Sequence and structural diversity of mouse Y chromosomes

Andrew P Morgan and Fernando Pardo-Manuel de Villena*

Department of Genetics and Linebarger Comprehensive Cancer Center

University of North Carolina, Chapel Hill, NC

Abstract

Over the 180 million years since their origin, the Y chromosomes of mammals have evolved a gene repertoire highly specialized for function in the male germline. The mouse Y chromosome is unique among mammal Y chromosomes studied to date in that it is large, gene-rich and euchromatic. Yet little is known about its diversity in natural populations. Because the Y chromosome is passed only through the male germline and is obligately transmitted from fathers to sons without recombination, it provides a rich view into male-specific mutational, selective and demographic processes. We therefore took advantage of a recent high-quality assembly of the mouse Y to perform a systematic survey of a diverse sample of Y chromosomes using published whole-genome sequencing datasets. Sequence diversity in non-repetitive regions of Y chromosomes is $< 10\%$ that on autosomes and the site frequency spectrum is skewed towards low-frequency alleles, consistent with a recent population bottleneck. But copy number of genes on the repetitive long arm of the Y is extremely variable: the total size of the Y chromosome varies by two-fold within *Mus musculus* and three-fold between *M. musculus* and *M. spretus*. We show that expression of Y-linked genes in the testis is rapidly evolving in murid rodents and especially within *M. musculus*, and is consistent with ongoing intragenomic conflict with the X chromosome. Our results provide insight on the demographic history of an important model organism and the biology of a rapidly-evolving sex chromosome.

Keywords: Y chromosome, sex chromosome evolution, intragenomic conflict

*Corresponding author

5049C Genetic Medicine Building

Department of Genetics

University of North Carolina

Chapel Hill NC 27599-7264

fernando@med.unc.edu

1 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 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, mammal 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 mammal 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 was also known to be large and almost entirely euchromatic (**?**). Spontaneous 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 (Burgoyne et al., 1992; Touré et al., 2004).

Sequencing, assembly and annotation of the mouse Y in the reference 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 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 distant 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 presumably function in sperm maturation.

Independent amplification of homologous genes on the X and Y chromosomes is thought to be a byproduct of competition

