Sequence and structural diversity of mouse Y chromosomes

Andrew P Morgan and Fernando Pardo-Manuel de Villena*

Department of Genetics and Lineberger Comprehensive Cancer Center

University of North Carolina, Chapel Hill, NC

*Corresponding author:

5049C Genetic Medicine Building

Department of Genetics

University of North Carolina

Chapel Hill NC 27599-7264

fernando@med.unc.edu

Keywords: sex chromosome evolution, intragenomic conflict, mouse evolution

Abstract

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 linage 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). Several lines of evidence suggest that the evolution of the acquired genes is driven by intragenomic conflict with the X chromosome for transmission to progeny (Ellis et al. 2011; Cocquet et al. 2012).

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 hinted that it consisted of mostly of repetitive sequences, with copy number in the hundreds, and that it was evolving rapidly (Nishioka and Lamothe 1986; Eicher et al. 1989). Unlike other mammalian Y chromosomes, which are dominated by large blocks of heterochromatin (Hughes and Page 2015), the mouse Y chromosome was also known to be large and almost entirely euchromatic. Spontaneous

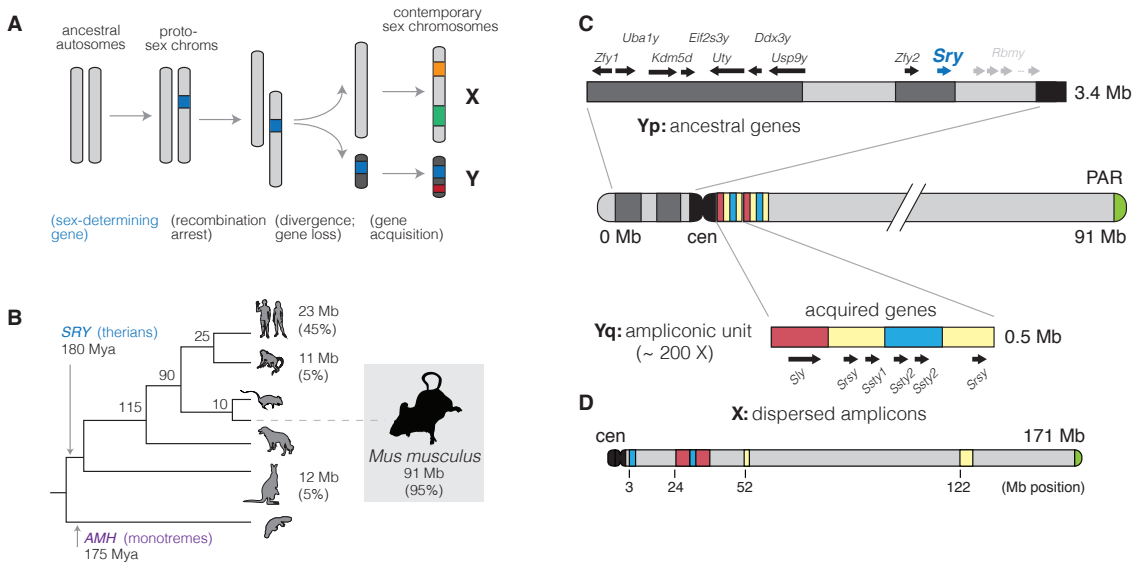

Figure 1: Evolution of mammalian Y chromosomes. (**A**) Evolution of heteromorphic sex chromosomes. (**B**) Y chromosomes of mammals. The Y chromosome of therian mammals, characterized by the sex-determining factor *SRY*, diverged from the X chromosome approximately 180 Mya. (The monotremata have a different 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 choromosome, 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.

mutations in laboratory stocks allowed the mapping of male-specific tissue antigens and the sex-determining factor *Sry* to the short arm of the chromosome (Yp) (McLaren et al. 1988), while lesions on the long arm (Yq) were associated with infertility and defects in spermatogenesis (Styrna et al. 1991; Burgoyne et al. 1992; Touré et al. 2004).

Sequencing, assembly and annotation of the mouse Y chromosome in the inbred strain C57BL/6J was finally completed in 2014 after more than a decade of painstaking effort (Soh et al. 2014). Ancestral genes are restricted to Yp and are fewer in number on the mouse Y chromosome than in other studied mammals. Yq was shown to consist of approximately 200 copies of a 500 kb unit — the "huge repeat array" — containing the acquired genes *Sly*, *Ssty1*, *Ssty2* and *Srsy* (**Figure 1C**). *Sly* and its X-linked homologs *Slx* and *Slxl1* are found only in the genus *Mus* and have sequence similarity to the synaptonemal complex protein SYCP3 (Ellis et al. 2011). *Ssty1/2* and *Sstx* are most similar to members of the spindlin family (Oh et al. 1997) and are present in taxa at least as phylogenetically distant as rats. The coding potential of *Srsy* and *Srsx* is unclear, but they have sequence similarity to melanoma-related cancer/testis antigens typified by the human MAGEA family. Their phylogenetic origins remain unresolved. The genes of the huge repeat array are expressed almost exclusively in post-meiotic round spermatids and function in chromatin condensation and sperm maturation (Burgoyne et al. 1992; Touré et al. 2004,

2005; Yamauchi et al. 2009, 2010).

Independent amplification of homologous genes on the X and Y chromosomes is thought to be a byproduct of competition between the X and Y chromosomes for transmission to the next generation. The current consensus favors an unidentified X-linked sex-ratio distorter whose action is suppressed by one or more Y-linked factors (Ellis et al. 2011). Consistent with this hypothesis, the *Sly* and *Slx* families act in opposing directions to maintain or relieve transcriptional silencing of the sex chromosomes after meiosis (post-meiotic sex chromosome repression, PSCR) (Hendriksen et al. 1995; Cocquet et al. 2012). Underexpression of *Sly* (via knockdown of *Sly*) in the testis results in sex ratio distortion in favor of females; the reverse is true for knockdown of *Slx*. In both case sex-ratio distortion is accompanied by defects in sperm morphology (Touré et al. 2004; Cocquet et al. 2009, 2010). Disruption of PSCR and the related process of meiotic sex chromosome inactivation (MSCI) is also associated with male sterility in inter-subspecific hybrids between *M. m. domesticus* and *M. m. musculus* (Campbell et al. 2013). Together these observations suggest that the intragenomic conflict between the sex chromosomes in mouse is played out in post-meiotic spermatids and may have mechanistic overlap with hybrid male sterility.

Intragenomic conflict can have a profound impact on the genetic diversity of sex chromosomes in natural populations. Sex-ratio-distorter systems in *Drosophila* provide the best known examples (Jaenike 2001; Derome et al. 2004; Kingan et al. 2010). The extent to which diversity on mouse sex chromosomes is influenced by intragenomic conflict remains an open question. The differential impact of selection on mouse X chromosome versus autosomes (the "faster-X" effect) is well-studied, mostly through the lens of speciation (Torgerson and Singh 2003; Kousathanas et al. 2014; Larson et al. 2016a,b). Larson et al. (2016b) used pairwise comparisons between wild-derived strains of *M. m. musculus* and *M. m. domesticus* to show that the "faster-X" effect is most prominent for genes expressed primarily in the testis and early in spermatogenesis (before MSCI). This set of genes is also prone to dysregulation in sterile hybrids Larson et al. (2016a). By contrast, selective pressures imposed by intragenomic conflict between the sex chromosomes should be exerted in spermatids after the onset of PSCR. Genes with spermatid-specific expression are expected to respond most rapidly, while those with broad expression are expected to be constrained by putative functional requirements in other tissues or cell types.

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

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

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

Results

A survey of Y-linked coding variation in mouse

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

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

One group of inbred strains in our dataset — C57BL/6J (reference genome), C57BL/10J, C57L/J and C57BR/cdJ — have a known common ancestor in the year 1929, and a common ancestor with the strain C58/J in 1915 (Beck et al. 2000). Assuming an average of three generations per year, the total branch length of the pedigree connecting the C57 and C58 strains is $5,280$ generations, during which time 3 mutations occurred. We 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 the effect of 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). 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 to western European populations. *M. m. castaneus* harbors two distinct and paraphyletic lineages: one corresponding to the Indian subcontinent and another represented only by the wild-derived inbred strain CAST/EiJ (from Thailand.) The latter haplogroup probably corresponds to the southeast Asian lineage identified in previous reports (Geraldes et al. 2008; Yang et al. 2011).

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

Copy-number variation is pervasive on the Y chromosome

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

The highly repetitive content of Yq precludes a similarly detailed characterization of copy-number variation along this chromosome arm. However, we can estimate the copy number of each of the three gene families present (*Sly*, *Ssty1/2* and *Srsy*) by counting the total number of reads mapped to each and normalizing for sequencing depth. The hypothesis of X-Y

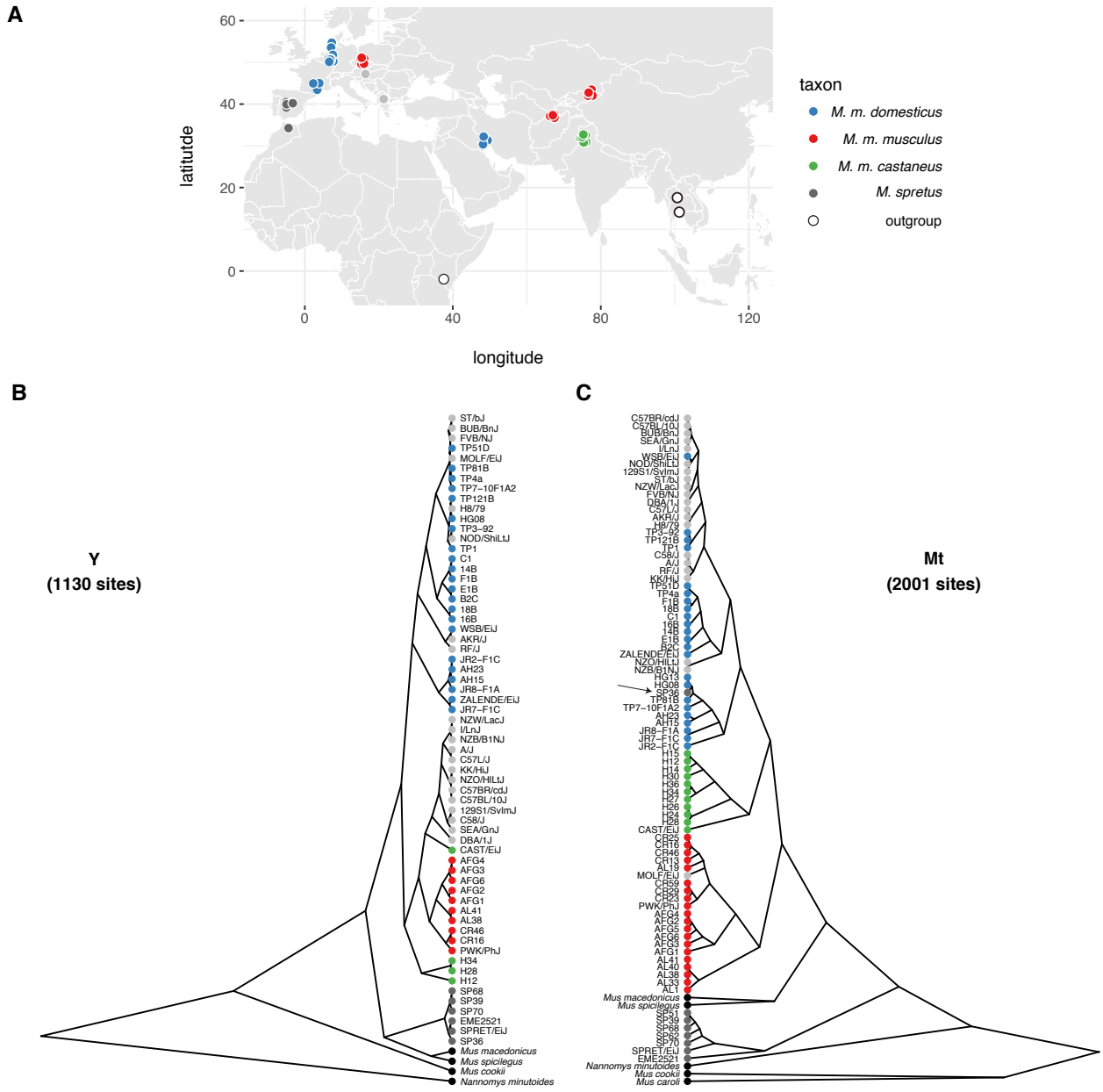

Figure 2: Patrilineal and matrilineal phylogeography in a geographically-diverse sample from the genus *Mus*. (**A**) Sampling locations of mice used in this study. (**B**) Phylogenetic tree from coding sites on the Y chromosome. Samples are colored according to nominal ancestry; laboratory strains are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

