* (Dyer 2012); of the clustering that was identified using STRUCTURE, the MT1, MT2 and ID populations formed one group, OR and PEI formed a second, and the rest of the range formed the main group. A pattern of more structure in the west is generally the pattern we see at the X-chromosome, especially once the effect of linkage with SR is removed; thus, we formally tested whether this pattern of differentiation is concordant with that of the autosomes using Mantel tests. We found that both the Random and ST data sets show a strong association between pairwise F_{ST} at the X-linked loci and the autosomes (Random: $r = 0.45$, $P = 0.009$; ST: $r = 0.43$, $P = 0.009$). In contrast, considering only populations that harbour SR, patterns of pairwise differentiation between the SR data set vs. the autosomes are not concordant ($r = 0.06$, $P = 0.38$), whereas the ST and SR data sets are ($r = 0.36$, $P = 0.043$). On balance, across the autosomes as well as the ST and SR X-chromosomes, there is much more genetic differentiation in the western part of the range of *D. neotestacea* than in the east, with less population differentiation among SR chromosomes. In general, however, similar to the autosomes, the population structure is weak; for example, no geographical clustering is found using STRUCTURE unless the sampling location is used as a prior, which was similar to the analyses of the autosomes (data not shown; Dyer 2012).

Effect of sex-ratio on X-linked diversity

The relatively high frequency of SR as well as the widespread structuring of variation between the ST and SR chromosomes suggests that SR may affect the general patterns of X-linked diversity in *D. neotestacea*. At the species-wide level, we found that the amount of genetic diversity in the ST data set was very similar to the Random sample (mean $H_{e\text{-}Random} = 0.736$ [95% CI 0.61–0.87], mean $H_{e\text{-}ST} = 0.702$ [95% CI 0.54–0.86]). In contrast, there was much lower diversity among the SR data set (mean $H_{e\text{-}SR} = 0.473$ [95% CI 0.28–0.66]), with SR individuals exhibiting a ~35% decrease in diversity relative to the Random data set. However, this is substantially more polymorphism than we would expect to see whether SR was at a long-term frequency equilibrium and the SR X-chromosome was genetically isolated from ST X-chromosomes, in which case we would expect the amount of relative heterozygosity of SR to be approximately equal to its frequency (i.e. 15%).

To investigate whether populations with a higher prevalence of SR may differ in the level of X-linked polymorphism compared with populations without SR, we asked whether the level of genetic diversity is correlated with the population prevalence of SR. We did not find a significant correlation between expected heterozygosity and local SR frequency for any data set (Fig. 4a; Autosomes: $r^2 = 0.26$; $F_{1,13} = 4.6$; $P = 0.05$; Random: $r^2 = 0.24$; $F_{1,13} = 4.1$; $P = 0.06$; ST: $r^2 = 0.2$; $F_{1,13} = 3.2$; $P = 0.10$; SR: $r^2 = 0.2$; $F_{1,9} = 2.6$; $P = 0.28$). We note that the diversity of Random data set is slightly higher than ST diversity in populations where SR is more common, with a slight but nonsignificant increase in diversity. This may be due to the combination of SR and ST diversity, as would be found with a Wahlund effect of genetically structured subpopulations (Wahlund 1928).

Under neutral equilibrium conditions and an equal population sex ratio, the X-chromosome is expected to harbour 75% of the polymorphism of the autosomes. When we compare the average heterozygosity of the X-chromosome to the autosomes at the species-wide level, the X/A ratio of the Random sample of X-chromosomes is 0.889 (95% CI 0.70–1.11), and the X/A ratio for the ST data

