

1 Article: Discoveries

2 **Sequence and structural diversity of mouse Y chromosomes**

3 Andrew P Morgan and Fernando Pardo-Manuel de Villena*

4 Department of Genetics and Lineberger Comprehensive Cancer Center
5 University of North Carolina, Chapel Hill, NC

6 *Corresponding author:

7 5049C Genetic Medicine Building

8 Department of Genetics

9 University of North Carolina

10 Chapel Hill NC 27599-7264

11 fernando@med.unc.edu

12 **Keywords:** sex chromosome evolution, intragenomic conflict, mouse evolution

14

Abstract

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

Over the 180 million years since their origin, the sex chromosomes of mammals have evolved a gene repertoire highly specialized for function in the male germline. The mouse Y chromosome is unique among mammalian Y chromosomes characterized to date in that it is large, gene-rich and euchromatic. Yet little is known about its diversity in natural populations. Here we take advantage of published whole-genome sequencing data to survey the diversity of sequence and copy number of sex-linked genes in three subspecies of house mice. Copy number of genes on the repetitive long arm of both sex chromosomes is highly variable, but sequence diversity in non-repetitive regions is decreased relative to expectations based on autosomes. We use simulations and theory to show that this reduction in sex-linked diversity is incompatible with neutral demographic processes alone, but is consistent with recent positive selection on genes active during spermatogenesis. Our results support the hypothesis that the mouse sex chromosomes are engaged in ongoing intragenomic conflict.

Introduction

Sex chromosomes have emerged many times in independent plant and animal lineages. The placental mammals share a sex chromosome pair that originated approximately 180 million years ago (Mya) (Hughes and Page 2015). In the vast majority of mammal species, the Y chromosome is sex-determining: presence of the Y-encoded protein SRY is sufficient to initiate the male developmental program (Berta et al. 1990). Since their divergence from the ancestral X chromosome, mammalian Y chromosomes have lost nearly all of their ancestral gene content (**Figure 1A**). Although these losses have occurred independently along different lineages within the mammals, the small subset of genes that are retained in each lineage tend to be dosage-sensitive and have housekeeping functions in core cellular processes such as transcription and protein degradation (Bellott et al. 2014; Cortez et al. 2014). Contrary to bold predictions that the mammalian Y chromosome is bound for extinction (Graves 2006), empirical studies of Y chromosomes have demonstrated that most gene loss occurs in early proto-sex chromosomes, and that the relatively old sex chromosomes of mammals are more stable (Bellott et al. 2014). The evolutionary diversity of Y chromosomes in mammals arises from the set of Y-acquired genes, which make up a small fraction of some Y chromosomes and a much larger fraction in others — from 5% in rhesus to 45% in human (Hughes and Page 2015) (**Figure 1B**). These genes are often present in many copies and are highly specialized for function in the male germline (Lahn and Page 1997; Soh et al. 2014).

The Y chromosome of the house mouse (*Mus musculus*) stands out among mammalian Y chromosomes both for its sheer size and its unusual gene repertoire. Early molecular studies of the mouse Y chromosome

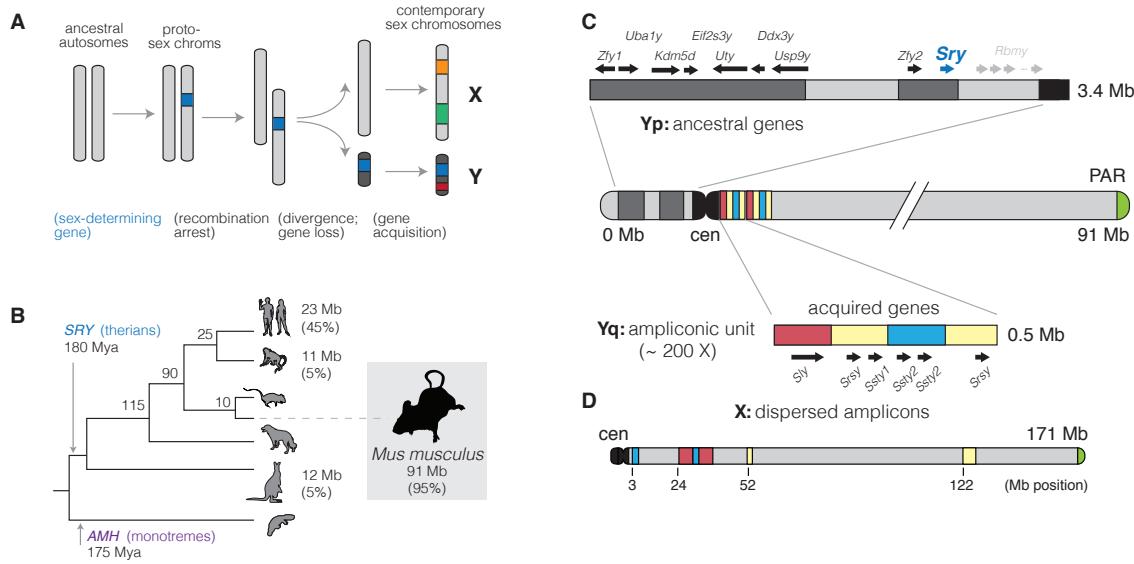


Figure 1: Evolution of mammalian Y chromosomes. **(A)** Evolution of heteromorphic sex chromosomes. **(B)** Y chromosomes of mammals. The Y chromosome of therian mammals, characterized by the sex-determining factor *SRY*, diverged from the X chromosome approximately 180 Mya. (The monotremata have a different sex-determining factor, *AMH*, and an idiosyncratic five-pair sex chromosome system.) Y chromosome sizes and the fraction of sequence occupied by multicopy, Y-acquired genes are shown at the tips of the tree. **(C)** Structure of the Y chromosome in the C57BL/6J reference strain. The short arm of the Y chromosome (Yq) consists primarily of genes shared with the X chromosome and retained since the sex chromosomes diverged from the ancestral autosome pair. These genes are interspersed with blocks of segmental duplications (light grey). The sex-determining factor *Sry* is encoded on the short arm. The long arm (Yq) consists of approximately 200 copies of a 500 kb repeating unit containing the acquired genes *Sly*, *Ssty1*, *Ssty2* and *Srsy*. The sequence in the repeat unit can be roughly divided into three families “red,” “yellow” and “blue” following (Soh et al. 2014). **(D)** The X chromosome, unlike the Y chromosome, is acrocentric. Homologs of the acquired genes on the Y chromosome (*Slx*, *Slxl1*, *Sstx* and *Srsx*; shown above using colored blocks as on the Y) are present in high copy number but are arranged in tandem chunks, rather than intermingled as on the Y.

46 hinted that it consisted of mostly of repetitive sequences, with copy number in the hundreds, and that it was
 47 evolving rapidly (Nishioka and Lamothe 1986; Eicher et al. 1989). Unlike other mammalian Y chromosomes,
 48 which are dominated by large blocks of heterochromatin (Hughes and Page 2015), the mouse Y chromosome
 49 was also known to be large and almost entirely euchromatic. Spontaneous mutations in laboratory stocks
 50 allowed the mapping of male-specific tissue antigens and the sex-determining factor *Sry* to the short arm of the
 51 chromosome (Yp) (McLaren et al. 1988), while lesions on the long arm (Yq) were associated with infertility and
 52 defects in spermatogenesis (Styrna et al. 1991; Burgoyne et al. 1992; Touré et al. 2004).

53 Sequencing, assembly and annotation of the mouse Y chromosome in the inbred strain C57BL/6J was finally
 54 completed in 2014 after more than a decade of painstaking effort (Soh et al. 2014). Ancestral genes are restricted
 55 to Yp and are fewer in number on the mouse Y chromosome than in other studied mammals. Yq was shown

56 to consist of approximately 200 copies of a 500 kb unit — the “huge repeat array” — containing the acquired
57 genes *Sly*, *Ssty1*, *Ssty2* and *Srsy* (**Figure 1C**). *Sly* and its X-linked homologs *Slx* and *Slxl1* are found only in
58 the genus *Mus* and have sequence similarity to the synaptonemal complex protein SYCP3 (Ellis et al. 2011).
59 *Ssty1/2* and *Sstx* are most similar to members of the spindlin family (Oh et al. 1997) and are present in taxa
60 at least as phylogenetically distant as rats. The coding potential of *Srsy* and *Srsx* is unclear, but they have
61 sequence similarity to melanoma-related cancer/testis antigens typified by the human MAGEA family. Their
62 phylogenetic origins remain unresolved. The genes of the huge repeat array are expressed almost exclusively
63 in post-meiotic round spermatids and function in chromatin condensation and sperm maturation (Burgoyne
64 et al. 1992; Touré et al. 2004, 2005; Yamauchi et al. 2009, 2010).

65 Independent amplification of homologous genes on the X and Y chromosomes is thought to be a byproduct
66 of competition between the X and Y chromosomes for transmission to the next generation. The current con-
67 sensus favors an unidentified X-linked sex-ratio distorter whose action is suppressed by one or more Y-linked
68 factors (Ellis et al. 2011). Consistent with this hypothesis, SLY acts directly to maintain transcriptional repres-
69 sion of post-meiotic sex chromatin (PSCR, Hendriksen et al. (1995)) by recruiting a suite of repressive histone
70 marks (Ellis et al. 2005; Cocquet et al. 2009; Moretti et al. 2016); its action is opposed by SLX and SLXL1. Imbal-
71 ance between SLY and SLX/SLXL1 tilts the progeny sex ratio in favor of the overexpressing chromosome and
72 causes defects in sperm morphology and sperm count (Touré et al. 2004; Cocquet et al. 2009, 2010). Disruption
73 of PSCR and the related process of meiotic sex chromosome inactivation (MSCI) is also associated with male
74 sterility in inter-subspecific hybrids between *M. m. domesticus* and *M. m. musculus* (Good et al. 2010; Campbell
75 et al. 2013; Bhattacharyya et al. 2013; Larson et al. 2016a). Together these observations suggest that the intrage-
76 nomic conflict between the sex chromosomes in mouse is played out in post-meiotic spermatids and may have
77 mechanistic overlap with hybrid male sterility.

78 Intragenomic conflict can have a profound impact on the genetic diversity of sex chromosomes in natural
79 populations. Sex-ratio-distorter systems in *Drosophila* provide some of the best-known examples (Jaenike 2001;
80 Derome et al. 2004; Kingan et al. 2010). The extent to which diversity on mouse sex chromosomes is influenced
81 by intragenomic conflict remains an open question. The differential impact of selection on mouse X chromo-
82 some versus autosomes (the “faster-X” effect) is well-studied, mostly through the lens of speciation (Torgerson
83 and Singh 2003; Kousathanas et al. 2014; Larson et al. 2016a,b). Larson et al. (2016b) used pairwise comparisons
84 between wild-derived strains of *M. m. musculus* and *M. m. domesticus* to show that the “faster-X” effect is most
85 prominent in two groups of genes: those expressed primarily in the testis and early in spermatogenesis (before
86 MSCI), and those up-regulated in spermatids (after PSCR). The former set of genes is also prone to aberrant
87 expression in sterile hybrids Larson et al. (2016a). By contrast, selective pressures imposed by intragenomic
88 conflict between the sex chromosomes should be exerted in spermatids after the onset of PSCR. Genes with

89 spermatid-specific expression are expected to respond most rapidly, while those with broad expression are
90 expected to be constrained by putative functional requirements in other tissues or cell types.

91 In this manuscript we take advantage of the relatively recent high-quality assembly of the mouse Y chro-
92 mosome (Soh et al. 2014) and public sequencing data from a diverse sample of wild mice to perform a survey
93 of sequence and copy-number diversity on the sex chromosomes. We use complementary gene-expression
94 data and annotations to partition the analysis into functionally-coherent groups of loci. We find that sequence
95 diversity is markedly reduced on both the X and Y chromosomes relative to expectations for a stationary pop-
96 ulation. This reduction cannot be fully explained by any of several demographic models fit to autosomal data,
97 but Y-linked diversity in *M. m. domesticus* is consistent with a recent selective sweep on Y chromosomes. Copy
98 number of genes expressed in spermatids supports the hypothesis that intragenomic conflict between the sex
99 chromosomes during spermiogenesis is an important selective pressure. These analyses broaden our under-
100 standing of the evolution of sex chromosomes in murid rodents and support an important role for positive
101 selection in the male germline.

102 Results

103 A survey of Y-linked coding variation in mouse

104 Whole-genome or whole-exome sequence data for 91 male mice was collected from published sources (Keane
105 et al. 2011; Doran et al. 2016; Harr et al. 2016; Morgan et al. 2016a; Neme and Tautz 2016; Sarver et al. 2017).
106 The final set consists of 62 wild-caught mice; 21 classical inbred strains; and 8 wild-derived inbred strains
107 (**Table 1** and **Table S1**). The three cardinal subspecies of *M. musculus* (*domesticus*, *musculus* and *castaneus*) are all
108 represented, with *Mus spretus* and *Mus spicilegus* as close outgroups and *Mus caroli*, *Mus cookii*, and *Nannomys*
109 *minutoides* as more distant outgroups. Our sample spans the native geographic range of the house mouse and
110 its sister taxa (**Figure 2A**).

111 Single-nucleotide variants (SNVs) and small indels were ascertained in 41.6 kb of sequence on Yp targeted
112 by the Roche NimbleGen exome-capture array. To mitigate the effect of alignment errors and cryptic copy-
113 number variation on our analyses, we discarded sites with evidence heterozygosity; fewer than 60 samples
114 with a called genotype; or evidence of strand bias (see **Materials and methods**). In total we identified 1,136
115 SNVs and 128 indels, with transition:transversion ratio 2.1.

116 One group of inbred strains in our dataset — C57BL/6J (reference genome), C57BL/10J, C57L/J and C57BR/cdJ
117 — have a known common ancestor in the year 1929, and a common ancestor with the strain C58/J in 1915
118 (Beck et al. 2000). Assuming an average of three generations per year, the total branch length of the pedi-
119 gree connecting the C57 and C58 strains is 5,280 generations, during which time 3 mutations occurred. We

Type	Population	Country	Males	Females
wild	<i>M. m. domesticus</i>	DE	8	1
		FR	7	0
		IR	5	0
	<i>M. m. musculus</i>	AF	5	1
		CZ	2	5
		KZ	2	4
	<i>M. m. castaneus</i>	IN	3	7
	<i>M. spretus</i>	ES	4	2
		MA	1	0
	<i>M. macedonicus</i>	MK	1	0
wild-derived	<i>M. spicilegus</i>	HU	1	0
	<i>M. caroli</i>	TH	0	1
	<i>M. cookii</i>	TH	1	0
	<i>Nannomys minutoides</i>	KE	1	0
	<i>M. m. domesticus</i>	IT	1	0
classical lab		US	1	1
	<i>M. m. musculus</i>	CZ	1	1
	<i>M. m. castaneus</i>	TH	1	1
	<i>M. spretus</i>	ES	1	0
	-	-	21	4

Table 1: Wild and laboratory mice used in this study.

used these values to obtain a direct estimate of the male-specific point mutation rate: 1.8×10^{-8} (95% Poisson CI $4.5 \times 10^{-9} - 4.7 \times 10^{-8}$) bp $^{-1}$ generation $^{-1}$. This interval contains the sex-averaged autosomal rate of 5.4×10^{-9} bp $^{-1}$ generation $^{-1}$ recently estimated from whole-genome sequencing of mutation-accumulation lines (Uchimura et al. 2015). Using the ratio between paternal to maternal mutations in mouse estimated in classic studies from Russell and colleagues (2.78; reviewed in Drost and Lee (1995)), that estimate corresponds to male-specific autosomal rate of 7.9×10^{-9} bp $^{-1}$ generation $^{-1}$, again within our confidence interval. We note that these estimates assume that selection has been negligible in laboratory colonies.