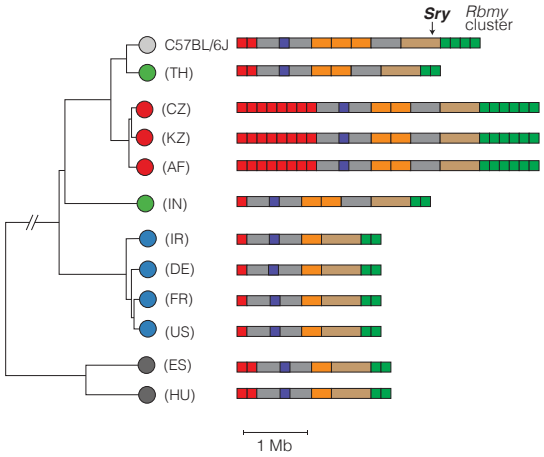

Figure 3: Copy-number variation on Yp. Schematic view of copy-number variable regions of the Y chromosome short arm (Yp) superposed on SNV-based phylogenetic tree. All CNVs shown overlap a segmental duplication in the reference sequence (strain C57BL/6J). One CNV overlaps a known protein-coding gene: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

intragenomic conflict predicts that, if expression levels are at least roughly proportional to copy number, amplification of gene families on Yq should be countered by amplification of their antagonistic homologs on the X chromosome (or vice versa.) We tested this hypothesis by comparing the copy number of X- and Y-linked homologs of the *Slx/y*, *Sstx/y* and *Srsx/y* families in wild mice. **Figure 4** shows that copy number on X and Y chromosomes are indeed correlated for *Slx/y*. The relationship between *Slx*-family and *Sly*-family copy number is almost exactly linear (slope = 0.98 [95% CI $0.87 - 1.09$]; $R^2 = 0.87$). We note that samples are not phylogenetically independent, so the statistical significance of the regression is exaggerated, but the qualitative result clearly supports previous evidence that conflict between X and Y chromosomes is mediated primarily through *Slx* and *Sly* (Cocquet et al. 2012).

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

The scale of copy number change within the *M. musculus* lineage suggests a high underlying mutation rate. We used whole-genome sequence data from a panel of 69 recombinant inbred lines (RILs) from the Collaborative Cross (CC; Srivastava et al. (2017)) to estimate the rate of copy-number change on Yq. Each CC line is independently derived from eight inbred founder strains via two generations of outcrossing followed by sibling matings until inbreeding is achieved (Consortium 2012). Distinct CC lines inheriting a Y chromosome from the same founder strain thus share an obligate male ancestor in the recent past, but no more recently than the start of inbreeding (**Figure 5A**). We estimated read depth in 100 kb bins across Yq, and normalized each bin against the median for CC lines inheriting a Y chromosome from the same founder strain. This normalization effectively removes noise from mapping of short reads to repetitive sequence, and uncovers megabase-sized

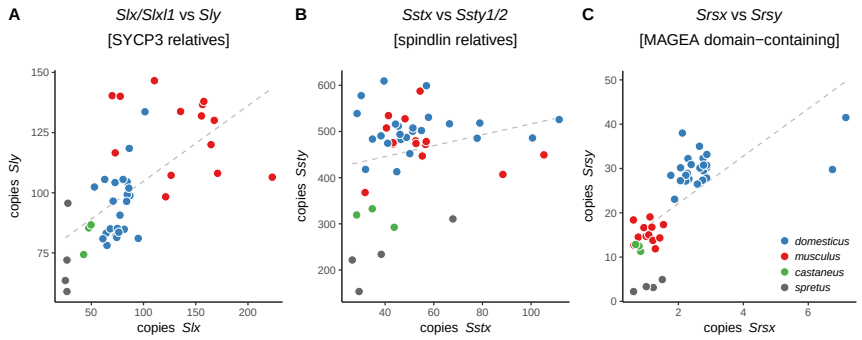

Figure 4: Approximate copy number of co-amplified gene families on X and Yq. Each dot represents a single individual. Grey dashed line is simple linear regression of Y-linked versus X-linked copy number.

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/HlLtJ	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.

CNVs in 5 CC lines carrying 3 different Y chromosomes (**Table 2**, **Table S2** and **Figure 5B**). Because the pedigree of each CC line is known, mutation rates — for each Y haplogroup, and overall — can be estimated directly, assuming each new allele corresponds to a single mutational event. Our estimate of 0.30 (95% Poisson CI 0.098 − 0.70) mutations per 100 father-son transmissions is about tenfold higher than ampliconic regions of the human Y chromosome (Repping et al. 2006), and places the mouse Yq among the most labile sequences known in mouse or human (Egan et al. 2007; Itsara et al. 2010; Morgan et al. 2016b).

Sequence diversity is markedly reduced on both sex chromosomes

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

We estimated nucleotide diversity in four classes of sites: intergenic sites (putatively neutral); introns; 4-fold degenerate

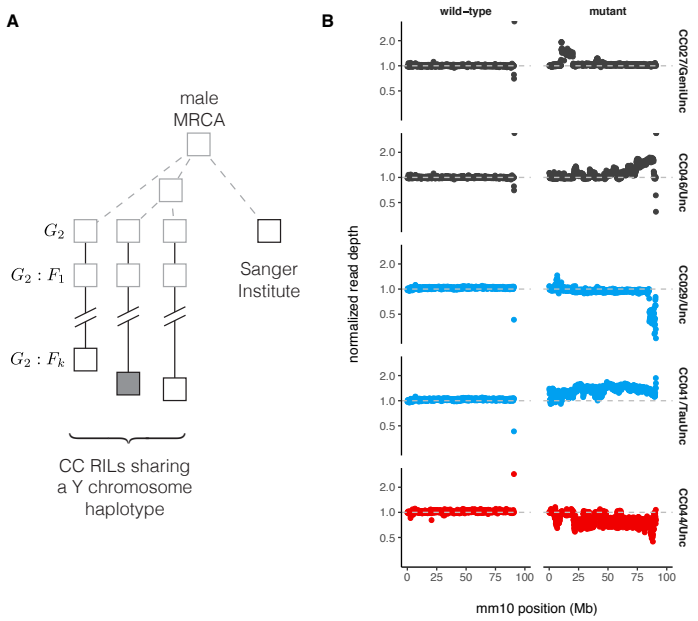

Figure 5: Copy-number variation on Yq in the Collaborative Cross. (**A**) Pedigree-based estimates of mutation rate on the Y chromosome long arm (Yq). Multiple recombinant inbred lines (RILs) from the Collaborative Cross (CC) panel share the same Y chromosome haplotype, with (filled shape) or without (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.

Table 3: Sequence diversity statistics across different classes of sites on the autosomes, X and Y chromosomes, by population. See main text for details. L, total size in bp; 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.

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

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

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

Sex chromosomes are affected differently than autosomes by both neutral forces, such as changes in population size (Pool and Nielsen 2007), and by natural selection (reviewed in *eg.* Ellegren (2011)). The X chromosomes of humans (Arbiza et al. 2014), several other primate species (Nam et al. 2015) are substantially less diverse than the demographic histories of these

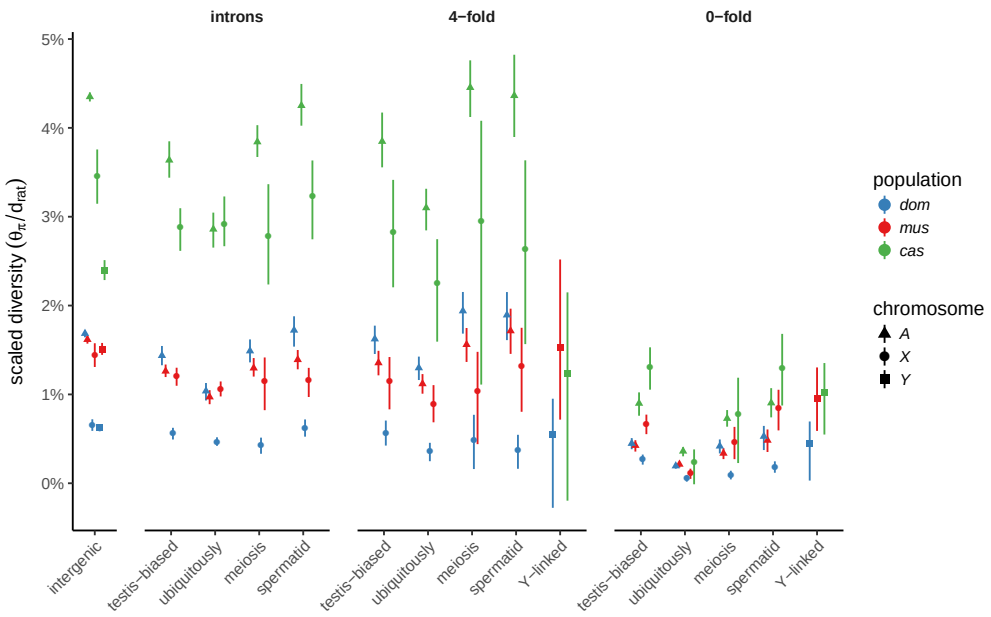

Figure 6: Scaled nucleotide diversity by population, site class and chromosome type. First panel from left shows estimates from intergenic sequence; remaining panels are site classes within protein-coding gene boundaries.

			Population		
Comparison	**Expected**	**Scaling**	*dom*	*mus*	*cas*
X:A	$3/4$	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	$1/4$	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 4: 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.

species would predict as a result of both purifying selection and recurrent selective sweeps. For humans, the pattern extends to the Y chromosome (Sayres et al. 2014). Having observed a deficit of polymorphism on both sex chromosomes in mouse, the central question arising in this paper is: to what extent is sex-chromosome diversity reduced by natural selection? A rich body of literature already exists for the influence of selection on the mouse X chromosome, especially in the context of speciation (Good et al. 2008; Teeter et al. 2008; Baines and Harr 2007; Kousathanas et al. 2014; Larson et al. 2016a,b), so we directed our focus to the lesser-studied Y chromosome.