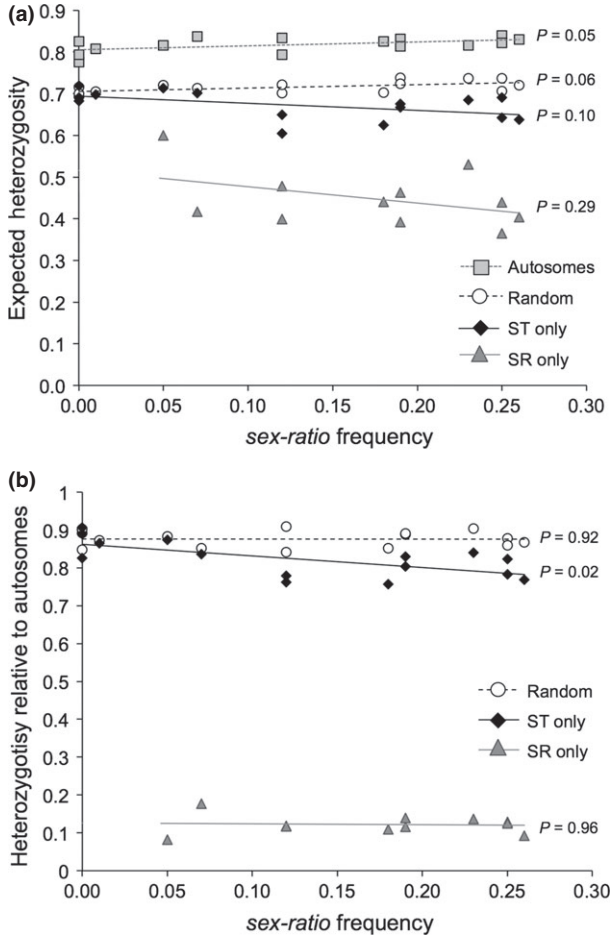

Fig. 4 The correlation of genetic diversity with the frequency of *sex-ratio* (SR). Shown here are (a) the total diversity for each X-chromosome data set and the autosomes and (b) the diversity of the X-chromosome relative to the autosomes. Only the *P*-value of each association is indicated here, see text for statistical details.

set is slightly less at 0.847 (95% CI 0.63–1.10). Both of these X/A ratios are higher than 0.75, though not significantly. Consistent with the overall patterns of diversity, the SR data set shows a lower X/A ratio of 0.571 (95% CI 0.33–0.83). However, because the species-wide prevalence of SR in *D. neotestacea* is only ~15%, this is nearly five times the ratio of 0.12 expected under neutrality (0.155*0.75).

Next, we asked whether the local prevalence of SR affects the local diversity of the X-chromosome relative to the autosomes. For each data set, we tested for an association between the local SR prevalence and the X/A ratio of expected heterozygosity. As seen in Fig. 4b, the X/A ratio of the ST data set decreased as the SR prevalence increased ($r^2 = 0.34$; $F_{1,13} = 6.8$; $P = 0.02$), but no effect was seen for the Random or SR data set (Random: $r^2 < 0.01$; $F_{1,13} = 0.01$; $P = 0.92$; SR: $r^2 < 0.01$; $F_{1,9} = 0.002$; $P = 0.96$). The decrease in ST diversity may result from its lower effective population size in populations where SR is more common. As with the total diversity, the combined diversity of the SR and ST X-chromosomes probably explains why there is no change in total population diversity (i.e. in the Random data set) as SR increases in prevalence.

The SR X-chromosome harbours more diversity than expected based on its prevalence; thus, we investigated whether it shows a signature of a different demographical history than the rest of the genome. We compared the observed heterozygosity with the heterozygosity expected based on the number of alleles under five different mutational models, and tested for an excess or deficiency of heterozygotes using two-tailed Wilcoxon sign rank tests (Table S7, Supporting information; Cornuet & Luikart 1996). For the autosomes, we did not find any deviation from mutation–drift equilibrium for any population, under any model. For the Random and ST X-chromosome data sets, the only evidence for population expansion was under the most conservative strict SMM model, except for in the ST data set where southernmost populations (NY and TN) both showed evidence of expansion (i.e. deficiency of heterozygosity) under all of the TPM and SMM models. In contrast, in the SR data set, every population showed a significant deficiency of heterozygosity relative to that expected by the number of observed alleles in the SMM and TPM model with 5% multistep changes, and four of the seven were significant with 20% multistep changes (Table S7, Supporting information). These results suggest that the demographical history of the SR chromosome is different from the ST and the autosomes, with a clear pattern of a recent population expansion.