Phylogeny of Y chromosomes recovers geographic relationships

Phylogenetic trees for exonic regions of the Y chromosome and mitochondrial genome were constructed with BEAST (**Figure 2B**). The estimated time to most recent common ancestor (MRCA) of *M. musculus* Y chromosomes is 900,000 years ago (95% highest posterior density interval [HPDI] 100,000 – 1,800,000) years ago. Within *M. musculus*, the *domesticus* subspecies diverges first, although the internal branch separating it from the MRCA of *musculus* and *castaneus* is very short. Consistent with several previous studies, we find that the “old” classical inbred strains share a single Y haplogroup within *M. m. musculus*. This haplogroup is distinct from that of European and central Asian wild mice and is probably of east Asian origin (Bishop et al. 1985; Tucker et al. 1992; Nagamine et al. 1992). Strains related to “Swiss” outbred stocks (FVB/NJ, NOD/ShiLtJ, HR8) and those of less certain American origin (AKR/J, BUB/BnJ) (Beck et al. 2000) have Y chromosomes with affinity

137 to western European populations. *M. m. castaneus* harbors two distinct and paraphyletic lineages: one corre-
138 sponding to the Indian subcontinent and another represented only by the wild-derived inbred strain CAST/EiJ
139 (from Thailand). The latter haplogroup corresponds to a southeast Asian lineage identified in previous reports
140 that sampled more extensively from that geographic region (Geraldes et al. 2008; Yang et al. 2011). It remains
141 unclear whether this haplogroup originated in *M. m. musculus* and displaced the *M. m. castaneus* Y chromo-
142 some in southeast Asia; or instead represents a deep branching within the (large and unsampled) population
143 ancestral to *musculus* and *castaneus* in central Asia.

144 The Y-chromosome tree otherwise shows perfect concordance between clades and geographic locations.
145 Within the *M. m. domesticus* lineage we can recognize two distinct haplogroups corresponding roughly to west-
146 ern Europe and Iran and the Mediterranean basin, respectively. Similarly, within *M. m. musculus*, the eastern
147 European mice (from Bavaria, Czech Republic) are well-separated from the central Asian mice (Kazakhstan and
148 Afghanistan). Relationships between geographic origins and phylogenetic affinity are considerably looser for
149 the mitochondrial genome. We even found evidence for inter-specific hybridization: one nominally *M. spretus*
150 individual from central Spain (SP36) carries a *M. spretus* Y chromosome and a *M. m. domesticus* mitochondrial
151 genome (arrowhead in **Figure 2B**). Several previous studies have found evidence for introgression between *M.*
152 *musculus* and *M. spretus* where their geographic ranges overlap (Orth et al. 2002; Song et al. 2011; Liu et al. 2015).

153 Copy-number variation is pervasive on the Y chromosome

154 We examined copy number along Yp using depth of coverage. Approximately 779 kb (24%) of Yp consists
155 of segmental duplications or gaps in the reference assembly (**Figure 1**); for duplicated regions we scaled the
156 normalized read depth by the genomic copy number in the reference sequence to arrive at a final copy-number
157 estimate for each individual. All of the known duplications on Yp are polymorphic in laboratory and natural
158 populations (**Figure 3A**). The distribution of CNV alleles follows the SNV-based phylogenetic tree. Only one
159 CNV region on Yp, adjacent to the centromere, contains a known protein-coding gene (*Rbmy*). Consistent with
160 a previous report (Ellis et al. 2011), we find that *musculus* Y chromosomes have more copies of *Rbmy* than
161 *domesticus* or *castaneus* chromosomes.

162 The highly repetitive content of Yq precludes a similarly detailed characterization of copy-number variation
163 along this chromosome arm. However, we can estimate the copy number of each of the three gene families
164 present (*Sly*, *Ssty1/2* and *Srsy*) by counting the total number of reads mapped to each and normalizing for
165 sequencing depth. The hypothesis of X-Y intragenomic conflict predicts that, if expression levels are at least
166 roughly proportional to copy number, amplification of gene families on Yq should be countered by amplifica-
167 tion of their antagonistic homologs on the X chromosome (or vice versa.) We tested this hypothesis by compar-
168 ing the copy number of X- and Y-linked homologs of the *Slx/y*, *Sstx/y* and *Srsx/y* families in wild mice. **Figure 4**

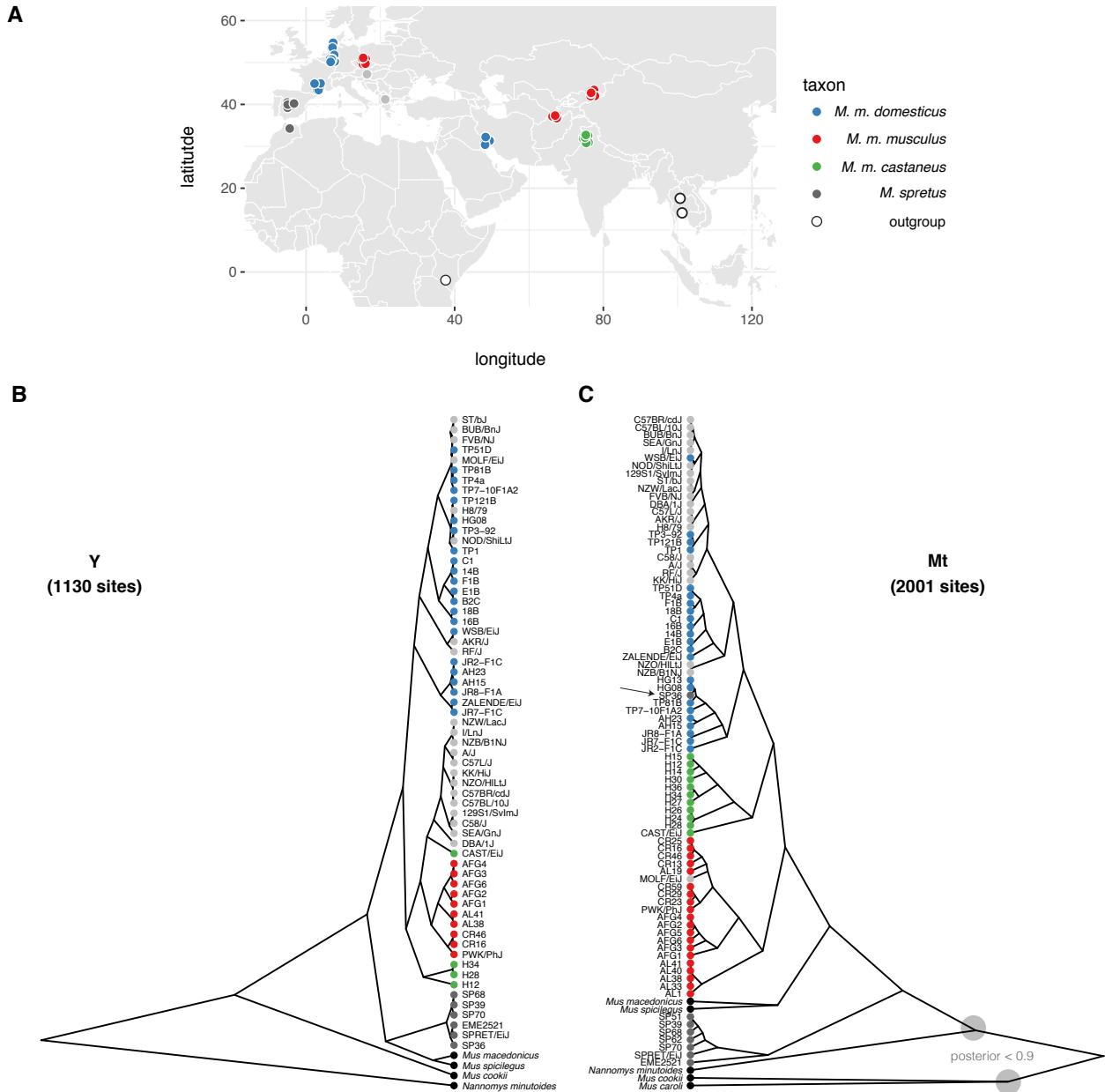


Figure 2: Patrilineal and matrilineal phylogeography in a geographically-diverse sample from the genus *Mus*. (A) Sampling locations of mice used in this study. (B) Phylogenetic tree from coding sites on the Y chromosome. Samples are colored according to nominal ancestry; laboratory strains are shown in light grey. (C) Phylogenetic tree from coding sites on the mitochondrial genome. Deep nodes with posterior support < 0.9 indicated with shaded circles.

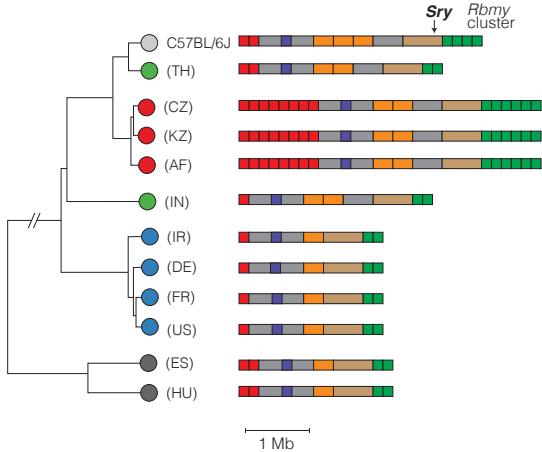


Figure 3: Copy-number variation on Yp. Schematic view of copy-number variable regions of the Y chromosome short arm (Yp) superposed on SNV-based phylogenetic tree. All CNVs shown overlap a segmental duplication in the reference sequence (strain C57BL/6J). One CNV overlaps a known protein-coding gene: an expansion of the ampliconic *Rbmy* cluster (green) in *M. m. musculus*. Color scheme for *Mus* taxa follows **Figure 2**.

169 shows that copy number on X and Y chromosomes are indeed correlated for *Slx/y*. The relationship between
170 *Slx*-family and *Sly*-family copy number is almost exactly linear (slope = 0.98 [95% CI 0.87 – 1.09]; $R^2 = 0.87$).
171 We note that samples are not phylogenetically independent, so the statistical significance of the regression is
172 exaggerated, but the qualitative result clearly supports previous evidence that conflict between X and Y chro-
173 mosomes is mediated primarily through *Slx* and *Sly* (Cocquet et al. 2012). Size differences estimated from *Sly*
174 copy number are also concordant with cytological observations that the Y chromosomes of wild-caught *M. m.*
175 *musculus* appear much larger than those of *M. spicilegus* or *M. spretus* (Bulatova and Kotenkova 1990; Yakimenko
176 et al. 1990).

177 It has recently been shown that two regions of the autosomes — on chromosomes 5 and 14 — have a suite
178 of epigenetic marks similar to the sex chromosomes in post-meiotic spermatids (Moretti et al. 2016). These
179 autosomal regions harbor many copies of a family of genes (known alternatively as *Speer* (Spiess et al. 2003)
180 or α -takusan (Tu et al. 2007)) expressed in spermatids. The copy number of *Speer* family members is, like *Sly*,
181 correlated with that of *Slx/Slxl1* (**Figure S1**). This finding supports the hypothesis that the *Speer* family may be
182 involved in sex-chromosome conflict in spermatids.

183 The scale of copy number change within the *M. musculus* lineage suggests a high underlying mutation rate.
184 We used whole-genome sequence data from a panel of 69 recombinant inbred lines (RILs) from the Collabora-
185 tive Cross (CC; Srivastava et al. (2017)) to estimate the rate of copy-number change on Yq. Each CC line is
186 independently derived from eight inbred founder strains via two generations of outcrossing followed by sib-
187 ling matings until inbreeding is achieved (Consortium 2012). Distinct CC lines inheriting a Y chromosome from
188 the same founder strain thus share an obligate male ancestor in the recent past, but no more recently than the

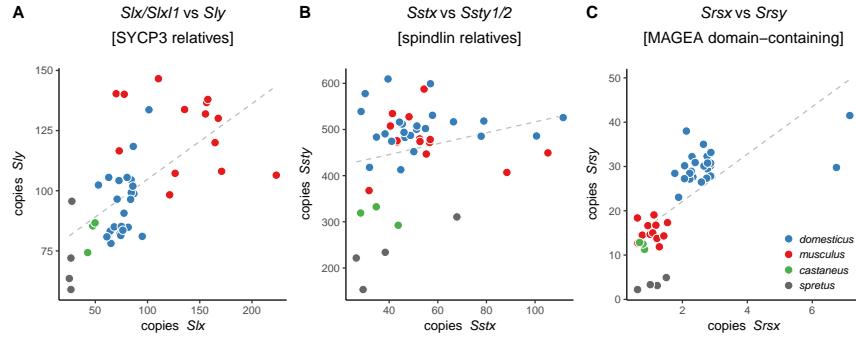


Figure 4: Approximate copy number of co-amplified gene families on X and Yq. Each dot represents a single individual. Grey dashed line is simple linear regression of Y-linked versus X-linked copy number.

Haplogroup	Events	N	G	Rate (events / 100 gen)
A/J	0	7	155	0.00 (0.00 – 2.4)
C57BL/6J	2	10	236	0.85 (0.10 – 3.1)
129S1/SvImJ	0	10	247	0.00 (0.00 – 1.5)
NOD/ShiLtJ	0	12	301	0.00 (0.00 – 1.2)
NZO/HILtJ	2	13	326	0.61 (0.074 – 2.2)
CAST/EiJ	0	8	194	0.00 (0.00 – 1.9)
PWK/PhJ	1	6	133	0.75 (0.019 – 4.2)
WSB/EiJ	0	3	68	0.00 (0.00 – 5.4)
overall	5	69	1660	0.30 (0.098 – 0.70)

Table 2: Pedigree-based estimates of mutation rates on Yq. *N*, number of CC lines with each Y chromosome haplogroup; *G*, total number of breeding generations.

189 start of inbreeding (**Figure 5A**). We estimated read depth in 100 kb bins across Yq and normalized each bin
 190 against the median for CC lines inheriting a Y chromosome from the same founder strain. This normalization
 191 effectively removes noise from mapping of short reads to repetitive sequence and uncovers CNVs from 6 to 30
 192 Mb in size in 5 CC lines carrying 3 different Y chromosomes (**Table 2**, **Table S2** and **Figure 5B**). Because the
 193 pedigree of each CC line is known, mutation rates — for each Y haplogroup, and overall — can be estimated
 194 directly, assuming each new allele corresponds to a single mutational event. Our estimate of 0.30 (95% Poisson
 195 CI 0.098 – 0.70) mutations per 100 father-son transmissions is about tenfold higher than ampliconic regions of
 196 the human Y chromosome (Repping et al. 2006), and places the mouse Yq among the most labile sequences
 197 known in mouse or human (Egan et al. 2007; Itsara et al. 2010; Morgan et al. 2016b). New Yq alleles also provide
 198 opportunities to investigate the effects of Yq copy number on fertility, sperm phenotypes and sex ratio (as in,
 199 among others, Styrna et al. (1991); Touré et al. (2004); Yamauchi et al. (2010); Cocquet et al. (2012); Fischer et al.
 200 (2016)).

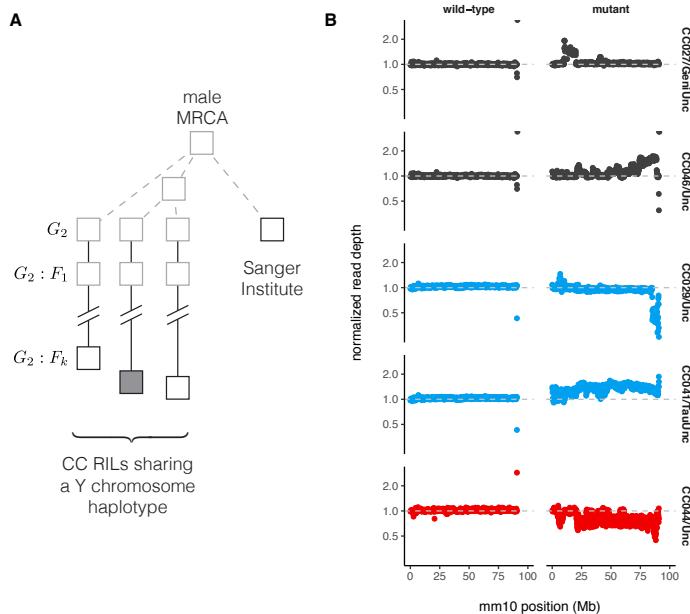


Figure 5: Copy-number variation on Yq in the Collaborative Cross. **(A)** Pedigree-based estimates of mutation rate on the Y chromosome long arm (Yq). Multiple recombinant inbred lines (RILs) from the Collaborative Cross (CC) panel share the same Y chromosome haplotype, with (filled shape) or without (open shape) a putative *de novo* CNV. These Y chromosome lineages are separated from their common male ancestor by an unknown number of generations prior to the initiation of the CC (grey dashed lines), plus a known number of generations of CC breeding (solid lines.) Representatives of the founder strains of the CC were sequenced at the Sanger institute; the number of generations separating the Sanger mouse from the common male ancestor is also unknown. **(B)** Normalized read depth across Yq for CC lines with *de novo* CNVs on Yq. Points are colored according to founder Y chromosome haplogroup. A representative wild-type line is shown for each mutant.