To establish an appropriate null against which to test hypotheses about natural selection on the sex chromosomes, we followed an approach similar to Sayres et al. (2014). We fit four simple demographic models to SFS from putatively-neutral intergenic sites on the autosomes using the maximum-likelihood framework implemented in $\partial a \partial i$ (Gutenkunst et al. 2009) (**Figure 7A**). Each model is parameterized by an initial effective population size (N_0), a size change (expressed as fraction f of starting size for models involving instantaneous size changes, or the ending population size N_e for the exponential growth models), and a time of onset of size change (τ). The relative fit of each model was quantified using the method of Aikake weights (Akaike 1978). All four models can be viewed as nested in the family of three-epoch, piecewise-exponential histories.

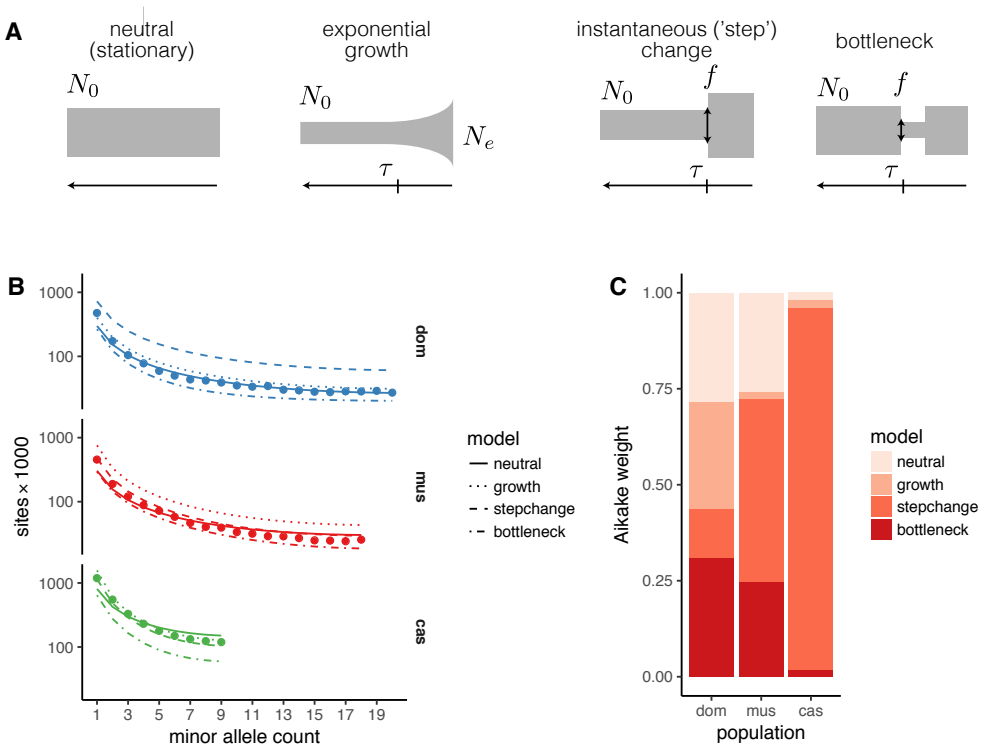

Figure 7: Inference of demographic histories from autosomal sites. (**A**) Four simple demographic models fit with $\partial a \partial i$. Each model is parameterized by one or more of an ancestral effective population size (N_0), time of population-size change (τ), change in population size as fraction 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.

In principle, such models are identifiable with sample size of $4 \times 3 = 12$ or more chromosomes (6 diploid individuals) (Bhaskar and Song 2014). In practice, more than one model fits each population about equally well (or equally poorly), with the exception of *M. m. castaneus* , which is best described by the "step-change" model (**Figure 7B-C**). Of course the true history of each population is almost certainly more complex than any of our models. Our goal is not to obtain a comprehensive description of mouse population history as such, but rather to pick an appropriate null model against which we can test hypotheses about selection. For *domesticus*, the stationary model is the most parsimonious; for *musculus*, the exponential-growth model.

We also compared our estimates of sex chromosome diversity to predictions from coalescent theory (Pool and Nielsen 2007; Polanski et al. 2017) for the four models considered above. For this analysis we focused on the ratios X:A and Y:A, which are independent of autosomal effective population size. Results are shown in **Figure S2**, with observed X:A and Y:A ratios superposed. We summarize some relevant trends here and refer to previous reviews (Pool and Nielsen 2007; Webster and Wilson Sayres 2016) for further details. Qualitatively, both X:A and Y:A are reduced after an instantaneous contraction in population size, eventually recovering to their stationary values after about $4N_e$ generations. For a bottleneck — a contraction followed by instantaneous recovery to the initial size — X:A and Y:A are at first sharply reduced and then increased relative to a stationary population, again returning to stationary values after about $4N_e$ generations. With exponential growth, X:A

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 5: 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.

and Y:A are actually increased relative to their stationary values. These patterns are modulated by the breeding sex ratio; X:A increases and Y:A decreases when females outnumber males, and vice versa. In brief, some combination of a male-biased breeding ratio and a very strong ($f \ll 0.1$) population contraction would be required to explain the observed reductions in X:A and Y:A in *domesticus*, with somewhat milder effects required to explain the reduction in *musculus* or *castaneus*. These histories are not consistent with population histories inferred from autosomal SFS. We hypothesize that this discrepancy is explained, at least in part, by selection.

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

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

Second, we used forward simulations from the models fit to autosomal SFS to explore the possible contribution of natural selection to the SFS of Y chromosomes. We simulated two modes of selection independently: purifying selection on linked deleterious alleles (background selection, BGS; (Hudson and Kaplan 1995)), and hard selective sweeps on newly-arising

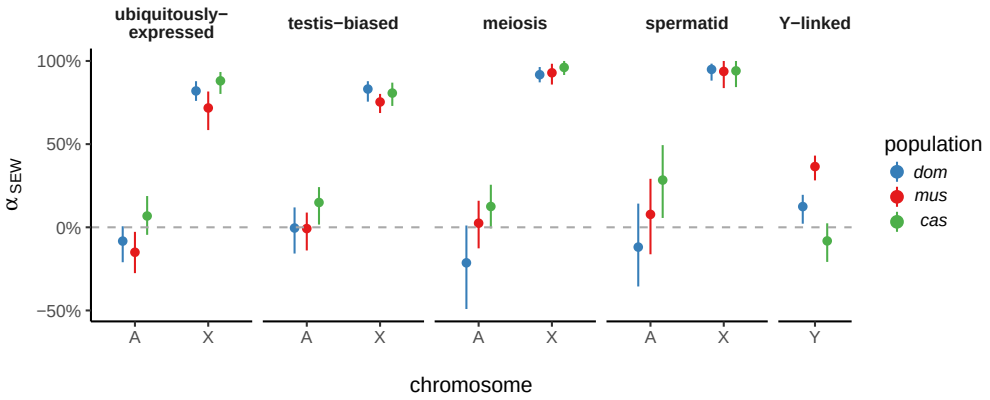

Figure 8: Proportion of sites fixed by positive selection (using the non-parametric estimator of Smith and Eyre-Walker (2002), α_{SEW}) according to gene-expression class, chromosome and population. Error bars represent 95% bootstrap CIs.

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

Table 6: 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.

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 9**. The Y chromosomes of *domesticus* are best approximated by the selective-sweep model. For *musculus* the result is less clear: the model with an exponentially-growing population and BGS has the best fit, but only marginally so. Furthermore, 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 9B**). 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 6**). 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 9D**). We note that, because the Y chromosome is inherited without

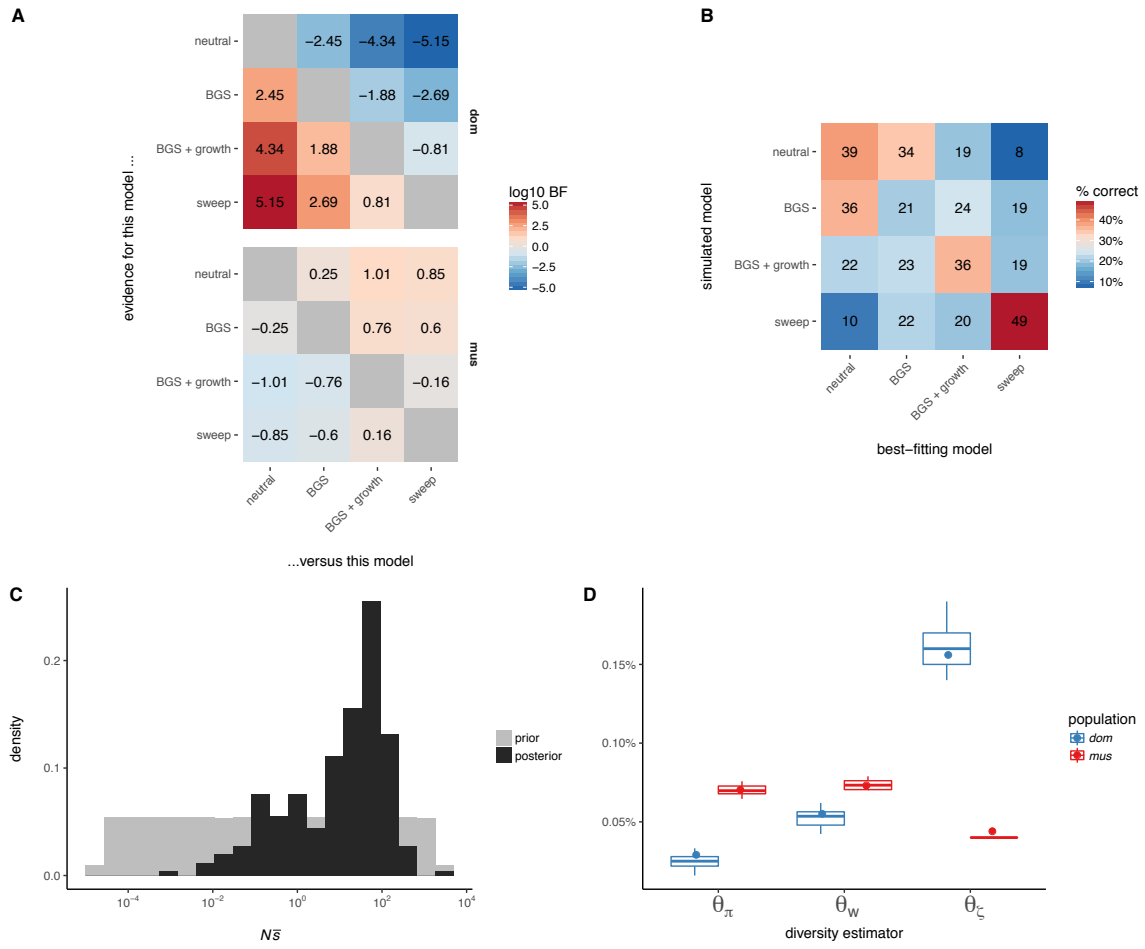

Figure 9: Modelling selection on the Y chromosome. (**A**) Pairwise comparison of goodness-of-fit using Bayes factors. (**B**) Percent recall of model-selection procedure estimated by leave-one-out cross-validation. (**C**) Prior and posterior distributions of the population-scaled mean selection coefficient $\gamma = N\bar{s}$ 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.

recombination, our estimate of $N\bar{s}$ reflects the cumulative selection intensity on the entire chromosome and not necessarily on a single site.