Discussion

In this study, we investigate how intragenomic conflict over the offspring sex ratio affects the population genetic patterns of the X-chromosome in the fly *D. neotestacea*. Populations of *D. neotestacea* vary in the prevalence of SR drive, from 0 to 30% of males. The cline in SR frequency is associated with local temperature and appears to be relatively stable, at least over the short term (James & Jaenike 1990; Dyer 2012). Across 12 randomly chosen X-linked loci, we find strong but imperfect LD between the SR and ST chromosomes. This association, in combination with the relatively high frequency of SR in some populations, drives the contemporary patterns of population structure of the X-chromosome. In addition, we find a reduction in diversity within ST chromosomes relative to the autosomes where SR is common, and we find that

SR genetic diversity is reduced compared with ST diversity, but is higher than expected based on its prevalence. The overall amount of genetic diversity remains constant throughout the geographical range and is not dependent on SR frequency, probably because both ST and SR contribute to overall diversity in populations where SR is common.

The genetic differentiation between the SR and ST chromosomes is probably due to chromosomal inversions that suppress crossing over in heterozygous SR/ST females. Inversions are a common feature of drive systems, and reduced gene flow is found in many other inversion systems, not just those that contain meiotic drive factors (reviewed in Andolfatto *et al.* 2001; Kirkpatrick & Barton 2006; Hoffmann & Rieseberg 2008). The population genetic patterns we find here are in line with that expected in inversion systems where selection is strong (Guerrero *et al.* 2012). The LD in *D. neotestacea* appears to be more extensive than nearly all other characterized meiotic drive systems; however, genetic mapping and cytological data are necessary to indicate the exact map distance of the loci we surveyed and the extent to which the SR chromosome in *D. neotestacea* is tied up in chromosomal inversions. That we find no LD among loci within either the SR or ST chromosomes suggests the loci we surveyed may be spread out over the X-chromosome and are probably not physically clustered.

The SR and ST chromosomes do not appear to be completely isolated from each other, and our data suggest that there is occasional gene flow between the two chromosome types, for example via gene conversion or double crossovers within inversions. First, while the level of polymorphism at SR chromosomes is reduced relative to ST chromosomes, it is higher than we would expect based simply on its prevalence. Specifically, we see only a ~35% reduction in polymorphism compared with ST, whereas if there was no gene flow between SR and ST and if SR was at equilibrium, we would expect SR to have much less polymorphism—around 80–90% reduced based on its current species-wide frequency. This excess of polymorphism in SR is probably not due to SR occurring at a higher frequency in the past; if this was the case, we would expect to find lower diversity in the ST chromosomes, especially compared with the autosomes, and possibly the signature of a bottleneck in the ST. Assuming that there was one origin of SR in *D. neotestacea*, this excess of diversity in SR also suggests that there probably has not been a very recent sweep of a single SR haplotype, as has been found in other systems (e.g. Derome *et al.* 2004; Presgraves *et al.* 2009). In support of this, the SR chromosome shows a signature of a recent population expansion (i.e. a deficiency of heterozygosity relative to the observed number of alleles), whereas the autosomes and ST data sets do not. One potential scenario to account for these patterns is that the SR chromosome has experienced a population expansion after a sweep to the current SR chromosome, perhaps due to the incorporation of a new X-linked inversion into the system. A signature of an expansion can also be caused by an influx of rare alleles from genetically distinct immigrants (i.e. a Wahlund effect); in the case of SR, this signature may be due in part to strong differentiation from but occasional gene exchange with the ST chromosome. In other words, alleles from ST may become associated with SR and remain as rare alleles that elevate SR allele number to a much greater extent than heterozygosity; though, heterozygosity is still elevated relative to the neutral expectation. In support of this, at several loci, the rare SR alleles are common alleles in ST (Fig. S1, Supporting information). Because microsatellite genotypes are allele size classes, DNA sequence data will be necessary to infer genealogical relationships among haplotypes and thus to distinguish between novel but homoplasic mutations on the SR and recombination events with ST.

Even though there is extensive LD between the SR and ST chromosomes in *D. neotestacea*, the SR system is not akin to 'degenerating' drive systems with extensive inversions, such as the SR system in *D. recens* (Dyer *et al.* 2007). In *D. neotestacea*, females homozygous for SR are fully fertile and are found in the wild, and the lack of LD within SR suggests that the SR chromosome recombines with itself in SR/SR females (this study; James & Jaenike 1990; James 1992; Dyer 2012). In addition, as described above, there is probably also occasional gene flow between the SR and ST chromosome in SR/ST females. Together, this means that evolutionarily the SR chromosome in *D. neotestacea* has the potential to purge deleterious mutations and also to accumulate novel beneficial mutations, both through novel mutations on the SR chromosome and mutations obtained from the ST chromosome.