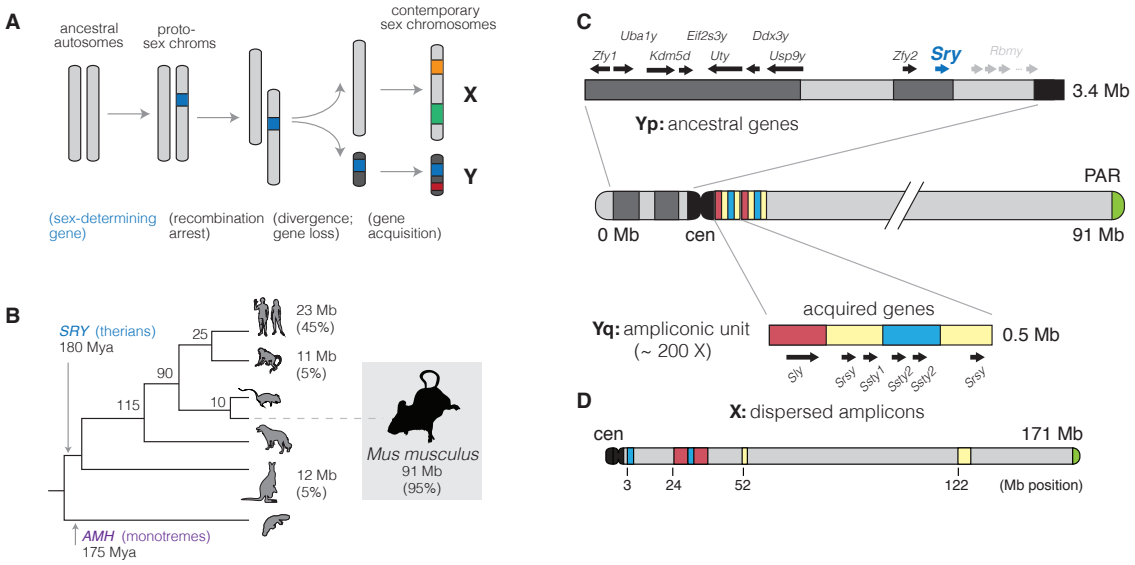

Figure 1: Evolution of mammal Y chromosomes. (**A**) Evolution of heteromorphic sex chromosomes. (**B**) Y chromosomes of mammals. The Y chromosome of therian mammals, characterized by the sex-determining factor *SRY*, diverged from the mammal X approximately 180 Mya. (The monotremata have a different 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 (Yq) consists primarily of genes shared the X and retained since the X and Y 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, is acrocentric. Homologs of the acquired genes from the Y (*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.

between the X and Y 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 (Cocquet et al., 2009). 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.

Because the Y chromosome is passed only through the male germline and is obligately transmitted from fathers to sons without recombination, it provides a rich view into male-specific mutational, selective and demographic processes. We therefore took advantage of the recent high-quality assembly of the mouse Y (Soh et al., 2014) to perform a systematic survey of a diverse sample of Y chromosomes using public whole-genome sequencing datasets. In this manuscript we characterize both sequence and structural variation in *Mus*, and use complementary gene expression data from natural populations and laboratory crosses to explore proximate functional consequences of this variation. We find that sequence diversity is dramatically reduced on both sex chromosomes relative to neutral expectations, most likely due to a recent population bottlenecks along some lineages. However, copy number of amplified genes on X and Yq is highly variable between populations. Expression patterns of Y-linked genes are rapidly evolving within the murid rodents and especially within *M. musculus*. In hybrids between subspecies, disrupted sex-linked gene expression in speramtids is consistent with sex-chromosome conflict — but this conflict cannot be understood in simple terms of X-versus-Y-linked copy number.

2 Results

2.1 A catalog of Y-linked sequence variation in mouse

Whole-genome sequence data for 68 male mice was collected from public sources (Keane et al., 2011; Doran et al., 2016; Harr et al., 2016; Morgan et al., 2016; Neme and Tautz, 2016). The final set consisted of 42 wild-caught mice; 20 classical inbred strains; 1 laboratory mouse derived from an outbred stock; and 5 wild-derived inbred strains (**Table 1**). All three cardinal subspecies of *M. musculus* (*domesticus*, *musculus* and *castaneus*) are represented. *Mus spretus* and *Mus spicilegus* served as close outgroups for analyses of the Y chromosome, and a female *Mus caroli* individual was used as a more distant outgroup in analyses of the mitochondrial genome.

SNVs and small indels were ascertained in 2.2 Mb of non-repetitive sequence on Yp and assigned ancestral or derived status based on the consensus call among the *M. spretus* samples. We identified $27,715$ high-confidence SNVs (transitions:transversions $= 1.72$) and $3,009$ high-confidence indels segregating in *M. musculus* after applying stringent filters for genotype quality (see **Materials and methods**). Of these 286 (0.9%) fall in protein-coding genes, and only 161 are predicted

Type	Population	Locality	N	
wild	*M. m. domesticus*	DE	8	
		FR	8	
		IR	8	
	M. m. musculus	CZ	2	
		KZ	3	
		AF	5	
	M. m. castaneus	IN	3	
	M. spretus	ES	4	
	M. spicilegus	HU	1	
wild-derived	*M. m. domesticus*	CH	1	ZALENDE/EiJ
		US	1	LEWES/EiJ
	M. m. musculus	CZ	1	PWK/PhJ
	M. m. castaneus	TH	1	CAST/EiJ
	M. spretus	ES	1	SPRET/EiJ
lab	-	-	21	

Table 1: Wild and laboratory mice used for Y chromosome analyses. Localities are given as two-letter country codes.

to impact protein function.