Sex-linked gene expression diverges rapidly in the testis

Given the dramatic differences in Y-linked gene content between even closely-related *Mus* taxa, we finally asked whether patterns of gene expression showed similar divergence. In particular, we sought to test the prediction that expression patterns of Y-linked genes diverge more rapidly than autosomal genes in the testis. To that end we re-analyzed published gene expression data from the brain, liver and testis of wild-derived outbred individuals representing seven (sub)species spanning an 8 million year evolutionary transect across the murid rodents (Neme and Tautz 2016) (**Figure 10A**). For genes on the autosomes and X chromosome, the great majority of expression variance lies between tissues rather than between (sub)species (PC1 and PC2, cumulative 77.1% of variance explained; **Figure 10B**). For Y-linked genes, highly enriched for function in the male germline, most variance (PC1, 59.6% of variance explained) naturally lies between the testis and the non-germline tissues.

To quantify divergence in gene expression patterns we computed the rank-correlation (Spearman's ρ) between species for each tissue type separately for autosomal, X-linked and Y-linked genes, and constructed trees by neighbor-joining (**Figure 10C**). We use total tree length as an estimator of expression divergence. The topology of these trees for the autosomes and X chromosome in brain and testis is consistent with known phylogenetic relationships within the Muridae. Consistent with previous comparative analyses of gene expression in mammals (Brawand et al. 2011), we find that expression patterns are most constrained in brain and least constrained in testis (**Figure 10D**). Expression divergence is equal between autosomes and X chromosome in brain and liver, but greater for X-linked genes in testis. Y-linked expression diverges much more rapidly in all three tissues, but the effect is most extreme in the testis. We caution that the precision of these estimates is limited by the small number of Y-linked relative to autosomal or X-linked genes.

This "faster-X" effect should be limited to functional elements subject to male-specific selection. Genes expressed in the male germline (testis-biased or expressed during spermatogenesis) might be enriched for such elements, relative to genes with ubiquitous expression. We therefore estimated expression divergence in autosomal, X- and Y-linked genes with four sets of genes with different expression patterns (**Figure 10E**). X-linked expression diverges more rapidly than autosomal expression only among genes with testis-biased expression. Curiously, the set of genes with the strongest "faster-X" effect at the sequence level — those expressed early in male meiosis — has no such effect at the expression level (**Figure 10F**). The number of Y-linked genes in each group is too small to permit any strong conclusions.

We attempted to bracket the date of origin of each of the ampliconic gene families on Yq under the assumption that, if present, they would be expressed at detectable levels (**Figure S3**). *Sly* is present only in the Palearctic clade (*M. musculus*, *M. spretus* and *M. spicilegus*). Its homolog *Slxl1* is restricted to the same clade, but *Slx* may be older as it is expressed — albeit at a lower level — in *Nannomys*. *Sstx* and *Ssty1/2* are present in all species, consistent with a previous report that they are ancestral to the mouse-rat divergence (Ellis et al. 2011). The distribution of *Srsx* and *Srsy* is more difficult to interpret: they are expressed in *M. musculus* and *Nannomys* but not in *M. spretus*, *M. spicilegus* or *Apodemus*. Since they are not present in rat,

the most parsimonious explanation is that they originated along the branch subtending the common ancestor of *Mus* and *Nannomys* and were independently lost in *M. spretus* and *M. spicilegus*.

Discussion

We have shown that nucleotide diversity is reduced on both sex chromosomes relative to expectations for a stationary population, and that the effect appears strongest in *domesticus* and weakest in *musculus* (**Table 4**). Sex differences in the long-term average mutation rate, estimated from synonymous-sites divergence to rat, are not sufficient to explain the deficit. Because sex chromosomes respond differently than autosomes to changes in population size, we fit several (simple) models of demographic history to autosomal site-frequency spectra (**Figure 7**) and compared their predictions to observed values. At least for the models we considered, neither gradual nor instantaneous changes in population size — of magnitude feasible given autosomal SFS — can account for the reduction in diversity on both the X and Y chromosomes, even if we allow for a sex ratio different than 1:1 (**Figure S2**). Estimates of effective size of each population (from autosomal sites) are in agreement with previous work on house mice (Din et al. 1996; Baines and Harr 2007; Salcedo et al. 2007; Geraldes et al. 2008).

Using demographic histories from autosomes as a baseline, we simulated two modes of selection — background selection and hard selective sweeps — on Y chromosomes of *domesticus* and *musculus*. Although discrimination between models was limited by both technical factors and theoretical constraints, we have shown that the Y-linked SFS in *domesticus* is consistent with a moderately strong selective sweep (**Figure 9**). The hard sweep model is also the best-fitting in *musculus*, but is only 1.4-fold more likely ($\log_{10}$ BF $= 0.16$) than the next-best model. We conclude that recent positive selection accounts, at least in part, for the reduction in Y-linked relative to autosomal diversity in *domesticus* and likely also in *musculus*.

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 S2C**). One important possibility that we have not considered is a fluctuating sex ratio. As we discuss later, 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.

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

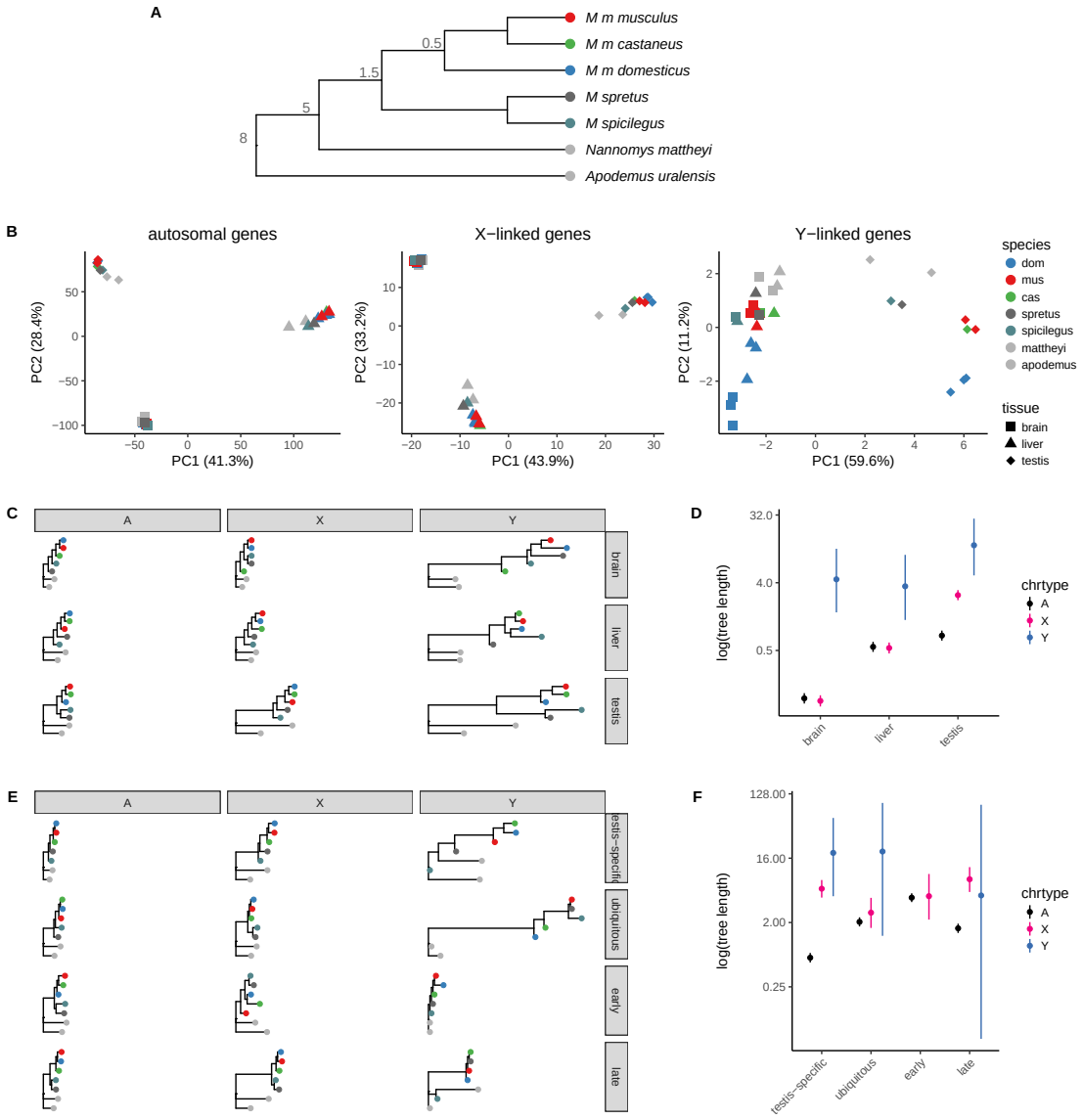

Figure 10: Divergence of sex-linked gene expression in murid rodents. (**A**) Schematic phylogeny of taxa in the multi-tissue expression dataset. Node labels are approximate divergence times (Mya); branch lengths not to scale. (**B**) Projection of samples onto the top two principal components of expression 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.

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 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 constrains diversity on the entire chromosome.

The mouse X chromosome has been studied intensively in the context of speciation (Forejt and Iványi 1974; Forejt 1996). Several major hybrid sterility loci have been mapped on the X chromosome in inter-subspecific crosses (Storchová et al. 2004; Good et al. 2008; Campbell et al. 2013; Turner et al. 2014), and it has been shown that gene flow is reduced on the X chromosome relative to the autosomes in the *domesticus-musculus* hybrid zone (Payseur et al. 2004; Teeter et al. 2008). Naturally-occurring hybrid males, like laboratory-bred hybrids, exhibit a wide range of reproductive defects (Turner et al. 2012). Using wild-derived inbred strains as proxies for natural populations, Larson et al. (2016b) showed that X-linked genes have, on average, faster rates of evolution than autosomal genes. The "faster-X" effect is corroborated by a previous analysis of the same *castaneus* mice used in this manuscript (Kousathanas et al. 2014). Taken together, this literature supports the idea that selection is pervasive on the mouse X chromosome. Both purifying selection and strong selective sweeps have similarly been proposed as explanations for the paucity of X-linked diversity in great apes (Hvilsom et al. 2012; Veeramah et al. 2014; Nam et al. 2015).