201 **Sequence diversity is markedly reduced on both sex chromosomes**

202 We next used whole-genome sequence data to examine patterns of nucleotide diversity within *Mus musculus*
203 in non-repetitive sequence on Yp compared to the autosomes and X chromosome. To do so we first identified a
204 subset of wild mice without evidence of cryptic relatedness (see **Materials and methods**); this left 20 male and 1
205 female *M. m. domesticus* (hereafter *dom*), 9 male and 10 female *M. m. musculus* (*mus*) and 3 male and 7 female *M.*
206 *m. castaneus* (*cas*). Analyses of autosomes used both males and females from each population; sex chromosome
207 analyses used males only to avoid introducing technical artifacts associated with differences in sample ploidy.
208 Diversity statistics were calculated from the joint site frequency spectrum (SFS), which in turn was estimated
209 directly from genotype likelihoods rather than hard genotype calls (Korneliussen et al. 2014).

210 We estimated nucleotide diversity in four classes of sites: intergenic sites (putatively neutral); introns; 4-fold
211 degenerate sites; and 0-fold degenerate sites. Putatively neutral sites are useful for estimating demographic
212 parameters, while the latter three classes are useful for assessing the impact of selection. Sites on the sex chro-
213 mosomes are subject to different selective pressures than autosomal sites, both because they are “exposed” in
214 the hemizygous state in males and because, in mammals, the sex chromosomes are enriched for genes with
215 sex-specific expression patterns. To evaluate these effects we further subdivided genic sites according to gene-
216 expression patterns inferred from two expression datasets, one in eighteen adult tissues and one a time course
217 across spermatogenesis (see **Materials and methods**). Genes on the autosomes and X chromosome were clas-
218 sified along two independent axes: testis-specific versus ubiquitously-expressed; or expressed early in meiosis,
219 prior to MSCI, versus expressed in post-meiotic spermatids. (Y chromosome genes are not subdivided, since
220 they are few in number and inherited as a single linkage block.) All diversity estimates are shown in **Ta-**
221 **ble S3**. For putatively neutral sites on the autosomes, our estimates of pairwise diversity ($\pi_{\text{dom}} = 0.339\%$,
222 $\pi_{\text{mus}} = 0.325\%$, $\pi_{\text{cas}} = 0.875\%$) are consistent with previous reports based on overlapping samples (Geraldes
223 et al. 2008; Halligan et al. 2013; Kousathanas et al. 2014; Harr et al. 2016). Within each chromosome type, levels
224 of diversity follow the expected rank order: intergenic sites > introns \approx 4-fold degenerate (synonymous) sites
225 > 0-fold degenerate (non-synonymous) sites.

226 For the X chromosome, we further examined the relationship between sequence diversity and local se-
227 quence features including recombination rate, X-Y gametologous amplicons, gene sets described above and
228 blocks of conserved synteny with rat (**Figure S2**). Diversity is reduced across the entire X chromosome in all
229 three populations, in marked contrast to local “troughs” observed in great apes (Nam et al. 2015). Regres-
230 sion of pairwise diversity (θ_π) on distance away from ubiquitously-expressed genes, meiosis genes, spermatid
231 genes, and X-Y ampliconic genes was significant only in *musculus* for ubiquitously expressed genes ($t = 6.6$,
232 Bonferroni-corrected $p = 6.8 \times 10^{-11}$). Similarly — and surprisingly — there was no relationship ($t = -1.2$,
233 $p = 0.23$) between sequence diversity and recombination rate at 100 kb resolution, as estimated from the Diver-

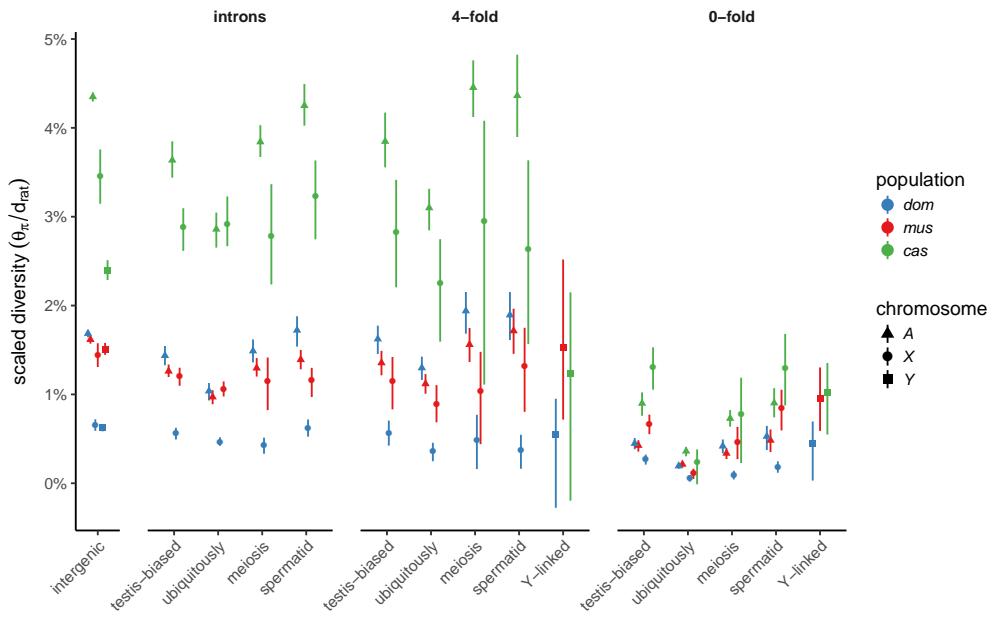


Figure 6: Scaled nucleotide diversity by population, site class and chromosome type. First panel from left shows estimates from intergenic sequence; remaining panels are site classes within protein-coding gene boundaries.

234 sity Outbred mouse stock (Morgan et al. 2017). (We speculate that characterizing recombination at finer scale
 235 from linkage disequilibrium (Auton and McVean 2007) would provide a more powerful test.)

236 In a panmictic population with equal effective number of breeding males and breeding females (*ie.* with
 237 equal variance in reproductive success between sexes), there are 3 X chromosomes and a single Y chromosome
 238 for every 4 autosomes. The expected ratios of X:A and Y:A diversity are therefore $\frac{3}{4}$ and $\frac{1}{4}$, respectively, if mu-
 239 tation rates in males and females are equal (Charlesworth et al. 1987). We estimated X:A and Y:A for putatively
 240 neutral sites and find that diversity on both sex chromosomes is markedly reduced relative to expectations in all
 241 three populations (Table 3). The effect is strongest in *M. m. domesticus* (X:A = 0.244, Y:A = 0.0858) and weakest
 242 in *M. m. musculus* (X:A = 0.563, Y:A = 0.216). The mutation rate is higher in the male than the female germline
 243 in most mammals (recently reviewed in Scally (2016)), including mice, which might contribute to differences in
 244 observed diversity between chromosomes. We used divergence between mouse and rat at synonymous, one-to-
 245 one orthologous sites (d_{rat}) on autosomes, X and Y chromosome as a proxy for the long-term average mutation
 246 rate, and corrected X:A and Y:A estimates for differences in mutation rate (“corrected” rows in Table 3). Even
 247 with this correction, X- and Y-linked diversity remains below expectations. Scaled diversity estimates for each
 248 class of sites are shown in Figure 6. Reduction in X:A diversity has been described previously on the basis of
 249 targeted sequencing of a few loci in all three subspecies (Baines and Harr 2007), and for *M. m. castaneus* on the
 250 basis of whole-genome sequencing (Halligan et al. 2013; Kousathanas et al. 2014). A reduction in Y:A has not,
 251 to our knowledge, been reported.

Comparison	Expected	Scaling	Population		
			dom	mus	cas
X:A	$\frac{3}{4}$	raw	0.244 (0.219 – 0.268)	0.563 (0.506 – 0.613)	0.501 (0.455 – 0.538)
		corrected	0.291 (0.261 – 0.319)	0.670 (0.603 – 0.729)	0.597 (0.542 – 0.641)
Y:A	$\frac{1}{4}$	raw	0.0858 (0.0805 – 0.0911)	0.216 (0.207 – 0.227)	0.128 (0.122 – 0.134)
		corrected	0.0924 (0.0867 – 0.0981)	0.233 (0.223 – 0.244)	0.137 (0.131 – 0.144)

Table 3: Diversity ratios between pairs of chromosome types relative to neutral expectations, with 95% confidence intervals. Both raw diversity and diversity corrected for divergence to rat are shown.

252 Reduction in sex-linked diversity is inconsistent with simple demographic models

253 Sex chromosomes are affected differently than autosomes by both neutral forces, such as changes in population
 254 size (Pool and Nielsen 2007), and by natural selection (reviewed in eg. Ellegren (2011)). The X chromosomes of
 255 humans (Arbiza et al. 2014) and several other primate species (Nam et al. 2015) are substantially less diverse
 256 than the demographic histories of these species would predict, as a result of both purifying selection and re-
 257 current selective sweeps. For humans, the pattern extends to the Y chromosome (Sayres et al. 2014). Having
 258 observed a deficit of polymorphism on both sex chromosomes in mouse, the central question arising in this
 259 paper is: to what extent is sex-chromosome diversity reduced by natural selection? A rich body of literature
 260 already exists for the influence of selection on the mouse X chromosome, especially in the context of speciation
 261 (Good et al. 2008; Teeter et al. 2008; Baines and Harr 2007; Kousathanas et al. 2014; Larson et al. 2016a,b), so we
 262 directed our focus to the lesser-studied Y chromosome.

263 To establish an appropriate null against which to test hypotheses about natural selection on the sex chro-
 264 mosomes, we followed an approach similar to Sayres et al. (2014). We fit four simple demographic models
 265 to SFS from putatively-neutral intergenic sites on the autosomes using the maximum-likelihood framework
 266 implemented in $\partial\alpha\partial i$ (Gutenkunst et al. 2009) (**Figure 7A**). Each model is parameterized by an initial effective
 267 population size (N_0), a size change (expressed as fraction f of starting size for models involving instantaneous
 268 size changes, or the ending population size N_e for the exponential growth models), and a time of onset of size
 269 change (τ). The relative fit of each model was quantified using the method of Akaike weights (Akaike 1978).
 270 All four models can be viewed as nested in the family of three-epoch, piecewise-exponential histories. In prin-
 271 ciple, such models are identifiable with sample size of $4 \times 3 = 12$ or more chromosomes (6 diploid individuals)
 272 (Bhaskar and Song 2014). In practice, more than one model fits each population about equally well (or equally
 273 poorly), with the exception of *M. m. castaneus*, which is best described by the “step-change” model (**Figure 7B-**
 274 **C**). Of course the true history of each population is almost certainly more complex than any of our models. Our
 275 goal is not to obtain a comprehensive description of mouse population history as such, but rather to pick an
 276 appropriate null model against which we can test hypotheses about selection. For *domesticus*, the stationary
 277 model is the most parsimonious; for *musculus*, the exponential-growth model.

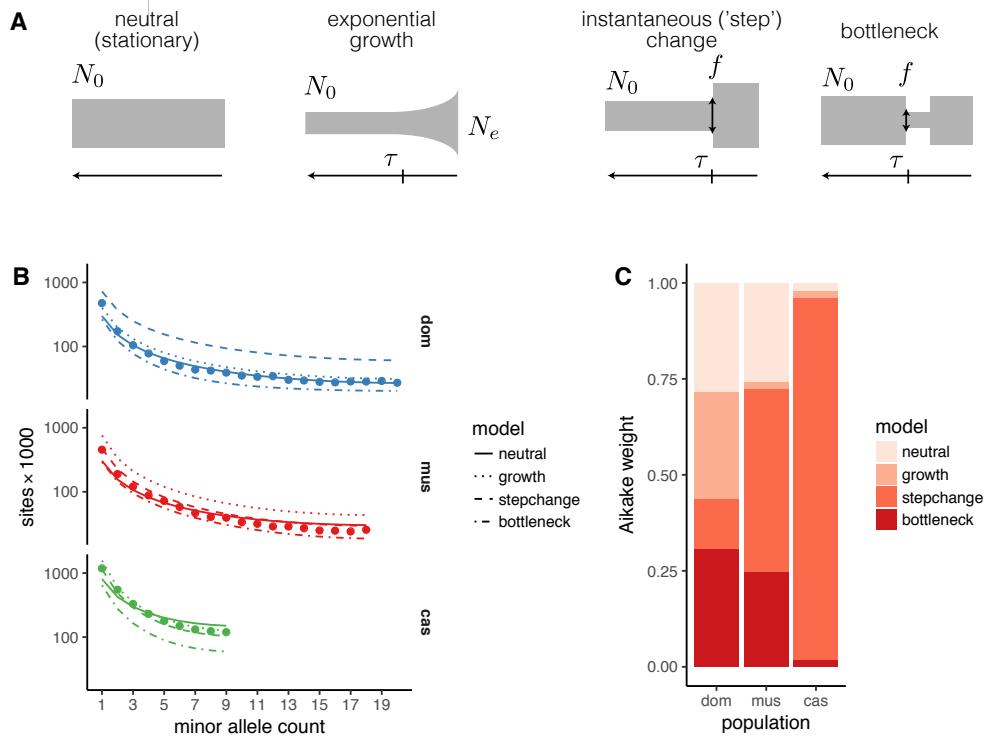


Figure 7: Inference of demographic histories from autosomal sites. (A) Four simple demographic models fit with $\partial a \partial i$. Each model is parameterized by one or more of an ancestral effective population size (N_0), time of population-size change (τ), change in population size as fraction of initial size (f), and present effective population size (N_e). (B) Observed site frequency spectra by population, with fitted spectra from the four models in panel A. (C) Relative support for each model, quantified by Aikake weight, by population.

Population	Model	Parameter			
		N_0	N_e	f	τ
dom	neutral	162 (2)	-	-	-
	growth	160 (50)	230 (80)	-	1.53 (0.06)
	step-change	230 (50)	400 (700)	2 (4)	7 (6)
	bottleneck	164 (3)	30 (50)	0.2 (0.3)	0 (2)
mus	neutral	165 (2)	-	-	-
	growth	150 (40)	500 (200)	-	7 (5)
	step-change	150 (50)	270 (100)	1.8 (0.5)	1 (3)
	bottleneck	164 (4)	20 (8)	0.12 (0.05)	1 (1)
cas	neutral	429 (3)	-	-	-
	growth	350 (100)	2000 (1000)	-	0 (2)
	step-change	300 (100)	800 (700)	3 (1)	1 (4)
	bottleneck	431 (7)	30 (10)	0.08 (0.03)	0.6 (0.9)

Table 4: Parameter estimates for models shown in Figure 7. Population sizes are given in thousands and times in units of N_0 ; bootstrap standard errors in parentheses.