One group of inbred strains in our dataset — C57BL/6J, C57BL/10J, C57L/J and C57BR/cdJ — have a known common ancestor in the year 1929. We used this fact to obtain a direct estimate of the male-specific point mutation rate: $5.4 \times 10^{-9} - 8.1 \times 10^{-9}$ bp^{-1} generation^{-1}, assuming an average of three generations per year. This interval just overlaps 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)), we obtain a male-specific autosomal rate of 7.9×10^{-9} bp^{-1} generation^{-1}, in good agreement with our estimate from the Y chromosome.

2.2 Phylogeny of Y chromosomes recovers geographic relationships

A phylogenetic tree for the Y chromosome and mitochondrial genome were constructed with `BEAST` (**Figure 2**). The approximate time to most recent common ancestor (MRCA) of *M. musculus* Y chromosomes is $275,000$ (95% highest posterior density interval [HPDI] $267,000 - 282,000$) years ago. Within *M. musculus*, the *musculus* subspecies diverges first, although the internal branch separating it from the MRCA of *domesticus* 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 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). A more detailed view of the Y chromosome phylogeny is shown in **Figure S1**.

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

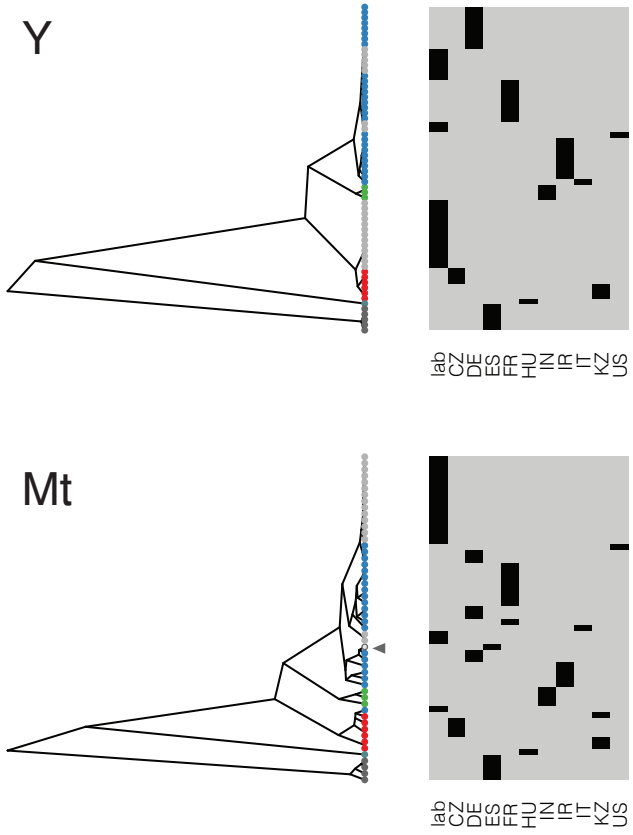

Figure 2: Phylogenetic trees for Y chromosomes (top) and complete mitochondrial genomes (bottom) of wild and laboratory mice. Samples are colored according to their taxonomic origin: blue, *M. m. domesticus*; red, *M. m. musculus*; green, *M. m. castaneus*; dark grey, *M. spretus*; and light grey, classical laboratory strains. Heatmaps at right are incidence matrices of samples onto countries (shown as two-letter country codes). More block structure indicates greater clustering of samples by geography. Arrowhead points to *M. spretus* sample with *M. m. domesticus* mitochondrial introgression.