Although we did not attempt to apply the ABC scheme to X chromosomes, we did inspect patterns of diversity across the X chromosome in relation to recombination rate, spermatogenesis genes, and blocks of conserved synteny with rat (**Figure 11**). Diversity is reduced across the entire X chromosome in all three populations, in marked contrast to local "troughs" observed in great apes. Regression of pairwise diversity (θ_π) on distance away from ubiquitously-expressed genes, meiosis genes, spermatid genes, and X-Y ampliconic genes was significant only in *musculus* for ubiquitously expressed genes ($t = 6.6$, Bonferoni-corrected $p = 6.8 \times 10^{-11}$). Similarly — and surprisingly — there was no relationship ($t = -1.2$, $p = 0.23$) between sequence diversity and recombination rate at 100 kb resolution, as estimated from the Diversity Outbred mouse stock (Morgan et al. 2017). (We speculate that characterizing recombination at finer scale from linkage disequilibrium (Auton and McVean 2007) would provide a more powerful test.) Two explanations or these patterns are possible. First, X-linked diversity could be reduced primarily by neutral processes, which should act chromosome-wide. Second, the targets of natural selection could be sufficiently dispersed that independent signatures of selection blur together. We favor the latter scenario, given the proportion of X-linked substitutions apparently fixed by selection (**Figure 8**). A larger sample size would provide better power for an explicit haplotype-based scan for selective sweeps.

To what extent are the properties of the mouse sex chromosomes a consequence of intragenomic conflict? One of the most striking feature of the mouse Y chromosome is its "huge repeat array," consisting of hundreds of copies of the rodent-specific gene families *Sly*, *Ssty1/2* and *Srsy* (Soh et al. 2014). Each family has one or more X-linked homologs that also exist in many copies (Li et al. 2013). SLY acts directly to maintain transcriptional repression of post-meiotic sex chromatin by recruiting a suite of repressive histone marks (Ellis et al. 2005; Cocquet et al. 2009; Moretti et al. 2016); its action is opposed by SLX and SLXL1, and imbalance between SLY and SLX/SLXL1 leads to spermiogenesis defects and distorts the progeny sex

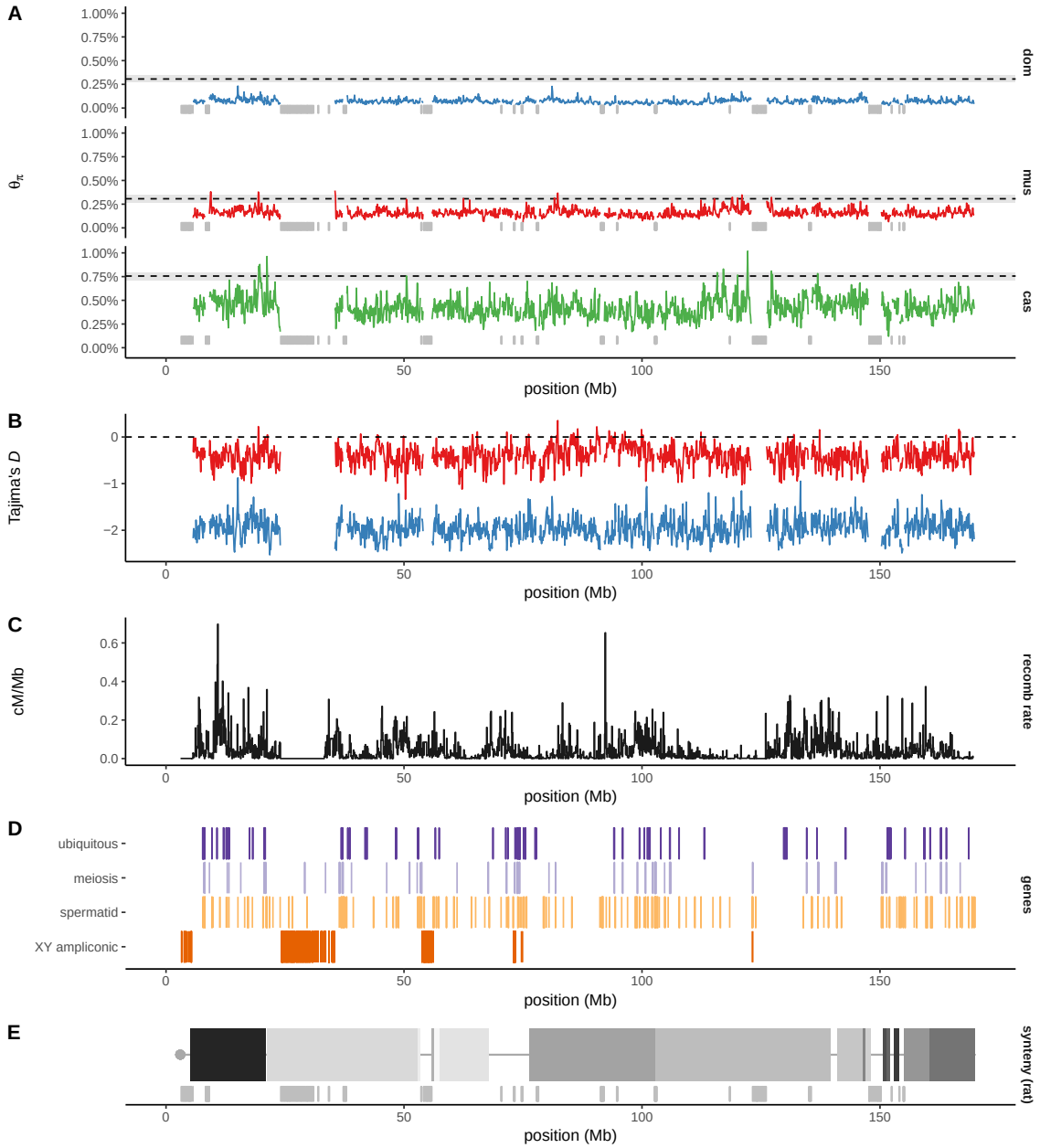

Figure 11: Global view of sequence diversity on the X chromosome. (**A**) Pairwise diversity in 100 kb windows by population, with expectation based on autosomal diversity shown as dashed line with 95% confidence band. Ampliconic regions from Mueller et al. (2008) were masked and are indicated with grey 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.

ratio in favor of the overexpressing chromosome in laboratory crosses (Cocquet et al. 2012). (Additional Y genes, including SSTY, may potentiate the action of SLY (Comptour et al. 2014).) This establishes the conditions for an intragenomic conflict between the X and Y chromosomes that implicates not only SLX/SLXL1 and SLY but *any* gene whose expression after meiosis is required or beneficial for sperm maturation and fertilizing ability. X-linked alleles that meet the functional requirement for post-meiotic expression in the face of repression by SLY — via a stronger promoter, a more stable transcript, a more active protein product, or increased copy number — should be favored by selection (Ellis et al. 2011). The same should be true, in reverse, for successful Y chromosomes.

Copy-number amplification appears to be an important mode of response to the X-Y conflict, facilitated by a high mutation rate (**Table 2**). Multi-copy genes with post-meiotic expression are highly enriched on the X chromosome (Mueller et al. 2008, 2013). We have shown that the copy number of *Slx/Slxl1* and *Sly* have increased three-fold within *M. musculus* and are correlated across populations, consistent with an "arms race" between the sex chromosomes (**Figure 4**). The pattern extends to two clusters of autosomal genes (from the *Speer* or α-takusan family) also expressed in round spermatids (**Figure S1**). Larson et al. (2016a) have shown that, in reciprocal F_1 hybrids between *domesticus* and *musculus* that are "mismatched" for *Slx/Slxl1* and *Sly*, X-linked expression is decreased when *Sly* > *Slx/Slxl1* and decreased when *Sly* < *Slx/Slxl1*. In contrast to (Ellis et al. 2011), we find that the absolute expression of ampliconic X genes and their Yq homologs (in whole testis) does increase with copy number across *Mus* (**Figure S3**). We interpret the correlation between X- and Y-linked copy number, the directional increase in copy number in the *Mus* genus, and consequent differences in gene expression as evidence of ongoing intragenomic conflict in spermatids.

We have also used simulations to show that strong selective sweeps are a plausible explanation for the low diversity of Y chromosomes, at least in *domesticus*. Although we cannot directly identify the putative targets of selection on the Y chromosome, we speculate that they lie in the ampliconic region of Yq and are involved in intragenomic conflict with the X chromosome. Several independent deletions of Yq in laboratory stocks converge on a similar phenotype, namely low fertility, abnormal sperm morphology and sex-ratio distortion in favor of females (Styrna et al. 1991; Conway et al. 1994; Touré et al. 2004; Fischer et al. 2016). Additional evidence for intragenomic conflict in natural populations of mice comes the observation that *musculus* Y chromosomes — the subspecies with highest *Sly* copy number — are more successful at introgressing across *domesticus-musculus* hybrid zone in Europe, and in localities where they do, the sex ratio is shifted towards males (Macholán et al. 2008). However, a major role for X-Y conflict in male hybrid sterility has been confidently excluded (Campbell et al. 2013; Campbell and Nachman 2014).

Many open questions remain with respect to the evolution of mouse Y chromosomes. How many of the hundreds of copies in each gene family retain coding potential? Which copies are functionally equivalent? What evolutionary trade-offs does success in the sex-chromosome conflict entail, in the context of sperm competition and polyandry in natural populations (Simmons and Fitzpatrick 2012)? All of these are important avenues of future study as we seek to understand the forces shaping sex chromosomes.

Materials and methods

Alignment and variant-calling

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

SNVs and short indels on Y were ascertained using `GATK HaplotypeCaller` v3.3-0-g37228af in the intersection of callable regions and exons targeted by the Roche NimbleGen exome-capture array, lifted over to mm10 with `CrossMap` v0.2.3 and the mm9-to-mm10 chain file from the UCSC Genome 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 with a "homozygous" (*ie.* single-copy hemizygous) call in all male samples were used. Sites called in fewer than 60 samples or with strand-bias p-value < 0.01 were filtered.

Estimation of site frequency spectra and summary statistics

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

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

Models of sex-chromosome diversity under neutral coalescent

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

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

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

$$\frac{\theta_{\pi,X}}{\theta_{\pi,A}} = \frac{h_X}{h_A}\frac{\mu_X}{\mu_A}\frac{\exp\left(\frac{(fg_1+g_2)(h_X-h_A)}{2Nh_X h_A f}\right)\left(1-f+\exp\left(\frac{g2}{2Nh_X f}\right)\left(f-1+\exp\left(\frac{g_1}{2Nh_X}\right)\right)\right)}{1-f+\exp\left(\frac{g_2}{Nh_A f}\right)\left(f-1+\exp\left(\frac{g1}{2Nh_X}\right)\right)}$$

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

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

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

In **Figure S2**, we plot X:A against Y:A. For the bottleneck and step-change models, X:A and Y:A vary with time since the onset of size change; these trajectories can be traced clockwise along each curve from $t = 0$ to $t = 4N$ (backwards in time.)

Demographic inference

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

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

Models of natural selection

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

We considered two modes of selection: background selection (BGS) due to purifying selection against deleterious mutations at linked sites; and hard selective sweeps on newly-arising beneficial mutations. Relative fitness in `SLiM` is modeled as $1 + s$ for sex-limited chromosomes. BGS was modeled by introducing mutations whose selection coefficients s were drawn from a mixture of a gamma distribution with mean $-\gamma = -N\bar{s}$ ($100 \times \alpha\%$ of mutations), and a point mass at zero (($100 \times (1 - \alpha)\%$ of mutations.) BGS simulations were run for $10N$ generations, sufficient to reach mutation-selection-drift equilibrium. For selective sweeps, the simulation was first run for $10N$ generations of burn-in, and then a beneficial variant was introduced with s drawn from a gamma distribution with mean $\gamma = N\bar{s}$. The simulation was then tracked until the beneficial variant was fixed 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 confirmed the integrity of simulations by checking that the pairwise diversity achieved by runs with selection coefficients fixed at zero matched the observed neutral values for each population (not shown.) Values of α were drawn from a uniform distribution on $(0, 1]$, and values of γ were drawn from a log-uniform distribution on $(10^{-6}, 10^{3}]$.