278 We also compared our estimates of sex chromosome diversity to predictions from coalescent theory (Pool
279 and Nielsen 2007; Polanski et al. 2017) for the four models considered above. For this analysis we focused on
280 the ratios X:A and Y:A, which are independent of autosomal effective population size. Results are shown in
281 **Figure S3**, with observed X:A and Y:A ratios superposed. We summarize some relevant trends here and refer
282 to previous reviews (Pool and Nielsen 2007; Webster and Wilson Sayres 2016) for further details. Qualitatively,
283 both X:A and Y:A are reduced after an instantaneous contraction in population size, eventually recovering to
284 their stationary values after about $4N_e$ generations. For a bottleneck — a contraction followed by instantaneous
285 recovery to the initial size — X:A and Y:A are at first sharply reduced and then increased relative to a stationary
286 population, again returning to stationary values after about $4N_e$ generations. With exponential growth, X:A
287 and Y:A are actually increased relative to their stationary values. These patterns are modulated by the breeding
288 sex ratio; X:A increases and Y:A decreases when females outnumber males, and vice versa. In brief, some
289 combination of a male-biased breeding ratio and a very strong ($f \ll 0.1$) population contraction would be
290 required to explain the observed reductions in X:A and Y:A in *domesticus*, with somewhat milder effects required
291 to explain the reduction in *musculus* or *castaneus*. These histories are not consistent with population histories
292 inferred from autosomal SFS. We hypothesize that this discrepancy is explained, at least in part, by selection.

293 **Both sex chromosomes have been shaped by positive selection in the male germline**

294 We used two approaches to investigate the role of selection on the sex chromosomes. First, we used a variant
295 of the McDonald-Kreitman test (McDonald and Kreitman 1991) to obtain a non-parametric estimate of the pro-
296 portion α of sites fixed by positive selection (loosely, the “evolutionary rate”) in genes with different expression
297 and inheritance patterns (Smith and Eyre-Walker 2002). The rate of adaptive evolution should be faster on the
298 X chromosome when new mutations tend to have greater fitness effect in males than in females, to be on aver-
299 age recessive, or both (Charlesworth et al. 1987). We might expect genes with testis-biased expression or genes
300 expressed during spermatogenesis to be targets of male-specific selection. Consistent with previous work on
301 the “faster-X” effect in mouse (Kousathanas et al. 2014; Larson et al. 2016a,b), we find that a greater proportion
302 of X-linked than autosomal substitutions are adaptive. The pattern holds in all three populations (**Figure 8**). In
303 *domesticus* and *musculus*, X-linked genes whose expression is biased towards early meiosis or round spermatids
304 evolve faster than X-linked genes with ubiquitous expression or expression across spermatogenesis. By con-
305 trast, non-amplionic Y-linked genes — all expressed during male meiosis — have evolutionary rates closer to
306 autosomal genes, with heterogeneity across populations. Unfortunately we cannot assess the rate of sequence
307 evolution in ampliconic gene families on the Y chromosome using short-read data.

308 Second, we used forward simulations from the models fit to autosomal SFS to explore the possible contri-
309 bution of natural selection to the SFS of Y chromosomes. We simulated two modes of selection independently:

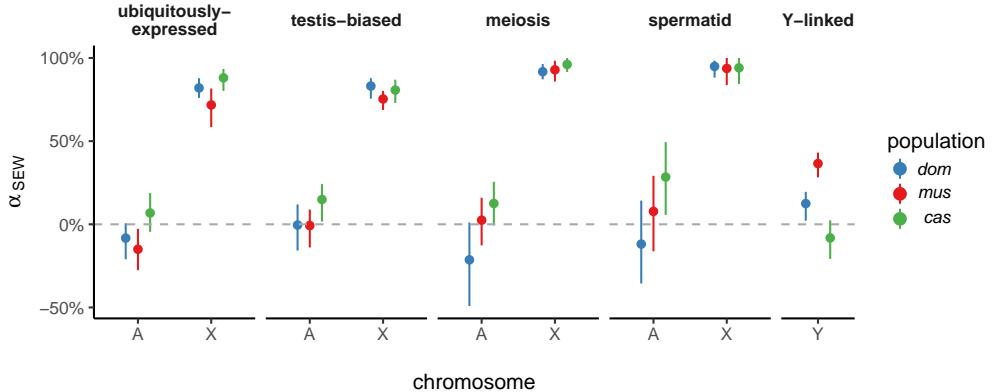


Figure 8: Proportion of sites fixed by positive selection (using the non-parametric estimator of Smith and Eyre-Walker (2002), α_{SEW}) according to gene-expression class, chromosome and population. Error bars represent 95% bootstrap CIs. Ampliconic genes on X and Y are excluded.

purifying selection on linked deleterious alleles (background selection, BGS; (Hudson and Kaplan 1995)), and hard selective sweeps on newly-arising beneficial alleles. For the BGS model, we varied the proportion of sites under selection α and the mean population-scaled selection coefficient $\gamma = N\bar{s}$; for the sweep model, we varied only the γ for the sweeping allele. (Simulation details are provided in **Materials and methods**.) Posterior distributions for these parameters were inferred using an approximate Bayesian computation (ABC) approach (Pritchard et al. 1999; Beaumont et al. 2002); Bayes factors were used for model comparison. The *castaneus* population was excluded from these analyses because sample size (only 3 chromosomes) was not sufficient for calculating some of the summary statistics chosen for ABC.

Results of the ABC procedure are shown in **Figure S4**. The Y chromosomes of *domesticus* are best approximated by the selective-sweep model. For *musculus* the result is less clear: the neutral null model actually provides the best fit, and among models with selection, the BGS model is superior. However, over the parameter ranges used in our simulations, we have limited power to discriminate between different models at the current sample size ($n \leq 20$ chromosomes) (**Figure S4B**). In the best case — the selective-sweep model — we achieve only 49% recall. This reflects both the constraints of a small sample and the more fundamental limits on model identifiability for a single non-recombining locus like the Y chromosome.

If a selective sweep did occur on *domesticus* Y chromosomes, it was moderately strong: we estimate $N\bar{s} = 9.29$ (50% HPDI 0 – 9.88) (**Table 5**). For comparison, $N\bar{s} \approx 500$ for adaptive alleles in the human lactase gene (*LCT*), a well-characterized example of recent positive selection (Tishkoff et al. 2007). Posterior distributions of several estimators of nucleotide diversity recapitulate the values observed in real data (**Figure S4D**). We note that, because the Y chromosome is inherited without recombination, our estimate of $N\bar{s}$ reflects the cumulative selection intensity on the entire chromosome and not necessarily on a single site.

Model	Population	Best?	Parameter	
			α	$N\bar{s}$
BGS	<i>dom</i>	\checkmark	0.675 (0.624 – 1.00)	43.7 (12.2 – 96.4)
	<i>mus</i>		0.481 (0.118 – 0.553)	2.4 (0 – 20.2)
BGS+growth	<i>dom</i>	\checkmark	0.647 (0.643 – 0.960)	25.7 (0 – 26.2)
	<i>mus</i>		0.473 (0.317 – 0.649)	8.39 (0 – 12.9)
sweep	<i>dom</i>	\checkmark	-	9.29 (0 – 9.88)
	<i>mus</i>		-	0.663 (0 – 0.934)

Table 5: Parameter estimates from ABC. Values are shown as posterior median and 50% highest posterior density interval (HPDI). Best-fitting model for each population indicated by check mark.

331 Sex-linked gene expression diverges rapidly in the testis

332 Given the dramatic differences in Y-linked gene content between even closely-related *Mus* taxa, we finally asked
 333 whether patterns of gene expression showed similar divergence. In particular, we sought to test the prediction
 334 that expression patterns of Y-linked genes diverge more rapidly than autosomal genes in the testis. To that
 335 end we re-analyzed published gene expression data from the brain, liver and testis of wild-derived outbred
 336 individuals representing seven (sub)species spanning an 8 million year evolutionary transect across the murid
 337 rodents (Neme and Tautz 2016) (**Figure 9A**). For genes on the autosomes and X chromosome, the great majority
 338 of expression variance lies between tissues rather than between (sub)species (PC1 and PC2, cumulative 77.1%
 339 of variance explained; **Figure 9B**). For Y-linked genes, highly enriched for function in the male germline, most
 340 variance (PC1, 59.6% of variance explained) naturally lies between the testis and the non-germline tissues.

341 To quantify divergence in gene expression patterns we computed the rank-correlation (Spearman's ρ) be-
 342 tween species for each tissue type separately for autosomal, X-linked and Y-linked genes, and constructed trees
 343 by neighbor-joining (**Figure 9C**). We use total tree length as an estimator of expression divergence. The topology
 344 of these trees for the autosomes and X chromosome in brain and testis is consistent with known phylogenetic
 345 relationships within the Muridae. Consistent with previous comparative analyses of gene expression in mam-
 346 mals (Brawand et al. 2011), we find that expression patterns are most constrained in brain and least constrained
 347 in testis (**Figure 9D**). Expression divergence is equal between autosomes and X chromosome in brain and liver,
 348 but greater for X-linked genes in testis. Y-linked expression diverges much more rapidly in all three tissues, but
 349 the effect is most extreme in the testis. We caution that the precision of these estimates is limited by the small
 350 number of Y-linked relative to autosomal or X-linked genes.

351 This "faster-X" effect should be limited to functional elements subject to male-specific selection. Genes
 352 expressed in the male germline (testis-biased and/or expressed during spermatogenesis) might be enriched
 353 for such elements, relative to genes with ubiquitous expression. We therefore estimated expression divergence
 354 in autosomal, X- and Y-linked genes with four sets of genes with different expression patterns (**Figure 9E**).
 355 X-linked expression diverges more rapidly than autosomal expression only among genes with testis-biased

356 expression. In contrast to Larson et al. (2016b) but in keeping with other predictions (Good and Nachman
357 2005), we find that the “faster-X” effect on expression is larger for genes expressed late than early in meiosis
358 (**Figure 9F**). The number of Y-linked genes in each group is too small to permit any strong conclusions.

359 Discussion

360 We have shown that nucleotide diversity in *M. musculus* is reduced on both sex chromosomes relative to expec-
361 tations for a stationary population, and that the effect appears strongest in *M. m. domesticus* and weakest in *M.*
362 *m. musculus* (**Table 3**). Sex differences in the long-term average mutation rate, estimated from synonymous-sites
363 divergence to rat, are not sufficient to explain the deficit. Because sex chromosomes respond differently than
364 autosomes to changes in population size, we fit several (simple) models of demographic history to autosomal
365 site-frequency spectra (**Figure 7**) and compared their predictions to observed values. At least for the models we
366 considered (see **Supplement**), neither gradual nor instantaneous changes in population size — of magnitude
367 feasible given autosomal SFS — can account for the reduction in diversity on both the X and Y chromosomes,
368 even if we allow for a sex ratio different than 1:1 (**Figure S3**). Estimates of effective size of each population
369 (from autosomal sites) are in agreement with previous work on house mice (Din et al. 1996; Baines and Harr
370 2007; Salcedo et al. 2007; Geraldes et al. 2008).

371 Using demographic histories from autosomes as a baseline, we simulated two modes of selection — back-
372 ground selection and hard selective sweeps — on Y chromosomes of *domesticus* and *musculus*. Although dis-
373 crimination between models was limited by both technical factors and theoretical constraints, we have shown
374 that the Y-linked SFS in *domesticus* is consistent with a moderately strong selective sweep (**Figure S4**). The
375 background selection model is the best-fitting in *musculus*, but is only 1.4-fold more likely ($\log_{10} \text{BF} = 0.16$)
376 than the next-best model. We conclude that recent positive selection accounts, at least in part, for the reduc-
377 tion in Y-linked relative to autosomal diversity in *domesticus*. Furthermore, coding sequences of X-linked genes
378 with germline expression are disproportionately shaped by positive selection (**Figure 8**). Both X- and Y-linked
379 genes have rapidly-diverging expression patterns in the testis, especially in spermatids (**Figure 9**). Together
380 these findings provide strong support for the idea that positive selection in the male germline is a potent and
381 ongoing force shaping both mammalian sex chromosomes (Mueller et al. 2013; Larson et al. 2016b).

382 To what extent are these pressures a consequence of intragenomic conflict? The reciprocal actions of SLX/SLX1
383 and SLY on sex-linked gene expression in spermatids establish the conditions for conflict between the X and Y
384 chromosomes that implicates *any* gene whose expression after meiosis is beneficial for sperm maturation and
385 fertilizing ability. X-linked alleles that meet the functional requirement for post-meiotic expression in the face
386 of repression by SLY — via a stronger promoter, a more stable transcript, a more active protein product, or
387 increased copy number — should be favored by selection (Ellis et al. 2011). The same should be true, in reverse,

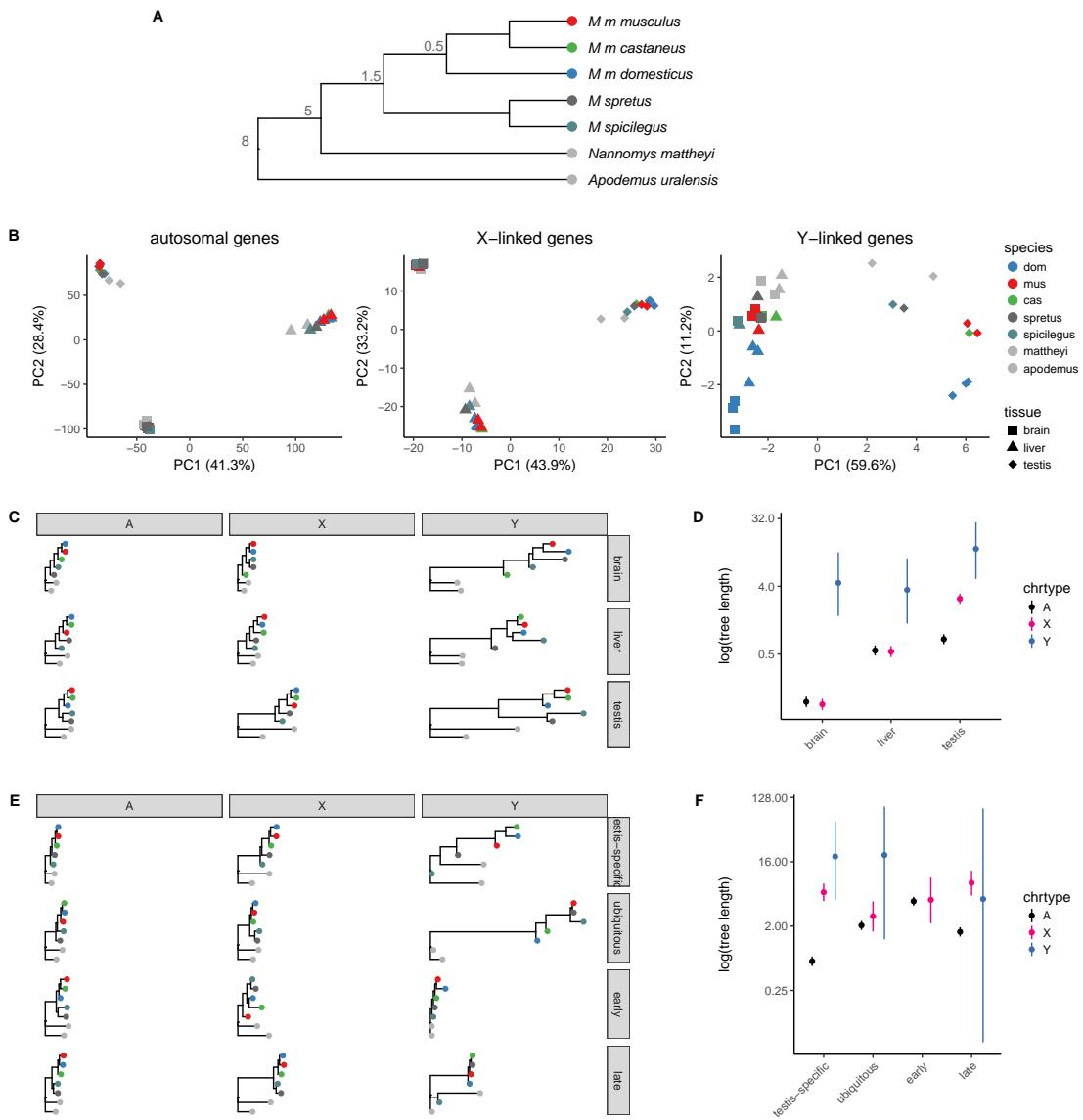


Figure 9: Divergence of sex-linked gene expression in murid rodents. (A) Schematic phylogeny of taxa in the multi-tissue expression dataset. Node labels are approximate divergence times (Mya); branch lengths not to scale. (B) Projection of samples onto the top two principal components of expression values for autosomal, X-linked and Y-linked genes. (C) Expression trees computed from rank-correlations between taxa for autosomal (A), X-linked (X) and Y-linked (Y) genes (across columns) for brain, liver and testis (across rows.) (D) Total tree length by chromosome type and tissue. (E) Expression trees as in panel C, with genes partitioned according to expression pattern: testis-specific; ubiquitously-expressed; early spermatogenesis (meiosis prior to MSCI); and late spermatogenesis (spermatids). (F) Total tree length by chromosome type and expression pattern.