Locus	Pop	L	N	S	θ_w	θ_π	D	D_{FL}	F	Y
A	*dom*	235019	52	3961	0.373 (0.006)	0.339 (0.006)	−0.33	2.35	1.47	−0.42
	mus		20	3875	0.465 (0.008)	0.409 (0.008)	−0.51	1.72	1.16	−0.40
	cas		6	4496	0.838 (0.013)	0.786 (0.013)	−0.40	1.87	1.57	−0.15
X	*dom*	77654	26	349	0.118 (0.008)	0.056 (0.005)	−2.09	−5.11	−4.92	−0.89
	mus		10	263	0.120 (0.009)	0.073 (0.005)	−1.97	−3.51	−3.77	−0.53
	cas		4	485	0.341 (0.017)	0.315 (0.015)	−0.80	−1.80	−1.91	−0.073
Y	*dom*	995467	26	2199	0.058 (0.001)	0.029 (0.001)	−1.97	−4.86	−4.66	−0.84
	mus		10	1613	0.057 (0.001)	0.037 (0.001)	−1.78	−3.12	−3.35	−0.49
	cas		4	3493	0.191 (0.003)	0.177 (0.003)	−0.79	−1.24	−1.37	−0.07
M	*dom*	979	26	18	0.482 (0.177)	0.142 (0.052)	−2.45	−4.82	−4.87	−1.21
	mus		10	9	0.335 (0.169)	0.190 (0.096)	−1.83	−1.51	−1.91	−0.65
	cas		4	3	0.141 (0.111)	0.130 (0.101)	−0.63	−1.14	−1.21	−0.09

Table 2: Sequence diversity statistics for autosomes, X and Y chromosomes and mitochondrial genome, by population. L, total sizes; S, segregating sites; θ_w, Watterson's θ; θ_π, Tajima's pairwise θ; D, Tajima's D; D_{FL}, Fu and Li's D; F, Fu and Li's F; Y, Achaz's Y. Both estimators of θ are expressed as percentages with bootstrap standard errors in parentheses.

		Population		
Comparison	Expected	*dom*	*mus*	*cas*
X:A	$^3/_4$	0.163 (0.012)	0.176 (0.013)	0.399 (0.026)
Y:A	$^1/_4$	0.0868 (0.0028)	0.0909 (0.0030)	0.225 (0.005)
Y:X	$^1/_3$	0.531 (0.047)	0.5120 (0.0392)	0.5640 (0.0294)

Table 3: Diversity ratios between pairs of chromosome types relative to neutral expectations. Bootstrap standard errors are shown in parentheses.

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 but a *M. m. domesticus* mitochondrial genome (arrowhead in **Figure 2**). 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).

2.3 Sequence diversity is markedly reduced on both sex chromosomes

We estimated nucleotide diversity within subspecies directly from genotype likelihoods on Yp (Korneliussen et al., 2014), rather than from called genotypes at variable sites. The rank ordering of subspecies by Y chromosome diversity parallels what has previously been shown for autosomes: *castaneus* >> *domesticus* > *musculus* (**Table 2**). Our estimates of diversity at Y-linked sites ($\pi_{\text{dom}} = 0.029\% \pm 0.001\%$, $\pi_{\text{mus}} = 0.037\% \pm 0.001\%$, $\pi_{\text{cas}} = 0.177\% \pm 0.003\%$) are in line with previous reports (Salcedo et al., 2007; Geraldes et al., 2008). To provide context for observed levels of Y-linked variation, we compared relative diversity in pairwise combinations of the autosomes, X and Y chromosomes within subspecies to neutral expectations (**Table 3**). We found a deficit of variation on both sex chromosomes relative to the autosomes. The effect is stronger on the X (approximately 80% lower nucleotide diversity than expected) than the Y chromosome (40%), and is stronger in *domesticus* and *musculus* than in *castaneus*.