Owing to its inheritance pattern, the X-chromosome may differ from the autosomes in many aspects of its molecular evolution, including the mutation rate, recombination rate, migration rate, and the strength and efficacy of selection (reviewed in Schaffner 2004). In particular, deviations from the expected X/A ratio of polymorphism can indicate differences among populations in selection, demography or mating system (reviewed in Vicoso & Charlesworth 2006). In *D. neotestacea*, while we find reduced polymorphism on the X-chromosome relative to the autosomes, the X/A ratio is generally higher than expected under neutrality. This result has been found in many organisms (reviewed in Mank *et al.* 2009; Frankham 2011), including ancestral African populations of *Drosophila melanogaster*, where it has been investigated in depth (e.g. Andolfatto 2001;

Hutter *et al.* 2007; Singh *et al.* 2007). Potential explanations for a higher than expected X/A ratio include a population expansion (e.g. Hutter *et al.* 2007; Pool & Nielsen 2007), an excess of females in the population or high reproductive variance in male mating success (e.g. Caballero 1995; Charlesworth 2001), or differences in selection or recombination between the X-chromosome and autosomes (e.g. Betancourt *et al.* 2004; Vicoso & Charlesworth 2009; Charlesworth 2012). Interestingly, we find that the X_{ST}/A ratio is highest where SR is absent and decreases as SR becomes more common (Fig. 4b). One factor that may account for this pattern is a reduction in the effective population size of ST as SR increases in prevalence. It is also possible that other factors may contribute to this pattern; for example, there may be reduced variance in male reproductive rate among ST males where SR is at higher frequency, perhaps through decreased sperm competition because males are rarer. Further investigation with additional loci is necessary to investigate the underlying causes of both the general excess of polymorphism on the X-chromosome and also what factors may lower the X_{ST}/A ratio in populations where SR is at higher prevalence.

Based on differences in the strength and direction of selection, genetic drift and sex-based migration, the pattern of geographical differentiation can also differ among parts of the genome (reviewed in Holsinger & Weir 2009). This study highlights the consequences of genetic conflict for patterns of population genetic variation and differentiation in the nuclear genome. To date, perhaps the best-studied source of genetic conflict for population genetic structure has been that of selfish endosymbionts such as *Wolbachia* on the host mtDNA, which results because both the mtDNA and the endosymbiont are maternally inherited and thus genetically linked (reviewed in Ballard & Whitlock 2004; Galtier *et al.* 2009). In many arthropods, including *D. neotestacea* (Jaenike *et al.* 2010), the amount of genetic variation and the pattern of population structure at the mtDNA are dependent on how long ago the population was infected with the selfish endosymbiont and the particular phenotype it exerts on the host. This study highlights the effect genetic conflict can have on the nuclear genome, and like for endosymbionts, this is only uncovered when the presence and phenotype of the selfish element is known *a priori*. Without accounting for the genetic association with SR, the pattern of population differentiation at the X-chromosomes would be quite discordant with the autosomes, with the potential to outweigh any signature from the autosomes. Perhaps not surprisingly, in a STRUCTURE analysis that included both X-linked and autosomal loci, the signal of the SR association obscures any geographical signal, even when an equal number of X and autosomal loci were included (results not shown).