388 for successful Y chromosomes.

389 Although we cannot directly identify the putative target(s) or mechanism(s) of selective sweeps on the Y
390 chromosome, several independent lines of evidence point to the ampliconic genes on Yq active in the X-Y
391 conflict. First, the copy number of *Slx/Slxl1* and *Sly* have increased three-fold within *M. musculus* and are
392 correlated across populations (**Figure 4**), consistent with an “arms race” between the sex chromosomes in which
393 the Y chromosome is the lagging player. The absolute expression of ampliconic X genes and their Yq homologs
394 (in whole testis) increases with copy number across *Mus* (**Figure S5**). Larson et al. (2016a) have shown that,
395 in spermatids from reciprocal *F*₁ hybrids between *domesticus* and *musculus* that are “mismatched” for *Slx/Slxl1*
396 and *Sly*, global X-chromosome expression is indeed perturbed in the direction predicted by the copy number
397 and actions of SLX/SLX1 and SLY. Second, several independent deletions of Yq in laboratory stocks converge
398 on a similar phenotype, namely low fertility, abnormal sperm morphology due to problems with chromatin
399 compaction, and sex-ratio distortion in favor of females (Styrna et al. 1991; Conway et al. 1994; Touré et al. 2004;
400 Fischer et al. 2016; MacBride et al. 2017). Third, Y chromosomes from *musculus* — the subspecies with highest
401 *Sly* copy number — are more successful at introgressing across *domesticus-musculus* hybrid zone in Europe, and
402 in localities where they do, the census sex ratio is shifted towards males (Macholán et al. 2008). Consomic
403 strains with differing only by their Y chromosomes show similar deviation in the sex ratio from parity (Case
404 et al. 2015). Finally, although modeling predicts moderately strong positive selection on Y, there is little evidence
405 that it occurs within coding sequences of single-copy genes on Yp (**Figure 8**). This observation permits several
406 explanations but is consistent with the idea that Yp alleles are hitchhiking with favorable alleles on Yq.

407 It is more difficult to ascertain the contribution of intragenomic conflict to the paucity of diversity on the
408 X chromosome. Although the mammalian X chromosome is enriched for genes with expression in the male
409 germline (eg. Rice (1984); Mueller et al. (2013)), its functional portfolio is considerably more broad than that
410 of the Y chromosome (Bellott et al. 2014, 2017). The X chromosome also has a major role in hybrid sterility in
411 mouse (Forejt and Iványi 1974; Forejt 1996; Storchová et al. 2004; Payseur et al. 2004; Teeter et al. 2008; Good et al.
412 2008; Campbell et al. 2013; Turner et al. 2014); the Y chromosome does not (Turner et al. 2012; Campbell and
413 Nachman 2014). We corroborate the “faster-X” effect on protein evolution that has been previously described
414 by others (Kousathanas et al. 2014; Larson et al. 2016b) and show that it is strongest for genes expressed in
415 the male germline (**Figure 8**), which are widely scattered across the X chromosome (**Figure S2D**). We conclude
416 that selection is pervasive on the mouse X chromosome and reduces diversity chromosome-wide. This stands
417 in contrast to the pattern observed in great apes, which has apparently been driven by a few strong selective
418 sweeps (Hvilsom et al. 2012; Veeramah et al. 2014; Nam et al. 2015).

419 Many open questions remain with respect to the evolution of mouse Y chromosomes. How many of the
420 hundreds of copies in each gene family retain coding potential? Which copies are functionally equivalent? Does

421 suppression of recombination promote the spread of clusters of genes like *Slx*, similar to sex-ratio drivers in
422 other species (Jaenike 2001)? What evolutionary trade-offs does success in the sex-chromosome conflict entail,
423 in the context of sperm competition and polyandry in natural populations (Simmons and Fitzpatrick 2012)?
424 Does the conflict lead to oscillations between a male-biased and female-biased population over time, and if so,
425 what is the effect on patterns of diversity on the sex chromosomes? All of these are important avenues of future
426 study as we seek to understand the forces shaping sex chromosomes.

427 Materials and methods

428 Alignment and variant-calling

429 Whole-genome sequencing reads were obtained from the European Nucleotide Archive (PRJEB9450, PRJEB11742,
430 PRJEB14673, PRJEB14167, PRJEB2176, PRJEB15190) and whole-exome reads from the NCBI Short Read Archive
431 (PRJNA323493). Reads were aligned to the mm10 reference sequence using `bwa mem v0.7.15-r1140` (Li 2013)
432 with default parameters. Optical duplicates were marked using `samblaster` and excluded from downstream
433 analyses. Regions of the Y chromosome accessible for variant calling were identified using the `CallableLoci`
434 tool in the `GATK v3.3-0-g37228af` (McKenna et al. 2010). To be declared “callable” within a single sample, sites
435 were required to have depth consistent with a single haploid copy ($3 < \text{depth} < 50$) and $< 25\%$ of overlapping
436 reads having mapping quality (MQ) zero. The analysis was restricted to Yp. The final set of callable sites was
437 defined as any site counted as callable within > 10 samples. In total, 2 289 336 bp (77% of the non-gap length of
438 Yp) were deemed callable.

439 SNVs and short indels on the Y chromosome were ascertained using `GATK HaplotypeCaller v3.3-0-`
440 `g37228af` in the intersection of callable regions and exons targeted by the Roche NimbleGen exome-capture
441 array, lifted over to mm10 with `CrossMap v0.2.3` and the mm9-to-mm10 chain file from the UCSC Genome
442 Browser (<http://hgdownload.soe.ucsc.edu/goldenPath/mm9/liftOver/mm9ToMm10.over.chain.gz>). To minimize artifacts from cryptic copy-number variation, X-Y homology, and the like, only biallelic sites
443 with a “homozygous” (ie. single-copy hemizygous) call in all male samples were used. Sites called in fewer
444 than 60 samples or with strand-bias p -value < 0.01 were filtered. Raw VCF files are provided in **File S1**.

446 For the Y chromosome phylogenetic tree shown in **Figure 2A**, data from Collaborative Cross lines carry-
447 ing A/J, 129S1/SvImJ, NOD/ShiLtJ, NZO/HILtJ, CAST/EiJ, PWK/PhJ, and WSB/EiJ Y chromosomes were
448 used in place of the inbred strains themselves. (See aliases in **Table S1**.) Whole-genome sequence from male
449 representatives of these lines has not (to our knowledge) been published.

450 Estimation of site frequency spectra and summary statistics

451 Site frequency spectra (SFS) were calculated from genotype likelihoods at callable sites using ANGSD v0.917 (Ko-
452 rneliussen et al. 2014). Genotype likelihoods for the autosomes were caculated under the GATK diploid model
453 after applying base alignment quality (BAQ) recalibration with the recommended settings for bwa alignments
454 (-baq 1 -c 50, effectively discarding evidence from reads aligning at < 95% identity). Sites were filtered
455 to have per-individual coverage consistent with the presence of a single diploid copy ($3 < \text{depth} < 80$), to be
456 non-missing in at least 3 individuals per population. Genotype likelihoods for the X and Y chromosomes were
457 calculated under the GATK haploid model with depth filters appropriate for haploid sites ($3 < \text{depth} < 40$).
458 Only reads with MQ > 20 and bases with call quality > 13 were considered, and ampliconic regions (plus a 100
459 kb buffer on each side) were masked. Site-wise allele frequencies were computed within each population sepa-
460 rately, and the joint SFS across non-missing sites in the three populations was estimated from these frequencies.
461 The consensus genotype from a single *Mus spicilegus* male was used as the ancestral sequence to polarize alleles
462 as ancestral or derived. Ensembl v87 reference annotations were used to define interneine sites, intronic sites,
463 0-fold and 4-fold degenerate sites.

464 Diversity statistics and neutrality tests were calculated from joint SFS using standard formulae implemented
465 in a custom Python package, sfspy (<http://github.com/andrewparkermorgan/sfspy>). Uncertainties
466 for autosomal and X-linked sites were obtained by bootstrapping over loci, since the X and autosomes recom-
467 bine; and for Y-linked sites using the built-in bootstrapping method of ANGSD.

468 Models of sex-chromosome diversity under neutral coalescent

469 The expected ratio of X-to-autosome (X:A) and Y-to-autosome (Y:A) pairwise diversity was obtained from the
470 formulae derived in Pool and Nielsen (2007). Define the inheritance factors $h_A = 1$, $h_X = 3/4$ and $h_Y = 1/4$; and
471 mutation rates μ_A , μ_X , μ_Y . For an instantaneous change in population size of depth f from starting size N , the
472 expected value of X:A is:

$$\frac{\theta_{\pi,X}}{\theta_{\pi,A}} = \frac{h_X \mu_X}{h_A \mu_A} \frac{(f - (f - 1) \left(1 - \frac{1}{2Nh_X f}\right)^g)}{(f - (f - 1) \left(1 - \frac{1}{2Nh_A f}\right)^g)}$$

473 The expression for Y:A can be written similarly. Note that X:A and Y:A depend only on the ratio between
474 mutation rates on different chromosomes, not the absolute mutation rate. For a bottleneck of depth f , starting
475 g_1 generations before the present and ending at $g_1 + g_2$ generations before the present:

$$\frac{\theta_{\pi,X}}{\theta_{\pi,A}} = \frac{h_X \mu_X}{h_A \mu_A} \frac{\exp\left(\frac{(fg_1+g_2)(h_X-h_A)}{2Nh_X h_A f}\right) \left(1 - f + \exp\left(\frac{g_2}{2Nh_X f}\right) \left(f - 1 + \exp\left(\frac{g_1}{2Nh_X}\right)\right)\right)}{1 - f + \exp\left(\frac{g_2}{Nh_A f}\right) \left(f - 1 + \exp\left(\frac{g_1}{2Nh_X}\right)\right)}$$

476 For a model with exponential growth with rate constant r , we used the approximation provided in Polanski
477 et al. (2017):

$$\frac{\theta_{\pi,X}}{\theta_{\pi,A}} = \frac{h_X \mu_X}{h_A \mu_A} \frac{\log \left\{ 2rN h_X \left(1 - \frac{1}{N} \right) + 1 \right\}}{2Nr}$$

478 Unequal sex ratios were modeled by calculating the number of X and Y chromosomes per autosome, given
479 fixed autosomal effective population size, using standard formulae as in Sayres et al. (2014), and passing these
480 into the equations above via the parameter h_X or h_Y .

481 In **Figure S3**, we plot X:A against Y:A. For the bottleneck and step-change models, X:A and Y:A vary with
482 time since the onset of size change; these trajectories can be traced clockwise along each curve from $t = 0$ to
483 $t = 4N$ (backwards in time.)

484 Demographic inference

485 The four demographic models illustrated in **Figure 7A** were fit to autosomal SFS using $\partial a \partial i$ v1.7 (Gutenkunst
486 et al. 2009). We fit each model separately to each population, using the sum of marginal spectra from 1000
487 approximately unlinked, putatively neutral intergenic regions each 100 kb in size, spanning a total of 85.4
488 Mb of callable sites after removing those missing in one or more populations. Because the depth, duration
489 and onset of a bottleneck have are confounded in the SFS, we fixed the duration of the bottleneck to be short
490 ($0.1 N_e$ generations) and attempted to estimate the remaining two parameters. We additionally constrained the
491 bottleneck model to include recovery to exactly the starting population size.

492 Convergence of model fits was assessed qualitatively by re-fitting each model from 10 sets of randomly-
493 drawn initial values. We confirmed that the best-fitting models shown in **Figure 7** represent the “modal” result,
494 in that a majority of independent runs reach a solution within 5 log-likelihood units of the one shown. Parameter
495 estimates should nonetheless be interpreted with caution, as their uncertainties are wide.

496 Models of natural selection

497 To model the effect of natural selection on Y-linked diversity while accounting for possible non-stationary de-
498 mographic processes, we used forward simulations implemented in **SLiM** v2.2.1 (Haller and Messer 2017). For
499 *M. m. domesticus* we simulated from a stationary model; for *M. m. musculus*, from an exponential growth model.

500 We considered two modes of selection: background selection (BGS) due to purifying selection against dele-
501 terious mutations at linked sites; and hard selective sweeps on newly-arising beneficial mutations. Relative
502 fitness in **SLiM** is modeled as $1 + s$ for sex-limited chromosomes. BGS was modeled by introducing muta-
503 tions whose selection coefficients s were drawn from a mixture of a gamma distribution with mean $-\gamma = -N\bar{s}$
504 ($100 \times \alpha\%$ of mutations), and a point mass at zero ($100 \times (1 - \alpha)\%$ of mutations.) BGS simulations were run for

505 $10N$ generations, sufficient to reach mutation-selection-drift equilibrium. For selective sweeps, the simulation
506 was first run for $10N$ generations of burn-in, and then a beneficial variant was introduced with s drawn from a
507 gamma distribution with mean $\gamma = N\bar{s}$. The simulation was then tracked until the beneficial variant was fixed
508 or lost; in the case of loss, the run was re-started from the end of the burn-in period with a new mutation. We
509 confirmed the integrity of simulations by checking that the pairwise diversity achieved by runs with selection
510 coefficients fixed at zero matched the observed neutral values for each population (not shown.) Values of α
511 were drawn from a uniform distribution on $(0, 1]$, and values of γ were drawn from a log-uniform distribution
512 on $(10^{-6}, 10^3]$. Runs were scaled for computational performance.

513 Simulations were connected to an approximate Bayesian computation (ABC) inference procedure imple-
514 mented with the R package `abc` (Csilléry et al. 2012). Briefly, 500,000 simulations were performed for each
515 model. Five summary statistics were calculated from the SFS generated by each simulation: Watterson's esti-
516 mator θ_w ; Tajima's estimators θ_π and θ_ζ ; Tajima's D ; and Fu and Li's D . The same set of statistics was computed
517 for the observed joint SFS. The 0.1% of simulations with smallest Euclidean distance to the observed sum-
518 mary statistics were retained, accounting for collinearity between summary statistics using the "neuralnet"
519 method of the function `abc::abc()`. Posterior distributions were computed via kernel smoothing over the
520 parameter values of the retained simulations using an Epanechnikov kernel and plug-in bandwidth estimate.

521 Models were compared via their Bayes factors, calculated using the `abc::postpr()` function. To confirm
522 the fidelity of the best-fitting model, summary statistics for pseudo-observed datasets (*i.e.* simulations from the
523 posterior distributions) were checked against the observed summary statistics.

524 **Size estimation of co-amplified regions of Yq and X**

525 Copy number of ampilconic genes on Yq and X was estimated as follows. First, all paralogs in each family were
526 identified by BLAT and BLAST searches using the sequences of canonical family members from Ensembl. These
527 searches were necessary because many member of each family are annotated only as "predicted genes" (gene
528 symbols "GmXXXX"). Based on BLAST results we assigned the *Spin2/4* family – with members in several clus-
529 ters on the proximal X chromosome – as *Sstx*. Normalized coverage was estimated for each non-overlapping
530 paralog by counting the total number of reads mapped and dividing by the genome-wide average read depth.