		Population		
Locus		*dom*	*mus*	*cas*
A	*dom*	-	0.275 (0.0031)	0.388 (0.0033)
	mus	0.822 (0.013)	-	0.284 (0.0028)
	cas	0.91 (0.015)	0.847 (0.011)	-
X	*dom*	-	0.622 (0.012)	0.681 (0.0094)
	mus	0.165 (0.011)	-	0.567 (0.0098)
	cas	0.231 (0.011)	0.238 (0.01)	-
Y	*dom*	-	0.68 (0.0034)	0.71 (0.0026)
	mus	0.165 (0.0031)	-	0.616 (0.0033)
	cas	0.191 (0.0029)	0.171 (0.0029)	-
M	*dom*	-	0.65 (0.12)	0.87 (0.054)
	mus	0.52 (0.19)	-	0.529 (0.14)
	cas	0.566 (0.21)	0.289 (0.11)	-

Table 4: Population differentiation (F_{st}, above diagonal) and divergence per site (d_{xy} as percentage, below diagonal) for autosomes and sex chromosomes. Bootstrap standard errors in parentheses.

	Population		
Loci	*dom*	*mus*	*cas*
Y vs X	0.79 (0.374)	0.45 (0.502)	0.03 (0.853)
Y vs M	*15.35 (8.9×10^{-5})	*4.42 (0.0356)	0.10 (0.755)
X vs M	*8.36 (3.8×10^{-3})	2.10 (0.148)	0.20 (0.653)

Table 5: Hudson-Kreitman-Aguadé (HKA) tests for neutral evolution of Y chromosomes compared to X-linked or mitochondrial loci. Entries in the table are the χ^2 statistic from the HKA test with p-values in parentheses. Comparisons for which the null hypothesis is rejected are marked with asterisks (*).

Levels of population differentiation, measured by F_{st}, are also increased on the sex chromosomes relative to autosomal loci. Here the effect is strongest for the Y chromosome (**Table 4**), with F_{st} values ranging from 0.62 (*musculus-castaneus*) to 0.71 (*domesticus-castaneus*).

To investigate possible causes of reduction in diversity on the Y chromosome, we used two complementary families of tests: the Hudson-Kreitman-Aguade (HKA) test (Hudson et al., 1987) and variations on Tajima's D statistic (Tajima, 1989). We used to HKA test to compare the Y chromosomes to the mitochondria and to X chromosome separately in *domesticus*, *musculus* and *castaneus*. The null hypothesis is rejected for *domesticus* ($p = 8.9 \times 10^{-5}$) and *musculus* ($p = 0.04$) but not for *castaneus* ($p = 0.76$) in the Y-mitochondria comparison. In both *musculus* and *domesticus*, the Y chromosome shows a deficit of polymorphism relative to the mitochondria (**Table 5**). No excess of divergence relative to polymorphism was detected in the Y-X comparison. Tajima's D, Fu and Li's D and F (Fu and Li, 1993) and Achaz's Y (Achaz, 2008) all take significantly negative values on the Y chromosome in *domesticus* and *musculus* but not *castaneus* (**Table 2**).

2.4 Reduction in Y-linked diversity is consistent with a bottleneck

Because it is inherited only through the male line and does not undergo recombination, the Y chromosome is a sensitive marker for the male-specific demographic history of populations. We used an approximate Bayesian computation (ABC) (Tavaré et al., 1997; Pritchard et al., 1999) strategy to evaluate models for patrilineal demography against our Y chromosome

dataset. Neutral coalescent simulations were carried out under several demographic scenarios (**Figure 3A**). Simulated and observed polymorphism data at putatively neutral sites were compared using summaries over the joint SFS across *domesticus*, *musculus* and *castaneus* (see **Materials and methods**). In the ABC scheme, a subset of simulations yielding summary statistics within a threshold distance of the observed values are treated as a sample from the marginal posterior distribution over demographic model parameters (Beaumont et al., 2002).

We evaluated eight families of demographic models of increasing complexity and used Bayes factors for model selection (**Figure 3**). Models with gene flow (I – IV) generally provided better fit to the data than models without gene flow (V – VIII). The best-fitting model (model VII) includes a bottleneck shared by *domesticus* and *musculus* but not *castaneus*. It captures the key features of the observed data: reduced diversity in *domesticus* and *musculus*; excess of low-frequency alleles in *domesticus* and *musculus*; and approximately equal F_{st} between all population pairs. Under this model, N_e for castaneus is approximately 1.5-fold higher than in *domesticus* or *musculus* and the three Y chromosome lineages began to diverge $636,000$ generations in the past (**Figure S2** and **Table 6**). The inferred bottleneck is sharp, reducing N_e by 89% (50% HPDI $87-98\%$).