Our results support previous work in this system that suggested variation in SR prevalence across populations is due to local selection and not to geographical structure or to segregating suppressors of drive (Dyer 2012). First, similar to the autosomes, at the X-chromosome, we find substantial gene flow across populations that differ in frequency of SR. Second, previous work searched for autosomal dominant and Y-linked suppressors of drive in a few populations (Dyer 2012), and the results here extend this to the entire species range and also suggest that there are no recessive autosomal suppressors of drive segregating. If there were suppressors segregating, we would expect to see ST males with an SR haplotype, but we do not (Fig. 1a). This does not mean that there are not fixed suppressors or modifiers of drive; in fact, we would expect these to be present given the extensive LD between the SR and ST chromosome, as this suggests that the SR system in *D. neotestacea* is not as evolutionary recent as originally thought (James & Jaenike 1990). Finally, across populations, the frequency of SR is correlated with the local temperature, with warmer populations harbouring a higher prevalence of drive (Dyer 2012). Our current results indicate that most of the segregating X-chromosome polymorphism, not just a small region around SR, is structured on a temperature gradient. This raises the possibility that variation in SR prevalence may be due to local selection on the SR system or to a different trait(s) where alleles with different ecological tolerances are also differentiated between the two types of chromosomes. Based on empirical work in other systems, the potential for selection to act on inversions that do not contain meiotic drive factors can also be very strong (Huey *et al.* 2000; Balanyà *et al.* 2006). However, we suggest that the transmission advantage of SR is so strong that selection is probably acting on SR to result in local differences in SR frequency, rather than on another trait, but this has yet to be investigated empirically.

In summary, our results suggest that genetic conflict has played a significant role in shaping the evolution of the X-chromosome of *D. neotestacea*. The extensive genetic association with SR appears to extend over a large genetic distance and is probably due to the gradual accumulation of X-linked modifiers that must be in linkage to result in the full expression of drive. Further work is necessary to obtain a more complete understanding of the evolutionary history of the SR chromosome and its molecular evolutionary consequences in *D. neotestacea*. First, genetic mapping and cytological data will indicate the extent to which the X-chromosome is tied up in chromosomal inversions. Second, an increased number of both X-linked and autosomal loci will indicate whether the patterns we observe

here hold true at the broader genomic scale. In addition, DNA sequence data will allow us to more accurately infer the age of the SR system, including distinguishing which mutations are unique to the SR chromosome or are a result of gene exchange between the SR and ST chromosomes. Finally, it will be interesting to investigate patterns of molecular evolution to ask whether SR has affected the rate of protein evolution and the efficacy of selection, as might be expected because of the reduced effective population sizes of the SR and ST chromosomes.

Acknowledgements

We are grateful to David Hall, Cheryl Pinzone and three anonymous reviewers for comments on the manuscript, and the Ellison Medical Foundation and National Science Foundation for funding.

References

Andolfatto P (2001) Contrasting patterns of *X*-linked and autosomal nucleotide variation in African and non-African populations of *Drosophila melanogaster* and *D. simulans*. *Molecular Biology and Evolution*, **18**, 279–290.

Andolfatto P, Depaulis F, Navarro A (2001) Inversion polymorphisms and nucleotide variability in *Drosophila*. *Genetical Research*, **77**, 1–8.

Balanyà J, Oller JM, Huey RB, Gilchrist GW, Serra L (2006) Global genetic change tracks global climate warming in *Drosophila subobscura*. *Science*, **313**, 1773–1775.

Ballard JWO, Whitlock MC (2004) The incomplete natural history of mitochondria. *Molecular Ecology*, **13**, 729–744.

Bastide H, Cazemajor M, Ogereau D, Derome N, Hospital F, Montchamp-Moreau C (2011) Rapid rise and fall of selfish *sex-ratio* X chromosomes in *Drosophila simulans*: spatio-temporal analysis of phenotypic and molecular data. *Molecular Biology and Evolution*, **28**, 2461–2470.

Betancourt AJ, Kim Y, Orr HA (2004) A pseudohitchhiking model of X vs. autosomal diversity. *Genetics*, **168**, 2261–2269.

Burt A, Trivers R (2006) Genes in Conflict: The Biology of Selfish Genetic Elements. Harvard University Press, Cambridge.

Caballero A (1995) On the effective size of populations with separate sexes, with particular reference to sex-linked genes. *Genetics*, **139**, 1007–1011.

Charlesworth B (2001) The effect of life-history and mode of inheritance on neutral genetic variability. *Genetical Research*, **77**, 153–166.

Charlesworth B (2012) The role of background selection in shaping patterns of molecular evolution and variation: evidence from variability on the Drosophila X chromosome. *Genetics*, **191**, 233–246.

Christianson SJ, Brand CL, Wilkinson GS (2011) Reduced polymorphism associated with X chromosome meiotic drive in the stalk-eyed fly *Teleopsis dalmanni*. *PLoS ONE*, **6**, e27254.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001–2014.