531 **Identification of *de novo* CNVs in Collaborative Cross lines**

532 Whole-genome sequencing reads (2×150 bp paired-end) from a single male individual from each 69 distinct
533 Collaborative Cross (CC) lines were obtained from Srivastava et al. (2017). Alignment and quality control was
534 performed as for wild mice. Read depth was estimated in 100 kb bins across the Y chromosome for each indi-
535 vidual, and normalized for the effective depth of sequencing in that sample. Unambiguous alignment of 150 bp

536 reads to the highly repetitive sequence on Yq is clearly not possible. However, each of the 8 founder Y chromo-
537 some haplogroups in the CC produces a characteristic read depth profile when reads are aligned with `bwa-mem`.
538 We exploited this fact to remove noise from ambiguous read mapping by re-normalizing the estimated depth
539 in each bin for each sample against the median depth in that bin for CC lines sharing the same Y chromosome
540 haplogroup (listed in **Table S2**). Any remaining deviations in read depth represent variation among lines shar-
541 ing the same Y chromosome haplogroup, that is, candidate *de novo* CNVs. CNVs were ascertained by manual
542 inspection of the re-normalized read depth profile of each CC line.

543 **Analyses of gene expression**

544 *Multi-tissue, multi-species dataset.* Neme and Tautz (Neme and Tautz 2016) measured gene expression in whole
545 testis from wild-derived outbred mice from several species (**Figure 9A**) using RNA-seq. Reads were retrieved
546 from the European Nucleotide Archive (PRJEB11513). Transcript-level expression was estimated using `kallisto`
547 (Bray et al. 2016) using the Ensembl 85 transcript catalog augmented with all *Slx/y*, *Sstx/y* and *Srsx/y* transcripts
548 identified in (Soh et al. 2014). In the presence of redundant transcripts (*i.e.* from multiple copies of a co-
549 amplified gene family), `kallisto` uses an expectation-maximization algorithm to distribute the “weight” of
550 each read across transcripts without double-counting. Transcript-level expression estimates were aggregated
551 to the gene level for differential expression testing using the R package `tximport`. As for the microarray data,
552 “predicted” genes (with symbols “GmXXXX”) on the Y chromosome were assigned to a co-amplified family
553 where possible using Ensembl Biomart.

554 Gene-level expression estimates were transformed to log scale and gene-wise dispersion parameters esti-
555 mated using the `voom()` function in the R package `limma`. Genes with total normalized abundance (length-
556 scaled transcripts per million, TPM) < 10 in aggregate across all samples were excluded, as were genes with
557 TPM > 1 in fewer than three samples.

558 *Spermatogenesis time course.* Larson et al. (2016a) measured gene expression in isolated spermatids of three
559 males from each of four *F*₁ crosses — CZECHII/EiJ × PWK/PhJ; LEWES/EiJ × PWK/PhJ; PWK/PhJ × LEWES/EiJ;
560 and WSB/EiJ × LEWES/EiJ — using RNA-seq. Reads were retrieved from NCBI Short Read Archive (SRP065082).
561 Transcript-level expression was estimated using `kallisto` (Bray et al. 2016) using the Ensembl 85 transcript
562 catalog augmented with all *Slx/y*, *Sstx/y* and *Srsx/y* transcripts identified in (Soh et al. 2014). In the presence of
563 redundant transcripts (*i.e.* from multiple copies of a co-amplified gene family), `kallisto` uses an expectation-
564 maximization algorithm to distribute the “weight” of each read across transcripts without double-counting.
565 Transcript-level expression estimates were aggregated to the gene level for differential expression testing using
566 the R package `tximport`. As for the microarray data, “predicted” genes (with symbols “GmXXXX”) on the Y
567 chromosome were assigned to a co-amplified family where possible using Ensembl Biomart.

568 Gene-level expression estimates were transformed to log scale and gene-wise dispersion parameters esti-
569 mated using the `voom()` function in the R package `limma`. Genes with total normalized abundance (length-
570 scaled transcripts per million, TPM) < 10 in aggregate across all samples were excluded, as were genes with
571 TPM > 1 in fewer than three samples.

572 *Definition of tissue-specific gene sets.* The “tissue specificity index” (τ) of Yanai et al. (2005) was used to define
573 tissue- or cell-type-specific gene sets. The index was first proposed for microarray data, and was adapted for
574 RNA-seq as follows. Define T_i to be the mean log-scaled expression of a gene in tissue or cell type i (of N
575 total), as estimated by `limma`. We require expression values to be strictly positive, so let $q = \min_i T_i$ and define
576 $\tilde{T}_i = T_i + q$. Finally, calculate τ as

$$\tau = \frac{1}{N-1} \sum_i^N \frac{1 - \tilde{T}_i}{\tilde{T}_{\max}}$$

577 The set of testis-biased genes was defined as all those with $\tau > 0.5$ and higher expression in testis than
578 in any of the other seventeen tissues in the multi-tissue dataset (PRJEB11897, Harr et al. (2016)). The set of
579 ubiquitously-expressed genes was defined as those with $\tau < 0.25$ and whose expression was above the median
580 expression in the highest-expressing tissue. The set of early-meiosis genes was defined as those with $\tau > 0.5$
581 and highest expression in leptotene/zygotene spermatocytes; spermatid genes were defined as those with $\tau >$
582 0.5 and highest expression in round spermatids. We analyzed expression specificity during spermatogenesis
583 separately in the two intra-subspecific F_1 crosses, and took the union of the resulting gene sets.

584 Acknowledgments

585 The authors thank Jeff Good, Erica Larson, Michael Nachman, Megan Phifer-Rixey, Jacob Mueller, Alyssa
586 Kruger, Marty Ferris, Peter Ellis and additional anonymous reviewers for many insightful comments and
587 suggestions. This work was supported by the National Institutes of Health (F30MH103925, R01HD065024,
588 U19AI100625, U42OD010924).

589 References

- 590 Akaike H. 1978. On the Likelihood of a Time Series Model. *J Roy Stat Soc D-Sta.* 27:217–235.
- 591 Arbiza L, Gottipati S, Siepel A, Keinan A. 2014. Contrasting X-Linked and Autosomal Diversity across 14
592 Human Populations. *Am J Hum Genet.* 94:827–844.

- 593 Auton A, McVean G. 2007. Recombination rate estimation in the presence of hotspots. *Genome Res.* 17:1219–
594 1227.
- 595 Baines JF, Harr B. 2007. Reduced X-Linked Diversity in Derived Populations of House Mice. *Genetics*. 175:1911–
596 1921.
- 597 Beaumont MA, Zhang W, Balding DJ. 2002. Approximate Bayesian computation in population genetics. *Genet-
598 ics*. 162:2025–2035.
- 599 Beck JA, Lloyd S, Hafezparast M, Lennon-Pierce M, Eppig JT, Festing MFW, Fisher EMC. 2000. Genealogies of
600 mouse inbred strains. *Nat Genet.* 24:23–25.
- 601 Bellott DW, Hughes JF, Skaletsky H, et al. (30 co-authors). 2014. Mammalian Y chromosomes retain widely
602 expressed dosage-sensitive regulators. *Nature*. 508:494–499.
- 603 Bellott DW, Skaletsky H, Cho TJ, et al. (16 co-authors). 2017. Avian W and mammalian Y chromosomes conver-
604 gently retained dosage-sensitive regulators. *Nat Genet.* 49:387–394.
- 605 Berta P, Hawkins JB, Sinclair AH, Taylor A, Griffiths BL, Goodfellow PN, Fellous M. 1990. Genetic evidence
606 equating SRY and the testis-determining factor. *Nature*. 348:448–450.
- 607 Bhaskar A, Song YS. 2014. DESCARTES’ RULE OF SIGNS AND THE IDENTIFIABILITY OF POPULATION
608 DEMOGRAPHIC MODELS FROM GENOMIC VARIATION DATA. *Ann Stat.* 42:2469–2493.
- 609 Bhattacharyya T, Gregorova S, Mihola O, Anger M, Sebestova J, Denny P, Simecek P, Forejt J. 2013. Mechanistic
610 basis of infertility of mouse intersubspecific hybrids. *Proc Natl Acad Sci USA*. 110:E468–E477.
- 611 Bishop CE, Boursot P, Baron B, Bonhomme F, Hatat D. 1985. Most classical *Mus musculus domesticus* laboratory
612 mouse strains carry a *Mus musculus musculus* Y chromosome. *Nature*. 315:70–72.
- 613 Brawand D, Soumillon M, Necsulea A, et al. (18 co-authors). 2011. The evolution of gene expression levels in
614 mammalian organs. *Nature*. 478:343–348.
- 615 Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. *Nat
616 Biotech.* 34:525–527.
- 617 Bulatova N, Kotenkova E. 1990. Variants of the Y-chromosome in sympatric taxa of *Mus* in southern USSR.
618 *Bulletino di zoologia*. 57:357–360.
- 619 Burgoyne PS, Mahadevaiah SK, Sutcliffe MJ, Palmer SJ. 1992. Fertility in mice requires X-Y pairing and a Y-
620 chromosomal “Spermiogenesis” gene mapping to the long arm. *Cell*. 71:391–398.

- 621 Campbell P, Good JM, Nachman MW. 2013. Meiotic sex chromosome inactivation is disrupted in sterile hybrid
622 male house mice. *Genetics*. 193:819–828.
- 623 Campbell P, Nachman MW. 2014. X-y interactions underlie sperm head abnormality in hybrid male house mice.
624 *Genetics*. 196:1231–1240.
- 625 Case LK, Wall EH, Osmanski EE, Dragon JA, Saligrama N, Zachary JF, Lemos B, Blankenhorn EP, Teuscher
626 C. 2015. Copy number variation in Y chromosome multicopy genes is linked to a paternal parent-of-origin
627 effect on CNS autoimmune disease in female offspring. *Genome Biol*. 16:28.
- 628 Charlesworth B, Coyne JA, Barton NH. 1987. The Relative Rates of Evolution of Sex Chromosomes and Auto-
629 somes. *Am Nat*. 130:113–146.
- 630 Cocquet J, Ellis PJI, Mahadevaiah SK, Affara NA, Vaiman D, Burgoyne PS. 2012. A Genetic Basis for a Postmei-
631 otic X Versus Y Chromosome Intragenomic Conflict in the Mouse. *PLOS Genet*. 8:e1002900.
- 632 Cocquet J, Ellis PJI, Yamauchi Y, Mahadevaiah SK, Affara NA, Ward MA, Burgoyne PS. 2009. The Multicopy
633 Gene Sly Represses the Sex Chromosomes in the Male Mouse Germline after Meiosis. *PLOS Biol*. 7:e1000244.
- 634 Cocquet J, Ellis PJI, Yamauchi Y, Riel JM, Karacs TPS, Rattigan Á, Ojarikre OA, Affara NA, Ward MA, Burgoyne
635 PS. 2010. Deficiency in the Multicopy Sycp3-Like X-Linked Genes Slx and Slxl1 Causes Major Defects in
636 Spermatid Differentiation. *Mol Biol Cell*. 21:3497–3505.
- 637 Consortium CC. 2012. The Genome Architecture of the Collaborative Cross Mouse Genetic Reference Popula-
638 tion. *Genetics*. 190:389–401.
- 639 Conway SJ, Mahadevaiah SK, Darling SM, Capel B, Rattigan AM, Burgoyne PS. 1994. Y353/B: a candidate
640 multiple-copy spermiogenesis gene on the mouse Y chromosome. *Mamm Genome*. 5:203–210.
- 641 Cortez D, Marin R, Toledo-Flores D, Froidevaux L, Liechti A, Waters PD, Grützner F, Kaessmann H. 2014.
642 Origins and functional evolution of Y chromosomes across mammals. *Nature*. 508:488–493.
- 643 Csilléry K, François O, Blum MGB. 2012. abc: an R package for approximate Bayesian computation (ABC).
644 *Methods Ecol Evol*. 3:475–479.
- 645 Derome N, Métayer K, Montchamp-Moreau C, Veuille M. 2004. Signature of Selective Sweep Associated With
646 the Evolution of sex-ratio Drive in *Drosophila simulans*. *Genetics*. 166:1357–1366.
- 647 Din W, Anand R, Boursot P, Darviche D, Dod B, Jouvin-Marche E, Orth A, Talwar G, Cazenave PA, Bonhomme
648 F. 1996. Origin and radiation of the house mouse: clues from nuclear genes. *J Evol Biol*. 9:519–539.

- 649 Doran AG, Wong K, Flint J, Adams DJ, Hunter KW, Keane TM. 2016. Deep genome sequencing and variation
650 analysis of 13 inbred mouse strains defines candidate phenotypic alleles, private variation and homozygous
651 truncating mutations. *Genome Biol.* 17:167.
- 652 Drost JB, Lee WR. 1995. Biological basis of germline mutation: Comparisons of spontaneous germline mutation
653 rates among drosophila, mouse, and human. *Environ Mol Mutagen.* 25:48–64.
- 654 Egan CM, Sridhar S, Wigler M, Hall IM. 2007. Recurrent DNA copy number variation in the laboratory mouse.
655 *Nat Genet.* 39:1384–1389.
- 656 Eicher EM, Hutchison KW, Phillips SJ, Tucker PK, Lee BK. 1989. A repeated segment on the mouse Y chromo-
657 some is composed of retroviral-related, Y-enriched and Y-specific sequences. *Genetics.* 122:181–192.
- 658 Ellegren H. 2011. Sex-chromosome evolution: recent progress and the influence of male and female heteroga-
659 metry. *Nat Rev Genet.* 12:157–166.
- 660 Ellis PJI, Bacon J, Affara NA. 2011. Association of Sly with sex-linked gene amplification during mouse evolu-
661 tion: a side effect of genomic conflict in spermatids? *Hum Mol Genet.* 20:3010–3021.
- 662 Ellis PJI, Clemente EJ, Ball P, Touré A, Ferguson L, Turner JMA, Loveland KL, Affara NA, Burgoyne PS. 2005.
663 Deletions on mouse Yq lead to upregulation of multiple X- and Y-linked transcripts in spermatids. *Hum Mol*
664 *Genet.* 14:2705–2715.
- 665 Fischer M, Kosyakova N, Liehr T, Dobrowolski P. 2016. Large deletion on the Y-chromosome long arm (Yq) of
666 C57bl/6jbomtac inbred mice. *Mamm Genome.* pp. 1–7.
- 667 Forejt J. 1996. Hybrid sterility in the mouse. *Trends Genet.* 12:412–417.
- 668 Forejt J, Iványi P. 1974. Genetic studies on male sterility of hybrids between laboratory and wild mice (*Mus*
669 *musculus* L.). *Genet Res.* 24:189–206.
- 670 Geraldes A, Basset P, Gibson B, Smith KL, Harr B, Yu HT, Bulatova N, Ziv Y, Nachman MW. 2008. Inferring the
671 history of speciation in house mice from autosomal, X-linked, Y-linked and mitochondrial genes. *Mol Ecol.*
672 17:5349–5363.
- 673 Good JM, Dean MD, Nachman MW. 2008. A Complex Genetic Basis to X-Linked Hybrid Male Sterility Between
674 Two Species of House Mice. *Genetics.* 179:2213–2228.
- 675 Good JM, Giger T, Dean MD, Nachman MW. 2010. Widespread Over-Expression of the X Chromosome in
676 Sterile F1 Hybrid Mice. *PLOS Genet.* 6:e1001148.