2.5 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 4**). The distribution of CNV alleles follows the SNV-based phylogenetic tree. Only one region, at the centromeric end of Yq, contains a known protein-coding gene (the *Rbmy* family). 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. We identified one additional CNV overlapping a protein-coding gene: the wild-derived inbred strain LEWES/EiJ (from Delaware; *M. m. domesticus* ancestry) carries an 82 kb duplication containing *Eif2s3y*.

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 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. 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 5** shows that copy number on X and Y 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 $0.87-1.09$]; $R^2 = 0.87$). This supports previous evidence that conflict between X and Y, if it exists, is mediated primarily through expression of *Slx* and *Sly* (Cocquet et al., 2012).

As an alternative approach, we also aggregated the reads aligning to the fundamental ampliconic units identified by (Soh et al., 2014) and counted the total number of reads mapping to the interdigitated "red", "blue" and "yellow" sequence families (shown in **Figure 1**). Consistent with the hypothesis that Yq expands and contracts by gain or loss of copies of the

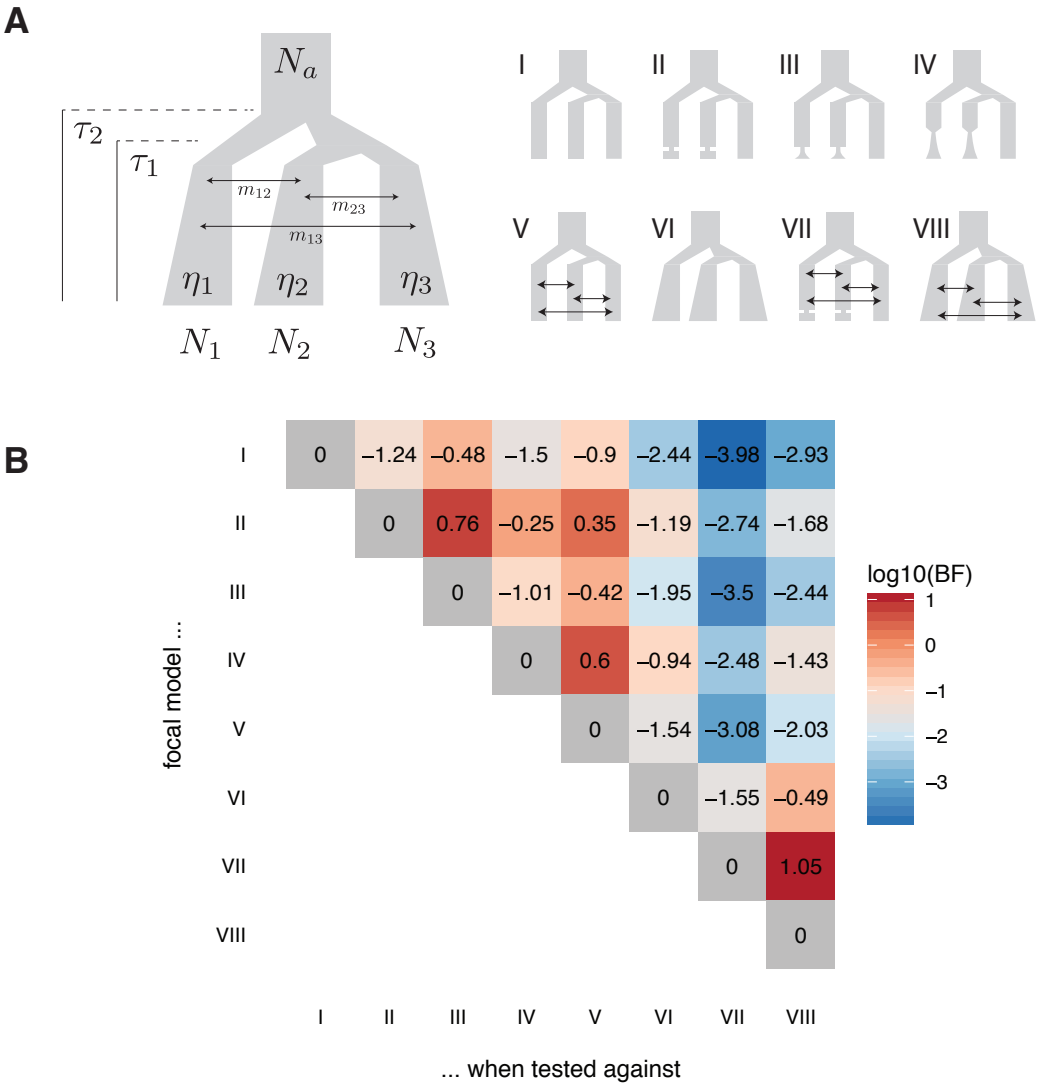