Simulations were connected to an approximate Bayesian computation (ABC) inference procedure implemented with the `R` package `abc` (Csilléry et al. 2012). Briefly, $500,000$ simulations were performed for each model. Five summary statistics were calculated from the SFS generated by each simulation: Watterson's estimator 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 for the observed joint SFS. The 0.1% of simulations with smallest Euclidean distance to the observed summary statistics were retained, accounting for collinearity between summary statistics using the "`neuralnet`" method of the function `abc::abc()`. Posterior distributions were computed via kernel smoothing over the parameter values of the retained simulations using an Epanechnikov kernel and plug-in bandwidth estimate.

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

Size estimation of co-amplified regions of Yq and X

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

To obtain an estimate for the gross size of Yq, all unmapped reads and reads mapping to mm10 Y were re-aligned to the Y chromosome contig of (Soh et al. 2014) using `bwa mem` with default parameters. Coverage was estimated over all reads,

regardless of mapping quality, in each of the "red", "yellow", "blue" and "grey" blocks in Figure 3 and Table S4 of (Soh et al. 2014). Read counts were normalized against a region of the X chromosome (chrX: $68.6 - 68.7$ Mb, containing the gene *Fmr1*) known to be present in a single haploid copy in all samples in the study. (This normalization implicitly accounts for mapping biases due to divergence between the target sample and the reference genome, provided the X and Y chromosomes diverge at roughly equal rates.) To estimate the total size of co-amplified regions of Yq we simply calculated the weighted sum of normalized coverage in the "red", "yellow" and "blue" blocks.

Identification of *de novo* CNVs in Collaborative Cross lines

Whole-genome sequencing reads (2×150 bp paired-end) from a single male individual from each 69 distinct Collaborative Cross (CC) lines were obtained from Srivastava et al. (2017). Alignment and quality control was performed as for wild mice. Read depth was estimated in 100 kb bins across the Y chromosome for each individual, and normalized for the effective depth of sequencing in that sample. Unambiguous alignment of 150 bp reads to the highly repetitive sequence on Yq is clearly not possible. However, each of the 8 founder Y chromosome haplogroups in the CC produces a characteristic read depth profile when reads are aligned with `bwa-mem`. We exploited this fact to remove noise from ambiguous read mapping by re-normalizing the estimated depth in each bin for each sample agains the median depth in that bin for CC lines sharing the same Y chromosome haplogroup (listed in **Table S2**). Any remaining deviations in read depth represent variation among lines sharing the same Y chromosome haplogroup, that is, candidate *de novo* CNVs. CNVs were ascertained by manual inspection of the re-normalized read depth profile of each CC line.

Analyses of gene expression

Multi-tissue, multi-species dataset. Neme and Tautz (Neme and Tautz 2016) measured gene expression in whole testis from wild-derived outbred mice from several species (**Figure 10A**) using RNA-seq. Reads were retrieved from the European Nucleotide Archive (PRJEB11513). Transcript-level expression was estimated using `kallisto` (Bray et al. 2016) using the Ensembl 85 transcript catatlog augmented with all *Slx/y*, *Sstx/y* and *Srsx/y* transcripts identified in (Soh et al. 2014). In the presence of redundant transcripts (*i.e.* from multiple copies of a co-amplified gene family), `kallisto` uses an expectation-maximization algorithm to distribute the "weight" of each read across transcripts without double-counting. Transcript-level expression estimates were aggregated to the gene level for differential expression testing using the `R` package `tximport`. As for the microarray data, "predicted" genes (with symbols "GmXXXX") on the Y chromosome were assigned to a co-amplified family where possible using Ensembl Biomart.

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

Spermatogenesis time course. Larson et al. (2016a) measured gene expression in isolated spermatids of three males from each of four F_1 crosses — CZECHII/EiJ×PWK/PhJ; LEWES/EiJ×PWK/PhJ; PWK/PhJ×LEWES/EiJ; and WSB/EiJ×LEWES/EiJ — using RNA-seq. Reads were retrieved from NCBI Short Read Archive (SRP065082). Transcript-level expression was

estimated using `kallisto` (Bray et al. 2016) using the Ensembl 85 transcript catatlog augmented with all *Slx/y*, *Sstx/y* and *Srsx/y* transcripts identified in (Soh et al. 2014). In the presence of redundant transcripts (*i.e.* from multiple copies of a co-amplified gene family), `kallisto` uses an expectation-maximization algorithm to distribute the "weight" of each read across transcripts without double-counting. Transcript-level expression estimates were aggregated to the gene level for differential expression testing using the `R` package `tximport`. As for the microarray data, "predicted" genes (with symbols "GmXXXX") on the Y chromosome were assigned to a co-amplified family where possible using Ensembl Biomart.

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

Definition of tissue-specific gene sets. The "tissue specificity index" (τ) of Yanai et al. (2005) was used to define tissue- or cell-type-specific gene sets. The index was first proposed for microarray data, and was adapted for 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 total), as estimated by `limma`. We require expression values to be strictly positive, so let $q = \min_i T_i$ and define $\tilde{T}_i = T_i + q$. Finally, calculate τ as

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

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

Acknowledgments

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

References

Akaike H. 1978. On the Likelihood of a Time Series Model. *Journal of the Royal Statistical Society. Series D (The Statistician)*. 27:217–235.

Arbiza L, Gottipati S, Siepel A, Keinan A. 2014. Contrasting X-Linked and Autosomal Diversity across 14 Human Populations. *The American Journal of Human Genetics*. 94:827–844.

Auton A, McVean G. 2007. Recombination rate estimation in the presence of hotspots. *Genome Res*. 17:1219–1227.

Baines JF, Harr B. 2007. Reduced X-Linked Diversity in Derived Populations of House Mice. *Genetics*. 175:1911–1921.

Beaumont MA, Zhang W, Balding DJ. 2002. Approximate Bayesian computation in population genetics. *Genetics*. 162:2025–2035.

Beck JA, Lloyd S, Hafezparast M, Lennon-Pierce M, Eppig JT, Festing MFW, Fisher EMC. 2000. Genealogies of mouse inbred strains. *Nat Genet*. 24:23–25.

Bellott DW, Hughes JF, Skaletsky H, et al. (30 co-authors). 2014. Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators. *Nature*. 508:494–499.

Bellott DW, Skaletsky H, Cho TJ, et al. (16 co-authors). 2017. Avian W and mammalian Y chromosomes convergently retained dosage-sensitive regulators. *Nat Genet*. 49:387–394.

Berta P, Hawkins JB, Sinclair AH, Taylor A, Griffiths BL, Goodfellow PN, Fellous M. 1990. Genetic evidence equating SRY and the testis-determining factor. *Nature*. 348:448–450.

Bhaskar A, Song YS. 2014. DESCARTES' RULE OF SIGNS AND THE IDENTIFIABILITY OF POPULATION DEMOGRAPHIC MODELS FROM GENOMIC VARIATION DATA. *Ann Stat*. 42:2469–2493.

Bishop CE, Boursot P, Baron B, Bonhomme F, Hatat D. 1985. Most classical Mus musculus domesticus laboratory mouse strains carry a Mus musculus musculus Y chromosome. *Nature*. 315:70–72.

Brawand D, Soumillon M, Necsulea A, et al. (18 co-authors). 2011. The evolution of gene expression levels in mammalian organs. *Nature*. 478:343–348.

Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. *Nat Biotech*. 34:525–527.

Burgoyne PS, Mahadevaiah SK, Sutcliffe MJ, Palmer SJ. 1992. Fertility in mice requires X-Y pairing and a Y-chromosomal "Spermiogenesis" gene mapping to the long arm. *Cell*. 71:391–398.

Campbell P, Good JM, Nachman MW. 2013. Meiotic sex chromosome inactivation is disrupted in sterile hybrid male house mice. *Genetics*. 193:819–828.

Campbell P, Nachman MW. 2014. X-y interactions underlie sperm head abnormality in hybrid male house mice. *Genetics*. 196:1231–1240.

Charlesworth B, Coyne JA, Barton NH. 1987. The Relative Rates of Evolution of Sex Chromosomes and Autosomes. *The American Naturalist*. 130:113–146.

Cocquet J, Ellis PJI, Mahadevaiah SK, Affara NA, Vaiman D, Burgoyne PS. 2012. A Genetic Basis for a Postmeiotic X Versus Y Chromosome Intragenomic Conflict in the Mouse. *PLOS Genet*. 8:e1002900.

Cocquet J, Ellis PJI, Yamauchi Y, Mahadevaiah SK, Affara NA, Ward MA, Burgoyne PS. 2009. The Multicopy Gene Sly Represses the Sex Chromosomes in the Male Mouse Germline after Meiosis. *PLOS Biol*. 7:e1000244.

Cocquet J, Ellis PJI, Yamauchi Y, Riel JM, Karacs TPS, Rattigan Á, Ojarikre OA, Affara NA, Ward MA, Burgoyne PS. 2010. Deficiency in the Multicopy Sycp3-Like X-Linked Genes Slx and Slxl1 Causes Major Defects in Spermatid Differentiation. *Mol. Biol. Cell*. 21:3497–3505.

Comptour A, Moretti C, Serrentino ME, Auer J, Ialy-Radio C, Ward MA, Touré A, Vaiman D, Cocquet J. 2014. SSTY proteins co-localize with the post-meiotic sex chromatin and interact with regulators of its expression. *FEBS J.* 281:1571–1584.

Consortium CC. 2012. The Genome Architecture of the Collaborative Cross Mouse Genetic Reference Population. *Genetics*. 190:389–401.

Conway SJ, Mahadevaiah SK, Darling SM, Capel B, Rattigan AM, Burgoyne PS. 1994. Y353/B: a candidate multiple-copy spermiogenesis gene on the mouse Y chromosome. *Mammalian Genome*. 5:203–210.

Cortez D, Marin R, Toledo-Flores D, Froidevaux L, Liechti A, Waters PD, Grützner F, Kaessmann H. 2014. Origins and functional evolution of Y chromosomes across mammals. *Nature*. 508:488–493.

Csilléry K, François O, Blum MGB. 2012. abc: an R package for approximate Bayesian computation (ABC). *Methods in Ecology and Evolution*. 3:475–479.

Derome N, Métayer K, Montchamp-Moreau C, Veuille M. 2004. Signature of Selective Sweep Associated With the Evolution of sex-ratio Drive in Drosophila simulans. *Genetics*. 166:1357–1366.

Din W, Anand R, Boursot P, Darviche D, Dod B, Jouvin-Marche E, Orth A, Talwar G, Cazenave PA, Bonhomme F. 1996. Origin and radiation of the house mouse: clues from nuclear genes. *Journal of Evolutionary Biology*. 9:519–539.