Derome N, Metayer K, Montchamp-Moreau C (2004) Signature of selective sweep associated with the evolution of *sex-ratio* drive in *Drosophila simulans*. *Genetics*, **166**, 1357–1366.

Di Rienzo A, Peterson AC, Garza JC, Valdes AM, Slatkin M, Freimer NB (1994) Mutational processes of simple-sequence repeat loci in human populations. *Proceedings of the National Academy of Sciences USA*, **91**, 3166–3170.

Dyer KA (2007) Identification and characterization of 21 polymorphic microsatellite loci from the mycophagous fly *Drosophila neotestacea*. *Molecular Ecology Notes*, **7**, 1120–1122.

Dyer KA (2012) Local selection underlies the geographic distribution of *sex-ratio* drive in *Drosophila neotestacea*. *Evolution*, **66**, 973–984.

Dyer KA, Charlesworth B, Jaenike J (2007) Chromosome-wide linkage disequilibrium as a consequence of meiotic drive. *Proceedings of the National Academy of Sciences USA*, **104**, 1587–1592.

Dyer KA, White BE, Bray MJ, Piqué DG, Betancourt AJ (2011) Molecular evolution of a Y to autosome translocation. *Molecular Biology and Evolution*, **28**, 1293–1306.

Estoup A, Jarne P, Cornuet JM (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetic analysis. *Molecular Ecology*, **11**, 1591–1604.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software Structure: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Fouvry L, Ogereau D, Berger A, Gavory F, Montchamp-Moreau C (2011) Sequence analysis of the segmental duplication responsible for Paris *sex-ratio* drive in *Drosophila simulans*. *G3 (Bethesda, Md.)*, **1**, 401–410.

Frankham R (2011) How closely does genetic diversity in finite populations conform to predictions of neutral theory? Large deficits in regions of low recombination. *Heredity*, **108**, 1–12.

Galtier N, Nabholz B, Glemin S, Hurst GDD (2009) Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Molecular Ecology*, **18**, 4541–4550.

Guerrero RF, Rousset F, Kirkpatrick M (2012) Coalescent patterns for chromosomal inversions in divergent populations. *Philosophical Transactions of the Royal Society Series B*, **367**, 430–438.

Hamilton WD (1967) Extraordinary sex ratios. *Science*, **156**, 477–488.

Hoffmann AA, Rieseberg L (2008) Revisiting the impact of inversions in evolution: from population genetic markers to drivers of adaptive shifts and speciation. *Annual Reviews of Ecology, Evolution, and Systematics*, **63**, 21–42.

Holsinger KE, Weir BS (2009) Genetics in geographically structured populations: defining, estimating, and interpreting F_{ST}. *Nature Reviews Genetics*, **10**, 639–650.

Huey RB, Gilchrist GW, Carlson ML, Berrigan D, Serra L (2000) Rapid evolution of a geographic cline in size in an introduced fly. *Science*, **287**, 308–309.

Hutter S, Li HP, Beisswanger S, De Lorenzo D, Stephan W (2007) Distinctly different sex ratios in African and European populations of *Drosophila melanogaster* inferred from chromosome-wide nucleotide polymorphism data. *Genetics*, **177**, 469–480.