- 677 Good JM, Nachman MW. 2005. Rates of Protein Evolution Are Positively Correlated with Developmental Tim-
678 ing of Expression During Mouse Spermatogenesis. *Mol Biol Evol.* 22:1044–1052.
- 679 Graves JAM. 2006. Sex Chromosome Specialization and Degeneration in Mammals. *Cell.* 124:901–914.
- 680 Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD. 2009. Inferring the Joint Demographic His-
681 tory of Multiple Populations from Multidimensional SNP Frequency Data. *PLOS Genet.* 5:e1000695.
- 682 Haller BC, Messer PW. 2017. SLiM 2: Flexible, Interactive Forward Genetic Simulations. *Mol Biol Evol.* 34:230–
683 240.
- 684 Halligan DL, Kousathanas A, Ness RW, Harr B, Eöry L, Keane TM, Adams DJ, Keightley PD. 2013. Contribu-
685 tions of Protein-Coding and Regulatory Change to Adaptive Molecular Evolution in Murid Rodents. *PLOS
686 Genet.* 9:e1003995.
- 687 Harr B, Karakoc E, Neme R, et al. (19 co-authors). 2016. Genomic resources for wild populations of the house
688 mouse, *Mus musculus* and its close relative *Mus spretus*. *Sci Data.* 3:160075.
- 689 Hendriksen PJ, Hoogerbrugge JW, Themmen AP, Koken MH, Hoeijmakers JH, Oostra BA, van der Lende T,
690 Grootegoed JA. 1995. Postmeiotic transcription of X and Y chromosomal genes during spermatogenesis in
691 the mouse. *Dev Biol.* 170:730–733.
- 692 Hudson RR, Kaplan NL. 1995. The Coalescent Process and Background Selection. *Philos Trans R Soc Lond B Biol
693 Sci.* 349:19–23.
- 694 Hughes JF, Page DC. 2015. The Biology and Evolution of Mammalian Y Chromosomes. *Ann Rev Genet.* 49:507–
695 527.
- 696 Hvilsom C, Qian Y, Bataillon T, et al. (17 co-authors). 2012. Extensive X-linked adaptive evolution in central
697 chimpanzees. *Proc Natl Acad Sci USA.* 109:2054–2059.
- 698 Itsara A, Wu H, Smith JD, Nickerson DA, Romieu I, London SJ, Eichler EE. 2010. De novo rates and selection
699 of large copy number variation. *Genome Res.* 20:1469–1481.
- 700 Jaenike J. 2001. Sex Chromosome Meiotic Drive. *Annu Rev Ecol Syst.* 32:25–49.
- 701 Keane TM, Goodstadt L, Danecek P, et al. (41 co-authors). 2011. Mouse genomic variation and its effect on
702 phenotypes and gene regulation. *Nature.* 477:289–294.
- 703 Kingan SB, Garrigan D, Hartl DL. 2010. Recurrent Selection on the Winters sex-ratio Genes in *Drosophila
704 simulans*. *Genetics.* 184:253–265.

- 705 Korneliussen TS, Albrechtsen A, Nielsen R. 2014. ANGSD: Analysis of Next Generation Sequencing Data. *BMC
706 Bioinformatics*. 15:356.
- 707 Kousathanas A, Halligan DL, Keightley PD. 2014. Faster-X Adaptive Protein Evolution in House Mice. *Genetics*.
708 196:1131–1143.
- 709 Lahn BT, Page DC. 1997. Functional Coherence of the Human Y Chromosome. *Science*. 278:675–680.
- 710 Larson EL, Keeble S, Vanderpool D, Dean MD, Good JM. 2016a. The composite regulatory basis of the large
711 X-effect in mouse speciation. *Mol Biol Evol*. p. msw243.
- 712 Larson EL, Vanderpool D, Keeble S, Zhou M, Sarver BAJ, Smith AD, Dean MD, Good JM. 2016b. Contrasting
713 Levels of Molecular Evolution on the Mouse X Chromosome. *Genetics*. 203:1841–1857.
- 714 Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv:1303.3997
715 [q-bio]*. ArXiv: 1303.3997.
- 716 Liu KJ, Steinberg E, Yozzo A, Song Y, Kohn MH, Nakhleh L. 2015. Interspecific introgressive origin of genomic
717 diversity in the house mouse. *Proc Natl Acad Sci USA*. 112:196–201.
- 718 MacBride MM, Navis A, Dasari A, Perez AV. 2017. Mild reproductive impact of a Y chromosome deletion on a
719 C57bl/6j substrain. *Mamm Genome*. 28:155–165.
- 720 Macholán M, Baird SJ, Munclinger P, Dufková P, Bímová B, Piálek J. 2008. Genetic conflict outweighs heteroga-
721 metic incompatibility in the mouse hybrid zone? *BMC Evol Biol*. 8:271.
- 722 McDonald JH, Kreitman M. 1991. Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature*. 351:652–
723 654.
- 724 McKenna A, Hanna M, Banks E, et al. (11 co-authors). 2010. The Genome Analysis Toolkit: a MapReduce
725 framework for analyzing next-generation DNA sequencing data. *Genome Res*. 20:1297–1303.
- 726 McLaren A, Simpson E, Epplen JT, Studer R, Koopman P, Evans EP, Burgoyne PS. 1988. Location of the genes
727 controlling H-Y antigen expression and testis determination on the mouse Y chromosome. *Proc Natl Acad Sci
728 USA*. 85:6442–6445.
- 729 Moretti C, Vaiman D, Tores F, Cocquet J. 2016. Expression and epigenomic landscape of the sex chromosomes
730 in mouse post-meiotic male germ cells. *Epigenetics Chromatin*. 9:47.
- 731 Morgan AP, Didion JP, Doran AG, Holt JM, McMillan L, Keane TM, Villena FPMd. 2016a. Genome Report:
732 Whole Genome Sequence of Two Wild-Derived *Mus musculus domesticus* Inbred Strains, LEWES/EiJ and
733 ZALENDE/EiJ, with Different Diploid Numbers. *G3*. p. g3.116.034751.

- 734 Morgan AP, Gatti DM, Najarian ML, Keane TM, Galante RJ, Pack AI, Mott R, Churchill GA, Villena FPMd. 2017.
735 Structural Variation Shapes the Landscape of Recombination in Mouse. *Genetics*. 206:603–619.
- 736 Morgan AP, Holt JM, McMullan RC, et al. (12 co-authors). 2016b. The Evolutionary Fates of a Large Segmental
737 Duplication in Mouse. *Genetics*. 204:267–285.
- 738 Mueller JL, Mahadevaiah SK, Park PJ, Warburton PE, Page DC, Turner JMA. 2008. The mouse X chromosome
739 is enriched for multicopy testis genes showing postmeiotic expression. *Nat Genet*. 40:794–799.
- 740 Mueller JL, Skaletsky H, Brown LG, Zaghlul S, Rock S, Graves T, Auger K, Warren WC, Wilson RK, Page DC.
741 2013. Independent specialization of the human and mouse X chromosomes for the male germ line. *Nat Genet*.
742 45:1083–1087.
- 743 Nagamine CM, Nishioka Y, Moriwaki K, Boursot P, Bonhomme F, Lau YFC. 1992. The musculus-type Y Chro-
744 mosome of the laboratory mouse is of Asian origin. *Mamm Genome*. 3:84–91.
- 745 Nam K, Munch K, Hobolth A, et al. (85 co-authors). 2015. Extreme selective sweeps independently targeted the
746 X chromosomes of the great apes. *Proc Natl Acad Sci USA*. 112:6413–6418.
- 747 Neme R, Tautz D. 2016. Fast turnover of genome transcription across evolutionary time exposes entire non-
748 coding DNA to de novo gene emergence. *eLife*. 5:e09977.
- 749 Nishioka Y, Lamothe E. 1986. Isolation and characterization of a mouse Y chromosomal repetitive sequence.
750 *Genetics*. 113:417–432.
- 751 Oh B, Hwang SY, Solter D, Knowles BB. 1997. Spindlin, a major maternal transcript expressed in the mouse
752 during the transition from oocyte to embryo. *Development*. 124:493–503.
- 753 Orth A, Belkhir K, Britton-Davidian J, Boursot P, Benazzou T, Bonhomme F. 2002. Natural hybridization be-
754 tween 2 sympatric species of mice, *Mus musculus domesticus* L. and *Mus spretus* Lataste. *C R Biol*. 325:89–97.
- 755 Payseur BA, Krenz JG, Nachman MW. 2004. Differential Patterns of Introgression across the X Chromosome in
756 a Hybrid Zone between Two Species of House Mice. *Evolution*. 58:2064–2078.
- 757 Polanski A, Szczesna A, Garbulowski M, Kimmel M. 2017. Coalescence computations for large samples drawn
758 from populations of time-varying sizes. *PLOS One*. 12:e0170701.
- 759 Pool JE, Nielsen R. 2007. Population Size Changes Reshape Genomic Patterns of Diversity. *Evolution*. 61:3001–
760 3006.

- 761 Pritchard JK, Seielstad MT, Perez-Lezaun A, Feldman MW. 1999. Population growth of human Y chromosomes:
762 a study of Y chromosome microsatellites. *Mol Biol Evol.* 16:1791–1798.
- 763 Repping S, van Daalen SKM, Brown LG, et al. (11 co-authors). 2006. High mutation rates have driven extensive
764 structural polymorphism among human Y chromosomes. *Nat Genet.* 38:463–467.
- 765 Rice WR. 1984. Sex Chromosomes and the Evolution of Sexual Dimorphism. *Evolution.* 38:735–742.
- 766 Salcedo T, Geraldes A, Nachman MW. 2007. Nucleotide variation in wild and inbred mice. *Genetics.* 177:2277–
767 2291.
- 768 Sarver BAJ, Keeble S, Cosart T, Tucker PK, Dean MD, Good JM. 2017. Phylogenomic Insights into Mouse
769 Evolution Using a Pseudoreference Approach. *Genome Biol Evol.* 9:726–739.
- 770 Sayres MAW, Lohmueller KE, Nielsen R. 2014. Natural Selection Reduced Diversity on Human Y Chromo-
771 somes. *PLOS Genet.* 10:e1004064.
- 772 Scally A. 2016. Mutation rates and the evolution of germline structure. *Philos Trans R Soc Lond B Biol Sci.*
773 371:20150137.
- 774 Simmons LW, Fitzpatrick JL. 2012. Sperm wars and the evolution of male fertility. *Reproduction.* 144:519–534.
- 775 Smith NGC, Eyre-Walker A. 2002. Adaptive protein evolution in Drosophila. *Nature.* 415:1022–1024.
- 776 Soh YQS, Alföldi J, Pyntikova T, et al. (21 co-authors). 2014. Sequencing the mouse Y chromosome reveals
777 convergent gene acquisition and amplification on both sex chromosomes. *Cell.* 159:800–813.
- 778 Song Y, Endepols S, Kleemann N, Richter D, Matuschka FR, Shih CH, Nachman MW, Kohn MH. 2011. Adaptive
779 Introgression of Anticoagulant Rodent Poison Resistance by Hybridization between Old World Mice. *Curr
780 Biol.* 21:1296–1301.
- 781 Spiess AN, Walther N, Müller N, Balvers M, Hansis C, Ivell R. 2003. SPEER—A New Family of Testis-Specific
782 Genes from the Mouse. *Biol Reprod.* 68:2044–2054.
- 783 Srivastava A, Morgan AP, Najarian ML, et al. (18 co-authors). 2017. Genomes of the Mouse Collaborative Cross.
784 *Genetics.* 206:537–556.
- 785 Storchová R, Gregorová S, Buckiová D, Kyselová V, Divina P, Forejt J. 2004. Genetic analysis of X-linked hybrid
786 sterility in the house mouse. *Mamm Genome.* 15:515–524.
- 787 Styryna J, Klag J, Moriwaki K. 1991. Influence of partial deletion of the Y chromosome on mouse sperm pheno-
788 type. *J Reprod Fertil.* 92:187–195.

- 789 Teeter KC, Payseur BA, Harris LW, Bakewell MA, Thibodeau LM, O'Brien JE, Krenz JG, Sans-Fuentes MA,
790 Nachman MW, Tucker PK. 2008. Genome-wide patterns of gene flow across a house mouse hybrid zone.
791 *Genome Res.* 18:67–76.
- 792 Tishkoff SA, Reed FA, Ranciaro A, et al. (19 co-authors). 2007. Convergent adaptation of human lactase persis-
793 tence in Africa and Europe. *Nat Genet.* 39:31–40.
- 794 Torgerson DG, Singh RS. 2003. Sex-linked mammalian sperm proteins evolve faster than autosomal ones. *Mol
795 Biol Evol.* 20:1705–1709.
- 796 Touré A, Clemente EJ, Ellis P, Mahadevaiah SK, Ojarikre OA, Ball PAF, Reynard L, Loveland KL, Burgoyne PS,
797 Affara NA. 2005. Identification of novel Y chromosome encoded transcripts by testis transcriptome analysis
798 of mice with deletions of the Y chromosome long arm. *Genome Biol.* 6:R102.
- 799 Touré A, Szot M, Mahadevaiah SK, Rattigan Á, Ojarikre OA, Burgoyne PS. 2004. A New Deletion of the Mouse
800 Y Chromosome Long Arm Associated With the Loss of Ssty Expression, Abnormal Sperm Development and
801 Sterility. *Genetics.* 166:901–912.
- 802 Tu S, Shin Y, Zago WM, States BA, Eroshkin A, Lipton SA, Tong GG, Nakanishi N. 2007. Takusan: A Large
803 Gene Family that Regulates Synaptic Activity. *Neuron.* 55:69–85.
- 804 Tucker PK, Lee BK, Lundrigan BL, Eicher EM. 1992. Geographic origin of the Y chromosomes in "old" inbred
805 strains of mice. *Mamm Genome.* 3:254–261.
- 806 Turner LM, Schwahn DJ, Harr B. 2012. Reduced male fertility is common but highly variable in form and
807 severity in a natural house mouse hybrid zone. *Evolution.* 66:443–458.
- 808 Turner LM, White MA, Tautz D, Payseur BA. 2014. Genomic Networks of Hybrid Sterility. *PLOS Genet.*
809 10:e1004162.
- 810 Uchimura A, Higuchi M, Minakuchi Y, Ohno M, Toyoda A, Fujiyama A, Miura I, Wakana S, Nishino J, Yagi T.
811 2015. Germline mutation rates and the long-term phenotypic effects of mutation accumulation in wild-type
812 laboratory mice and mutator mice. *Genome Res.* 25:1125–1134.
- 813 Veeramah KR, Gutenkunst RN, Woerner AE, Watkins JC, Hammer MF. 2014. Evidence for Increased Levels
814 of Positive and Negative Selection on the X Chromosome versus Autosomes in Humans. *Mol Biol Evol.*
815 31:2267–2282.
- 816 Webster TH, Wilson Sayres MA. 2016. Genomic signatures of sex-biased demography: progress and prospects.
817 *Curr Opin Genet Devel.* 41:62–71.

- 818 Yakimenko LV, Korobitsyna KV, Frisman LV, Muntianu AI. 1990. Cytogenetic and biochemical comparison of
819 *Mus musculus* and *Mus hortulanus*. *Experientia*. 46:1075–1077.
- 820 Yamauchi Y, Riel JM, Stoytcheva Z, Burgoyne PS, Ward MA. 2010. Deficiency in mouse Y chromosome long
821 arm gene complement is associated with sperm DNA damage. *Genome Biol.* 11:R66.
- 822 Yamauchi Y, Riel JM, Wong SJ, Ojarikre OA, Burgoyne PS, Ward MA. 2009. Live Offspring from Mice Lacking
823 the Y Chromosome Long Arm Gene Complement. *Biol Reprod.* 81:353–361.
- 824 Yanai I, Benjamin H, Shmoish M, et al. (12 co-authors). 2005. Genome-wide midrange transcription profiles
825 reveal expression level relationships in human tissue specification. *Bioinformatics*. 21:650–659.
- 826 Yang H, Wang JR, Didion JP, et al. (15 co-authors). 2011. Subspecific origin and haplotype diversity in the
827 laboratory mouse. *Nat Genet.* 43:648–655.

828

Supplement

829

Caveats to demographic models

830

Like all models, the scenarios of neutral demography and natural selection presented here are greatly simplified and almost certainly wrong. We chose to consider the history of each of the three subspecies independently, rather than in a joint isolation-with-migration or isolation-by-distance model, because the one-population models could be fit more robustly and more easily checked against analytical formulae. The simple models are nonetheless useful as guideposts along the way to a better approximation of the true history of the mouse sex chromosomes. On the basis of autosomal SFS, we can confidently reject scenarios with a single very sharp ($f \ll 0.1$) reduction in population size with or without unequal sex ratio as the sole explanation for lack of Y-linked diversity. Likewise we can rule out exponential growth alone, because it actually increases X:A and Y:A relative to a stationary model (Figure S3C). One important possibility that we have not considered is a fluctuating sex ratio. An “arms race” between the X and Y chromosomes for transmission in the male germline could lead to oscillations between a male-biased and female-biased population. Even if deviations in the sex ratio are transient, the net effect could be to reduce diversity on both sex chromosomes relative to autosomes out of proportion to the strength of selection.