Doran AG, Wong K, Flint J, Adams DJ, Hunter KW, Keane TM. 2016. Deep genome sequencing and variation analysis of 13 inbred mouse strains defines candidate phenotypic alleles, private variation and homozygous truncating mutations. *Genome Biology*. 17:167.

Drost JB, Lee WR. 1995. Biological basis of germline mutation: Comparisons of spontaneous germline mutation rates among drosophila, mouse, and human. *Environ. Mol. Mutagen.* 25:48–64.

Egan CM, Sridhar S, Wigler M, Hall IM. 2007. Recurrent DNA copy number variation in the laboratory mouse. *Nat Genet*. 39:1384–1389.

Eicher EM, Hutchison KW, Phillips SJ, Tucker PK, Lee BK. 1989. A repeated segment on the mouse Y chromosome is composed of retroviral-related, Y-enriched and Y-specific sequences. *Genetics*. 122:181–192.

Ellegren H. 2011. Sex-chromosome evolution: recent progress and the influence of male and female heterogamety. *Nat Rev Genet*. 12:157–166.

Ellis PJI, Bacon J, Affara NA. 2011. Association of Sly with sex-linked gene amplification during mouse evolution: a side effect of genomic conflict in spermatids? *Hum. Mol. Genet.* 20:3010–3021.

Ellis PJI, Clemente EJ, Ball P, Touré A, Ferguson L, Turner JMA, Loveland KL, Affara NA, Burgoyne PS. 2005. Deletions on mouse Yq lead to upregulation of multiple X- and Y-linked transcripts in spermatids. *Hum Mol Genet*. 14:2705–2715.

Fischer M, Kosyakova N, Liehr T, Dobrowolski P. 2016. Large deletion on the Y-chromosome long arm (Yq) of C57bl/6jbomtac inbred mice. *Mamm Genome*. pp. 1–7.

Forejt J. 1996. Hybrid sterility in the mouse. *Trends in Genetics*. 12:412–417.

Forejt J, Iványi P. 1974. Genetic studies on male sterility of hybrids between laboratory and wild mice (Mus musculus L.). *Genet. Res.* 24:189–206.

Geraldes A, Basset P, Gibson B, Smith KL, Harr B, Yu HT, Bulatova N, Ziv Y, Nachman MW. 2008. Inferring the history of speciation in house mice from autosomal, X-linked, Y-linked and mitochondrial genes. *Mol. Ecol.* 17:5349–5363.

Good JM, Dean MD, Nachman MW. 2008. A Complex Genetic Basis to X-Linked Hybrid Male Sterility Between Two Species of House Mice. *Genetics*. 179:2213–2228.

Graves JAM. 2006. Sex Chromosome Specialization and Degeneration in Mammals. *Cell*. 124:901–914.

Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD. 2009. Inferring the Joint Demographic History of Multiple Populations from Multidimensional SNP Frequency Data. *PLOS Genetics*. 5:e1000695.

Haller BC, Messer PW. 2017. SLiM 2: Flexible, Interactive Forward Genetic Simulations. *Mol Biol Evol*. 34:230–240.

Halligan DL, Kousathanas A, Ness RW, Harr B, Eöry L, Keane TM, Adams DJ, Keightley PD. 2013. Contributions of Protein-Coding and Regulatory Change to Adaptive Molecular Evolution in Murid Rodents. *PLOS Genet*. 9:e1003995.

Harr B, Karakoc E, Neme R, et al. (19 co-authors). 2016. Genomic resources for wild populations of the house mouse, Mus musculus and its close relative Mus spretus. *Scientific Data*. 3:160075.

Hendriksen PJ, Hoogerbrugge JW, Themmen AP, Koken MH, Hoeijmakers JH, Oostra BA, van der Lende T, Grootegoed JA. 1995. Postmeiotic transcription of X and Y chromosomal genes during spermatogenesis in the mouse. *Dev. Biol.* 170:730–733.

Hudson RR, Kaplan NL. 1995. The Coalescent Process and Background Selection. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 349:19–23.

Hughes JF, Page DC. 2015. The Biology and Evolution of Mammalian Y Chromosomes. *Annual Review of Genetics*. 49:507–527.

Hvilsom C, Qian Y, Bataillon T, et al. (17 co-authors). 2012. Extensive X-linked adaptive evolution in central chimpanzees. *PNAS*. 109:2054–2059.

Itsara A, Wu H, Smith JD, Nickerson DA, Romieu I, London SJ, Eichler EE. 2010. De novo rates and selection of large copy number variation. *Genome Res*. 20:1469–1481.

Jaenike J. 2001. Sex Chromosome Meiotic Drive. *Annual Review of Ecology and Systematics*. 32:25–49.

Keane TM, Goodstadt L, Danecek P, et al. (41 co-authors). 2011. Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature*. 477:289–294.

Kingan SB, Garrigan D, Hartl DL. 2010. Recurrent Selection on the Winters sex-ratio Genes in Drosophila simulans. *Genetics*. 184:253–265.

Korneliussen TS, Albrechtsen A, Nielsen R. 2014. ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics*. 15:356.

Kousathanas A, Halligan DL, Keightley PD. 2014. Faster-X Adaptive Protein Evolution in House Mice. *Genetics*. 196:1131–1143.

Lahn BT, Page DC. 1997. Functional Coherence of the Human Y Chromosome. *Science*. 278:675–680.

Larson EL, Keeble S, Vanderpool D, Dean MD, Good JM. 2016a. The composite regulatory basis of the large X-effect in mouse speciation. *Mol Biol Evol*. p. msw243.

Larson EL, Vanderpool D, Keeble S, Zhou M, Sarver BAJ, Smith AD, Dean MD, Good JM. 2016b. Contrasting Levels of Molecular Evolution on the Mouse X Chromosome. *Genetics*. 203:1841–1857.

Li G, Davis BW, Raudsepp T, Wilkerson AJP, Mason VC, Ferguson-Smith M, O'Brien PC, Waters PD, Murphy WJ. 2013. Comparative analysis of mammalian Y chromosomes illuminates ancestral structure and lineage-specific evolution. *Genome Res.* 23:1486–1495.

Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv:1303.3997 [q-bio]*. ArXiv: 1303.3997.

Liu KJ, Steinberg E, Yozzo A, Song Y, Kohn MH, Nakhleh L. 2015. Interspecific introgressive origin of genomic diversity in the house mouse. *PNAS*. 112:196–201.

Macholán M, Baird SJ, Munclinger P, Dufková P, Bímová B, Piálek J. 2008. Genetic conflict outweighs heterogametic incompatibility in the mouse hybrid zone? *BMC Evolutionary Biology*. 8:271.

McDonald JH, Kreitman M. 1991. Adaptive protein evolution at the Adh locus in Drosophila. *Nature*. 351:652–654.

McKenna A, Hanna M, Banks E, et al. (11 co-authors). 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20:1297–1303.

McLaren A, Simpson E, Epplen JT, Studer R, Koopman P, Evans EP, Burgoyne PS. 1988. Location of the genes controlling H-Y antigen expression and testis determination on the mouse Y chromosome. *Proc. Natl. Acad. Sci. U.S.A.* 85:6442–6445.

Moretti C, Vaiman D, Tores F, Cocquet J. 2016. Expression and epigenomic landscape of the sex chromosomes in mouse post-meiotic male germ cells. *Epigenetics & Chromatin*. 9:47.

Morgan AP, Didion JP, Doran AG, Holt JM, McMillan L, Keane TM, Villena FPMd. 2016a. Genome Report: Whole Genome Sequence of Two Wild-Derived Mus musculus domesticus Inbred Strains, LEWES/EiJ and ZALENDE/EiJ, with Different Diploid Numbers. *G3*. p. g3.116.034751.

Morgan AP, Gatti DM, Najarian ML, Keane TM, Galante RJ, Pack AI, Mott R, Churchill GA, Villena FPMd. 2017. Structural Variation Shapes the Landscape of Recombination in Mouse. *Genetics*. 206:603–619.

Morgan AP, Holt JM, McMullan RC, et al. (12 co-authors). 2016b. The Evolutionary Fates of a Large Segmental Duplication in Mouse. *Genetics*. 204:267–285.

Mueller JL, Mahadevaiah SK, Park PJ, Warburton PE, Page DC, Turner JMA. 2008. The mouse X chromosome is enriched for multicopy testis genes showing postmeiotic expression. *Nat. Genet.* 40:794–799.

Mueller JL, Skaletsky H, Brown LG, Zaghlul S, Rock S, Graves T, Auger K, Warren WC, Wilson RK, Page DC. 2013. Independent specialization of the human and mouse X chromosomes for the male germ line. *Nat. Genet.* 45:1083–1087.

Nam K, Munch K, Hobolth A, et al. (85 co-authors). 2015. Extreme selective sweeps independently targeted the X chromosomes of the great apes. *PNAS*. 112:6413–6418.

Neme R, Tautz D. 2016. Fast turnover of genome transcription across evolutionary time exposes entire non-coding DNA to de novo gene emergence. *eLife*. 5:e09977.

Nishioka Y, Lamothe E. 1986. Isolation and characterization of a mouse Y chromosomal repetitive sequence. *Genetics*. 113:417–432.

Oh B, Hwang SY, Solter D, Knowles BB. 1997. Spindlin, a major maternal transcript expressed in the mouse during the transition from oocyte to embryo. *Development*. 124:493–503.

Orth A, Belkhir K, Britton-Davidian J, Boursot P, Benazzou T, Bonhomme F. 2002. Natural hybridization between 2 sympatric species of mice, Mus musculus domesticus L. and Mus spretus Lataste. *C. R. Biol.* 325:89–97.

Payseur BA, Krenz JG, Nachman MW. 2004. Differential Patterns of Introgression across the X Chromosome in a Hybrid Zone between Two Species of House Mice. *Evolution*. 58:2064–2078.

Polanski A, Szczesna A, Garbulowski M, Kimmel M. 2017. Coalescence computations for large samples drawn from populations of time-varying sizes. *PLOS ONE*. 12:e0170701.

Pool JE, Nielsen R. 2007. Population Size Changes Reshape Genomic Patterns of Diversity. *Evolution*. 61:3001–3006.

Pritchard JK, Seielstad MT, Perez-Lezaun A, Feldman MW. 1999. Population growth of human Y chromosomes: a study of Y chromosome microsatellites. *Mol Biol Evol*. 16:1791–1798.

Repping S, van Daalen SKM, Brown LG, et al. (11 co-authors). 2006. High mutation rates have driven extensive structural polymorphism among human Y chromosomes. *Nat Genet*. 38:463–467.

Salcedo T, Geraldes A, Nachman MW. 2007. Nucleotide variation in wild and inbred mice. *Genetics*. 177:2277–2291.

Sarver BAJ, Keeble S, Cosart T, Tucker PK, Dean MD, Good JM. 2017. Phylogenomic Insights into Mouse Evolution Using a Pseudoreference Approach. *Genome Biol Evol*. 9:726–739.

Sayres MAW, Lohmueller KE, Nielsen R. 2014. Natural Selection Reduced Diversity on Human Y Chromosomes. *PLOS Genetics*. 10:e1004064.

Scally A. 2016. Mutation rates and the evolution of germline structure. *Phil. Trans. R. Soc. B*. 371:20150137.