Jaenike J (2001) Sex chromosome meiotic drive. *Annual Reviews of Ecology, Evolution, and Systematics*, **32**, 25–49.
Jaenike J, Unckless R, Cockburn SN, Boelio LM, Perlman SJ (2010) Adaptation via symbiosis: recent spread of a *Drosophila* defensive symbiont. *Science*, **329**, 212–215.
James AC (1992) "Sex-ratio" Meiotic Drive in *Drosophila neotestacea*. PhD dissertation, University of Rochester, Rochester, New York, USA.
James AC, Jaenike J (1990) "Sex-ratio" meiotic drive in *Drosophila testacea*. *Genetics*, **125**, 651–656.
Kaiser V, Bachtrog D (2010) Evolution of sex chromosomes in insects. *Annual Review of Genetics*, **44**, 91–112.
Kirkpatrick M, Barton N (2006) Chromosome inversions, local adaptation, and speciation. *Genetics*, **173**, 419–434.
Kovacevic M, Schaeffer SW (2000) Molecular population genetics of *X*-linked genes in *Drosophila pseudoobscura*. *Genetics*, **156**, 155–172.
Mank JE, Vicoso B, Berlin S, Charlesworth B (2009) Effective population size and the faster-X effect: empirical results and their interpretation. *Evolution*, **64**, 663–674.
Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**, 209–220.
Meiklejohn CD, Tao Y (2010) Genetic conflict and sex chromosome evolution. *Trends in Ecology and Evolution*, **25**, 215–223.
Montchamp-Moreau C, Ogereau D, Chamnade N, Colard A, Aulard S (2006) Organization of the *sex-ratio* meiotic drive region in *Drosophila simulans*. *Genetics*, **174**, 1365–1371.
Nei M, Maruyama T, Chakraborty R (1975) Bottleneck effect and genetic-variability in populations. *Evolution*, **29**, 1–10.
Patterson JT, Stone WS (1952) Evolution in the Genus *Drosophila*. The Macmillan Company, New York. 618 pp.
Piry S, Luikart G, Cornuet J-M (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity*, **90**, 502–503.
Pool JE, Nielsen R (2007) Population size changes reshape genomic patterns of diversity. *Evolution*, **61**, 3001–3006.
Presgraves DC (2010) The molecular evolutionary basis of species formation. *Nature Reviews Genetics*, **3**, 175–180.
Presgraves DC, Gérard PR, Cherukuri A, Lyttle TW (2009) Large-scale selective sweep among *Segregation Distorter* chromosomes in African populations of *Drosophila melanogaster*. *PLoS Genetics*, **5**, e1000463.
Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
Räisänen L, Roininen E, Liimatainen JO (2009) Characterization of novel microsatellites from *Drosophila transversa*. *Molecular Ecology Resources*, **9**, 645–647.
Raymond M, Rousset F (1995) Genepop, population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.
Rousset F (2008) Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103–106.
Schaffner SF (2004) The X chromosome in population genetics. *Nature Reviews Genetics*, **5**, 43–51.
Singh ND, Macpherson JM, Jensen JD, Petrov DA (2007) Similar levels of *X*-linked and autosomal nucleotide variation in African and non-African populations of *Drosophila melanogaster*. *BMC Evolutionary Biology*, **7**, 202.
Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, **139**, 457–462.
Vicoso B, Charlesworth B (2006) Evolution on the X chromosome: unusual patterns and processes. *Nature Reviews Genetics*, **7**, 645–653.
Vicoso B, Charlesworth B (2009) Recombination rates may affect the ratio of *X* to autosomal noncoding polymorphism in African populations of *Drosophila melanogaster*. *Genetics*, **181**, 1699–1701.
Wahlund S (1928) Zusammensetzung von Population und Korrelationserscheinung vom Standpunkt der Vererbungslehre aus betrachtet. *Hereditas*, **11**, 65–106.
Werren JH, Nur U, Wu C-I (1988) Selfish genetic elements. *Trends in Ecology and Evolution*, **3**, 297–302.

K.A.D. studies the population genetics and genetics of adaptation of speciation in Drosophila. She is especially interested in the consequences of genomic conflict. M.J.B. and S.J.L. are interested in all things genetics, and plan to study biomedical genetics in graduate school.

Data accessibility

Microsatellite data has been submitted to Dryad (doi:10.5061/dryad.2d315).

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Results of the Structure analyses to infer the number of genetic clusters (K).

Table S2 Proportion of membership of each population into each cluster, as defined by the Structure analyses on each dataset.

Table S3 Variability measures for each of the six microsatellites within each population and across all populations.

Table S4 Pairwise population differentiation at 12 X-linked microsatellite loci.

Table S5 Pairwise population differentiation at 12 X-linked microsatellite loci, using ST males and females inferred to be ST/ST.

Table S6 Pairwise population differentiation at 12 X-linked microsatellite loci, using SR males and females inferred to be SR/SR.

Table S7 Demographic inference based on the BOTTLENECK analyses.

Fig. S1 Histograms of allele sizes of X-linked microsatellite markers, including only males with a known sex-ratio phenotype based on offspring sex ratio.

Fig. S2 Additional Structure results.