843

Disentangling the effects of demography and selection on the Y chromosome is especially challenging because the Y has the smallest effective population size and is inherited without recombination, so it is most susceptible to changes in population size, to background selection and to selective sweeps. We have used ABC to show that Y-linked SFS are consistent with recent positive selection. Of course the fact that selective sweeps offer plausible fit, conditional on neutral demographic history, does not rule out background selection. Both mammalian Y and avian W chromosomes, which have independent evolutionary origins, retain a convergent set of dosage-sensitive genes with roles in core cellular processes (Bellott et al. 2017). The proportion of substitutions in these genes fixed by positive selection (α_{SEW} , Figure 8) is indeed much smaller than for X-linked genes. Together these results imply relatively strong purifying selection on ancestral Y genes, which in the absence of recombination should constrain diversity on the entire chromosome. This is an important alternative hypothesis to a hitchhiking effect associated with positive selection on Yq.

854

Supplementary tables

855

File S1. Raw VCF files with genotype calls on Y chromosome and mitochondrial genome, available from the Zenodo repository: <https://doi.org/10.5281/zenodo.817658>.

857

Table S1. List of samples used in this study (Excel spreadsheet). Sample manifest provided in first tab, column key in second tab.

859

Table S2. Y chromosome haplogroup assignment and CNV status for 69 Collaborative Cross strains.

860

Table S3. Sequence diversity statistics across different classes of sites on the autosomes, X and Y chromosomes, by population. See main text for details. L , total number of callable bases in target locus; θ_π , Tajima's pairwise θ ; θ_w , Watterson's θ ; D , Tajima's D ; D_{FL} , Fu and Li's D . Both estimators of θ are expressed as percentages with bootstrap standard errors in parentheses.

864

Supplementary figures

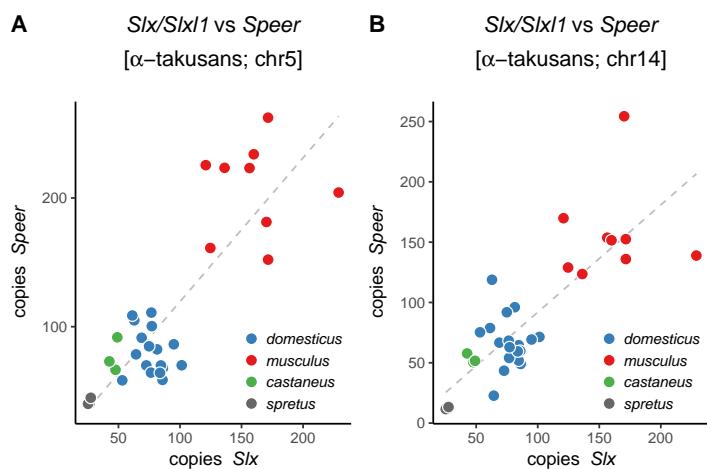


Figure S1: Copy number of *Speer* family members on chromosomes 5 (A) and 14 (B) compared to copy number of *Slx/Slxl1*.

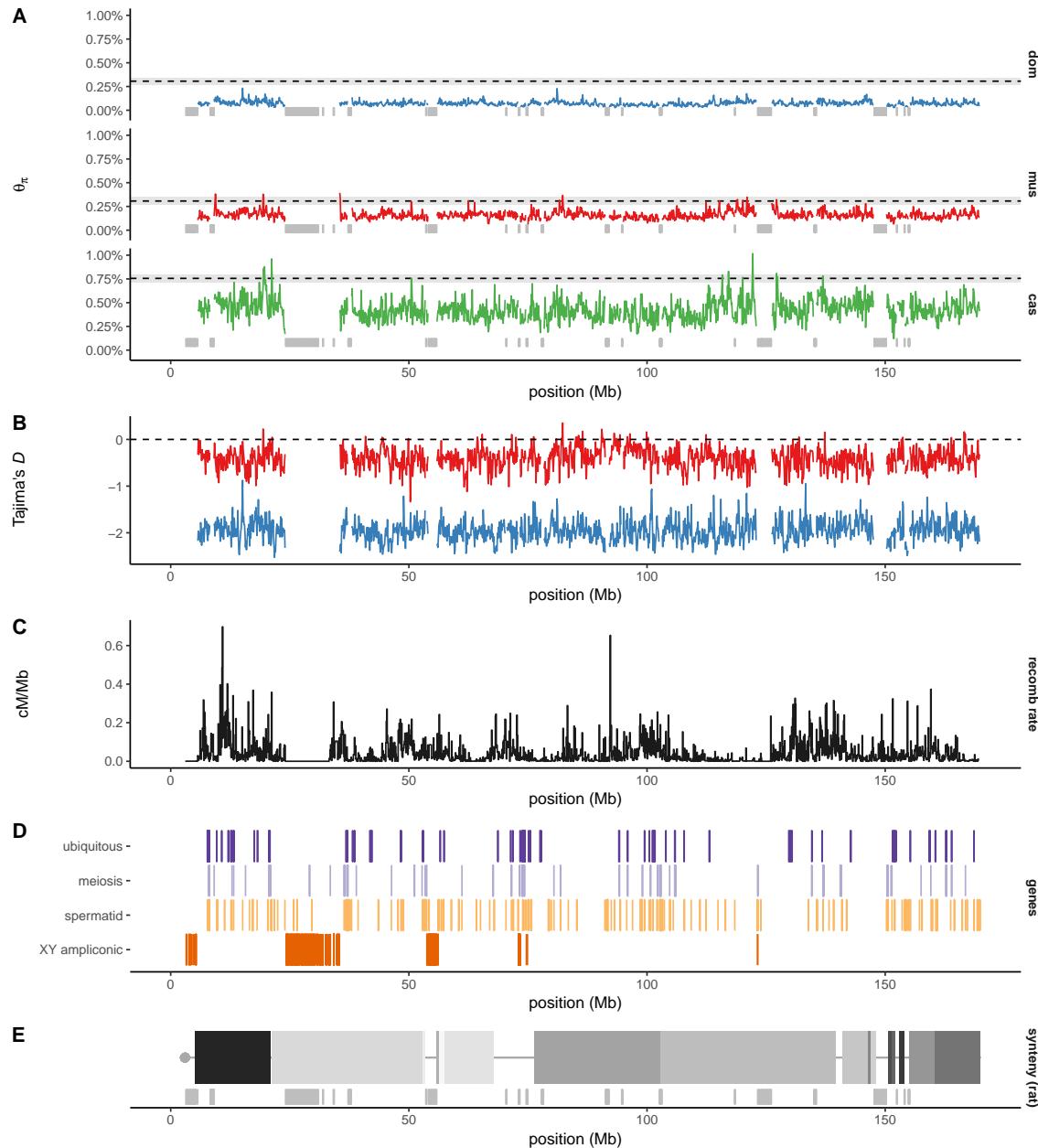


Figure S2: Global view of sequence diversity on the X chromosome. (A) Pairwise diversity in 100 kb windows by population, with expectation based on autosomal diversity shown as dashed line with 95% confidence band. Ampliconic regions from Mueller et al. (2008) were masked and are indicated with grey bars. (B) Tajima's D for *domesticus* and *musculus*, with masking as in panel A. (C) Recombination rate (centimorgans per Mb) estimated in the Diversity Outbred stock. (D) Locations of gene sets used elsewhere in this manuscript, plus X-Y ampliconic genes from Soh et al. (2014), as sketched in Figure 1. (E) Blocks of conserved synteny with rat. Boundaries between blocks define rearrangement breakpoints between mouse and the common rodent ancestor.

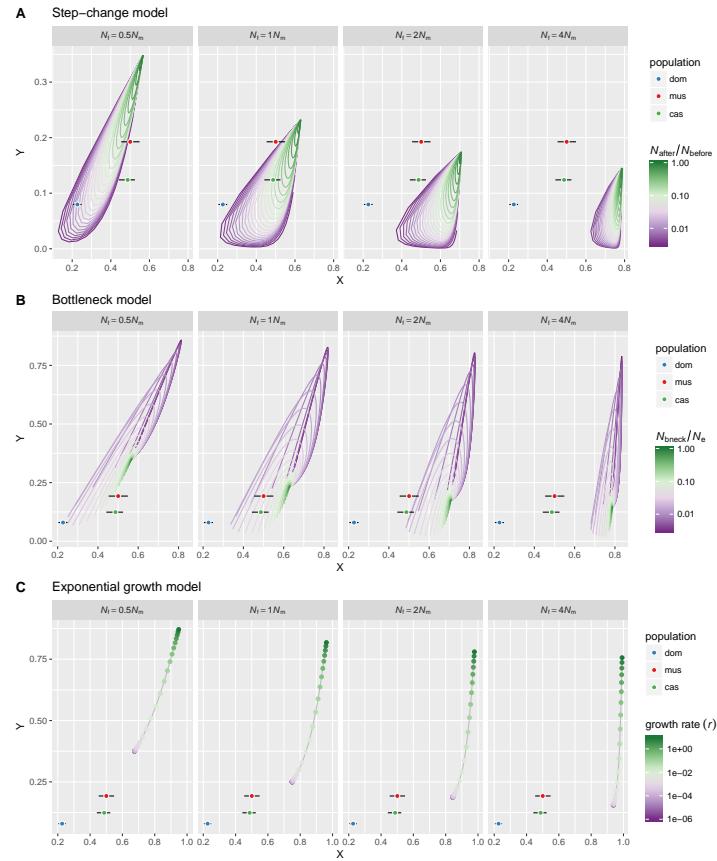


Figure S3: X:A and Y:A diversity ratios under several neutral coalescent models with varying sex ratio. Error bars represent bootstrap standard errors.

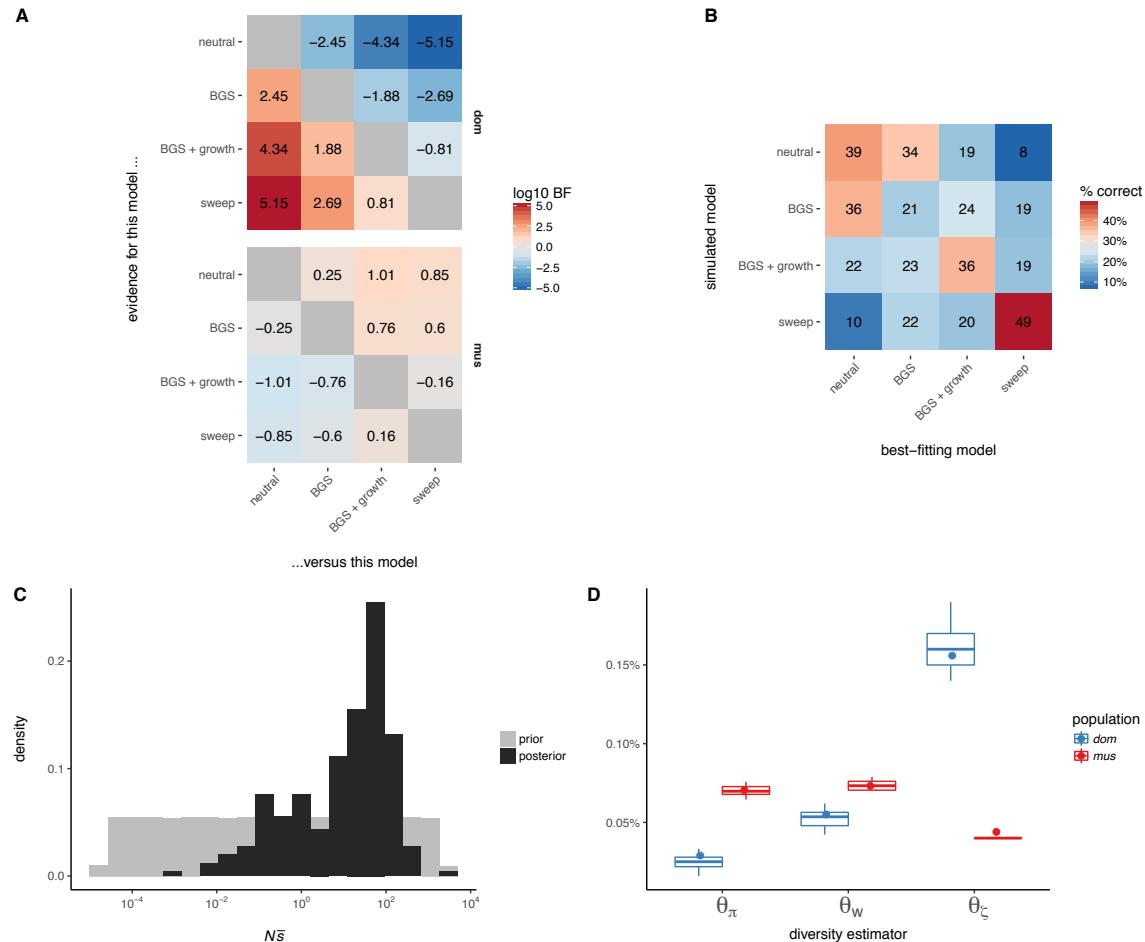


Figure S4: Modelling selection on the Y chromosome. **(A)** Pairwise comparison of goodness-of-fit using Bayes factors. **(B)** Percent recall of model-selection procedure estimated by leave-one-out cross-validation. **(C)** Prior and posterior distributions of the population-scaled mean selection coefficient $\gamma = N\bar{s}$ for the selective-sweep model in *domesticus*. **(D)** Posterior distribution of diversity statistics from the best-fitting model in each population (boxplots), compared to their observed values.

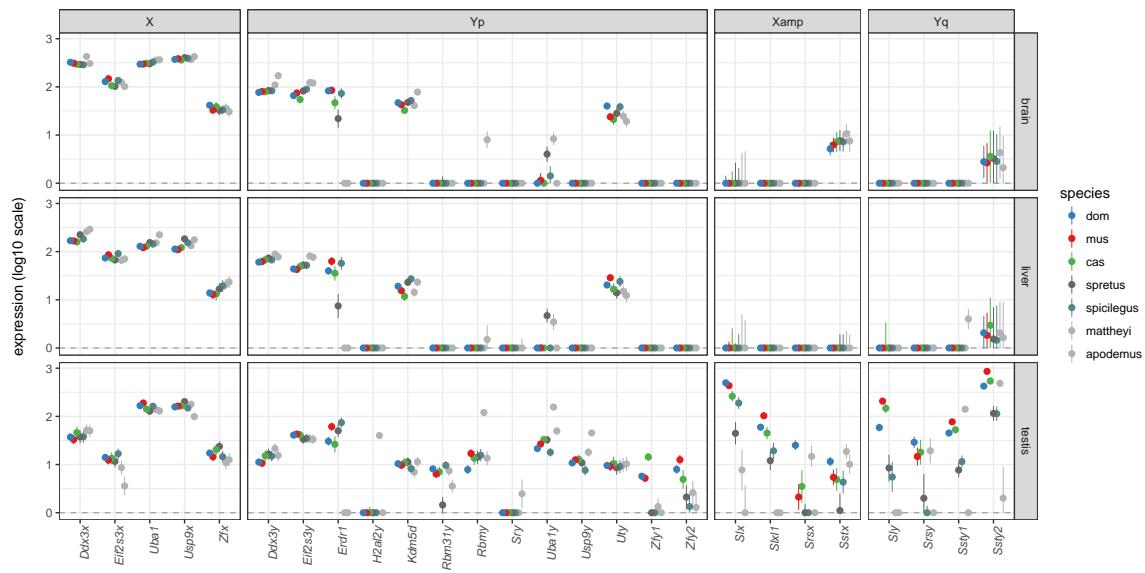


Figure S5: Expression of selected X- and Y-linked genes across *Mus* species. Genes are subdivided by chromosomal location: X, non-ampliconic X-linked genes with Y-linked homologs; Yp, non-ampliconic genes on Yp; Xamp, X-linked homologs of co-amplified genes; Yq, Y-linked homologs of co-amplified genes, residing on Yq.