Simmons LW, Fitzpatrick JL. 2012. Sperm wars and the evolution of male fertility. *Reproduction*. 144:519–534.

Smith NGC, Eyre-Walker A. 2002. Adaptive protein evolution in Drosophila. *Nature*. 415:1022–1024.

Soh YQS, Alföldi J, Pyntikova T, et al. (21 co-authors). 2014. Sequencing the mouse Y chromosome reveals convergent gene acquisition and amplification on both sex chromosomes. *Cell*. 159:800–813.

Song Y, Endepols S, Klemann N, Richter D, Matuschka FR, Shih CH, Nachman MW, Kohn MH. 2011. Adaptive Introgression of Anticoagulant Rodent Poison Resistance by Hybridization between Old World Mice. *Current Biology*. 21:1296–1301.

Spiess AN, Walther N, Müller N, Balvers M, Hansis C, Ivell R. 2003. SPEER—A New Family of Testis-Specific Genes from the Mouse. *Biol Reprod*. 68:2044–2054.

Srivastava A, Morgan AP, Najarian ML, et al. (18 co-authors). 2017. Genomes of the Mouse Collaborative Cross. *Genetics*. 206:537–556.

Storchová R, Gregorová S, Buckiová D, Kyselová V, Divina P, Forejt J. 2004. Genetic analysis of X-linked hybrid sterility in the house mouse. *Mamm Genome*. 15:515–524.

Styrna J, Klag J, Moriwaki K. 1991. Influence of partial deletion of the Y chromosome on mouse sperm phenotype. *J Reprod Fertil*. 92:187–195.

Teeter KC, Payseur BA, Harris LW, Bakewell MA, Thibodeau LM, O'Brien JE, Krenz JG, Sans-Fuentes MA, Nachman MW, Tucker PK. 2008. Genome-wide patterns of gene flow across a house mouse hybrid zone. *Genome Res*. 18:67–76.

Tishkoff SA, Reed FA, Ranciaro A, et al. (19 co-authors). 2007. Convergent adaptation of human lactase persistence in Africa and Europe. *Nat Genet*. 39:31–40.

Torgerson DG, Singh RS. 2003. Sex-linked mammalian sperm proteins evolve faster than autosomal ones. *Mol. Biol. Evol.* 20:1705–1709.

Touré A, Clemente EJ, Ellis P, Mahadevaiah SK, Ojarikre OA, Ball PAF, Reynard L, Loveland KL, Burgoyne PS, Affara NA. 2005. Identification of novel Y chromosome encoded transcripts by testis transcriptome analysis of mice with deletions of the Y chromosome long arm. *Genome Biol.* 6:R102.

Touré A, Szot M, Mahadevaiah SK, Rattigan Á, Ojarikre OA, Burgoyne PS. 2004. A New Deletion of the Mouse Y Chromosome Long Arm Associated With the Loss of Ssty Expression, Abnormal Sperm Development and Sterility. *Genetics*. 166:901–912.

Tu S, Shin Y, Zago WM, States BA, Eroshkin A, Lipton SA, Tong GG, Nakanishi N. 2007. Takusan: A Large Gene Family that Regulates Synaptic Activity. *Neuron*. 55:69–85.

Tucker PK, Lee BK, Lundrigan BL, Eicher EM. 1992. Geographic origin of the Y chromosomes in "old" inbred strains of mice. *Mamm. Genome*. 3:254–261.

Turner LM, Schwahn DJ, Harr B. 2012. Reduced Male Fertility Is Common but Highly Variable in Form and Severity in a Natural House Mouse Hybrid Zone. *Evolution*. 66:443–458.

Turner LM, White MA, Tautz D, Payseur BA. 2014. Genomic Networks of Hybrid Sterility. *PLOS Genet*. 10:e1004162.

Uchimura A, Higuchi M, Minakuchi Y, Ohno M, Toyoda A, Fujiyama A, Miura I, Wakana S, Nishino J, Yagi T. 2015. Germline mutation rates and the long-term phenotypic effects of mutation accumulation in wild-type laboratory mice and mutator mice. *Genome Res.* 25:1125–1134.

Veeramah KR, Gutenkunst RN, Woerner AE, Watkins JC, Hammer MF. 2014. Evidence for Increased Levels of Positive and Negative Selection on the X Chromosome versus Autosomes in Humans. *Mol Biol Evol*. 31:2267–2282.

Webster TH, Wilson Sayres MA. 2016. Genomic signatures of sex-biased demography: progress and prospects. *Current Opinion in Genetics & Development*. 41:62–71.

Yamauchi Y, Riel JM, Stoytcheva Z, Burgoyne PS, Ward MA. 2010. Deficiency in mouse Y chromosome long arm gene complement is associated with sperm DNA damage. *Genome Biology*. 11:R66.

Yamauchi Y, Riel JM, Wong SJ, Ojarikre OA, Burgoyne PS, Ward MA. 2009. Live Offspring from Mice Lacking the Y Chromosome Long Arm Gene Complement. *Biol Reprod*. 81:353–361.

Yanai I, Benjamin H, Shmoish M, et al. (12 co-authors). 2005. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics*. 21:650–659.

Yang H, Wang JR, Didion JP, et al. (15 co-authors). 2011. Subspecific origin and haplotype diversity in the laboratory mouse. *Nat Genet*. 43:648–655.

Supplementary material

Table S1: List of samples used in this study.

Table S2: Y chromosome haplogroups and CNV status for 69 Collaborative Cross strains.

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.

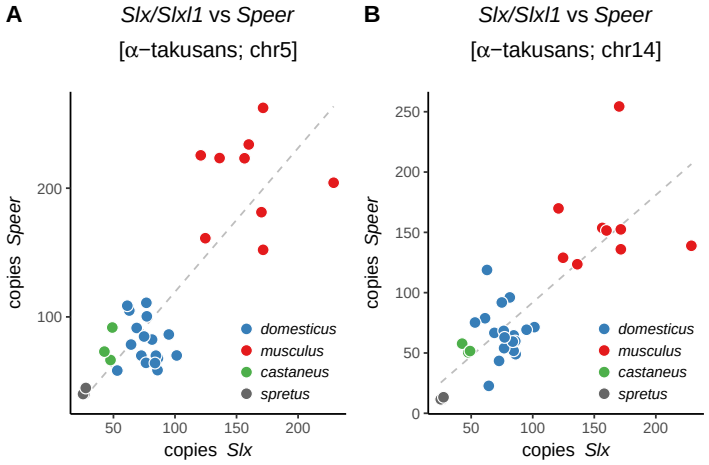

Figure S1: Copy number of *Speer* family members on chromosomes 5 (**A**) and 14 (**B**) compared to copy number of *Slx/Slxl1*.

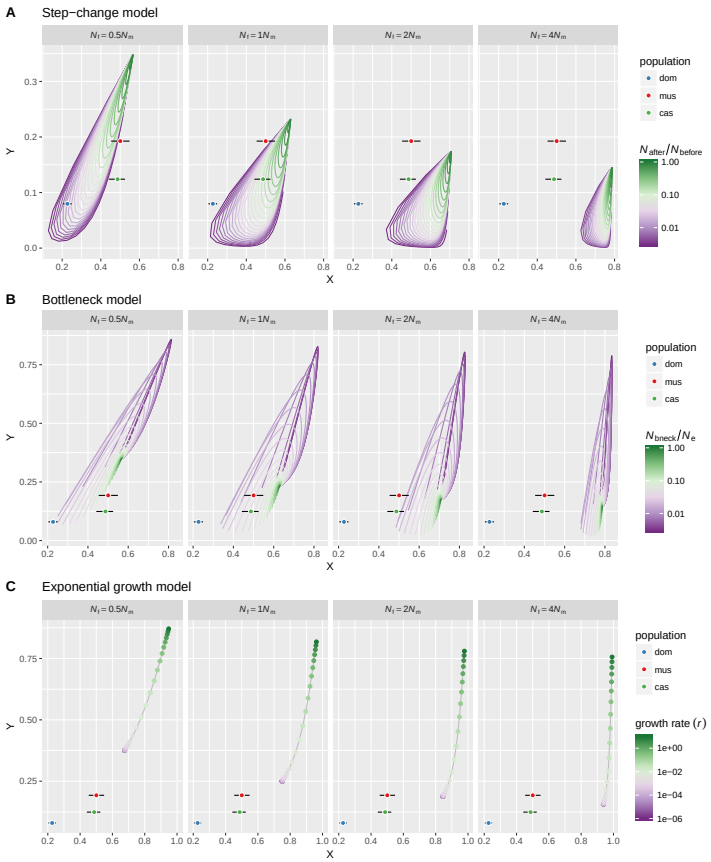

Figure S2: X:A and Y:A diversity ratios under several neutral coalscent models with varying sex ratio. Error bars represent boostrap standard errors.

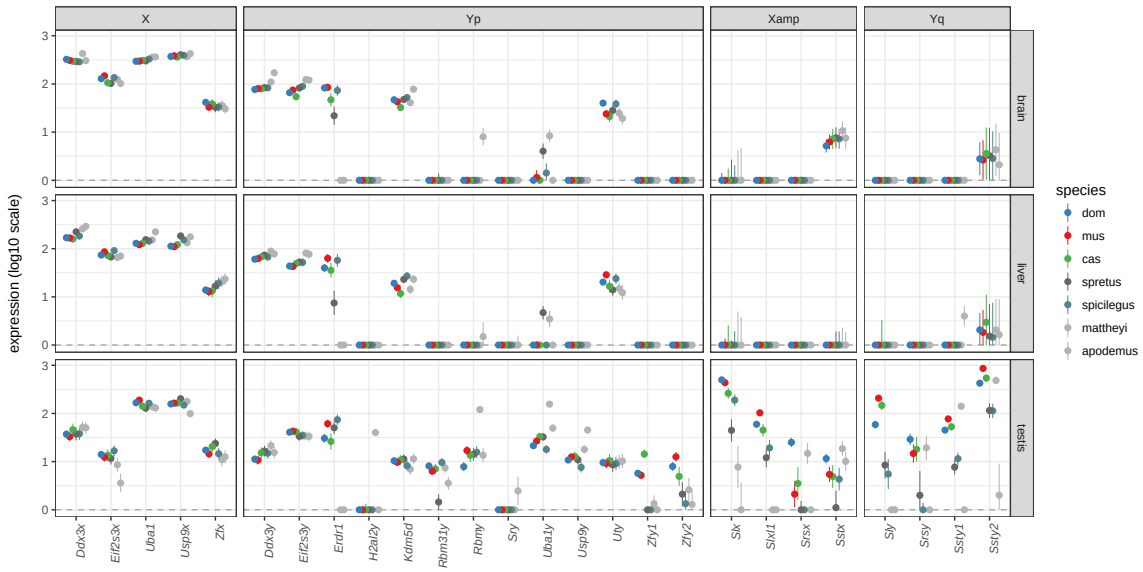

Figure S3: Expression of selected X- and Y-linked genes across *Mus* species. Genes are subdivided by chromosomal location: X, non-ampliconic X-linked genes with Y-linked homologs; Yp, non-ampliconic genes on Yp; Xamp, X-linked homologs of co-amplified genes; Yq, Y-linked homologs of co-amplified genes, residing on Yq.