Figure 3: Definition and selection of patrilineal demographic history. (**A**) Schematic representation of eight three-population scenarios used in the simulation step of the ABC procedure. (**B**) Pairwise comparisons of relative goodness-of-fit using Bayes factors (BF). Each cell shows the $\log_{10}$ BF of the fitted model on the y axis against the fitted model on the x-axis. One log-unit represents a 10-fold relative increase in posterior model probability.

Model	Parameter						
	N_0	N_{cas}	N_{dom}	N_{mus}	τ_1	τ_2	β
I	61.2 (23.5, 68.3)	44.2 (22.1, 53.9)	39.3 (17.9, 46.5)	40.4 (15, 44.9)	93 (39.4, 116)	201 (109, 259)	-
II	80.6 (45.7, 91.9)	52.4 (31.7, 62)	59 (36.3, 70.8)	54.5 (31.2, 66.8)	138 (60.6, 151)	260 (147, 297)	0.281 (0.262, 0.463)
III	66.2 (47.9, 72.5)	44.9 (24, 50)	43.4 (30.6, 47.1)	41.7 (26.1, 50.1)	83.5 (36.7, 94.6)	231 (133, 260)	0.242 (0.198, 0.419)
IV	82 (43.5, 98.8)	60.8 (29.8, 70.5)	54.6 (33.6, 74.1)	55.4 (22.3, 66)	133 (37.9, 145)	277 (114, 315)	0.132 (0.0682, 0.215)
V	78.7 (50.8, 85.4)	56.8 (35.1, 63.5)	42.4 (23, 47.2)	56.1 (37.9, 63.2)	185 (96.6, 203)	334 (207, 377)	-
VI	69.9 (39.1, 82.3)	41.8 (16.9, 48.6)	33.2 (11.8, 38.8)	36.3 (16.2, 44.9)	221 (84.9, 270)	456 (256, 577)	-
VII	98.5 (65.4, 116)	64.2 (35.6, 73.6)	54.7 (35.7, 69)	46.9 (13.2, 53.9)	317 (74.2, 352)	636 (269, 725)	0.107 (0.02, 0.128)
VIII	63.2 (42, 68.8)	42.2 (24, 46.7)	20.5 (12.5, 28.9)	27.8 (12.8, 33.6)	226 (137, 261)	427 (247, 458)	-

Model	Parameter					
	η_2	η_3	η_1	m_{12}	m_{13}	m_{23}
III	1.03 (0.494, 1.27)	0.988 (0.393, 1.11)	-	-	-	-
IV	0.941 (0.282, 1.21)	0.973 (0.833, 1.69)	-	-	-	-
V	1.04 (0.291, 1.13)	1.01 (0.453, 1.2)	0.629 (0.237, 0.959)	-	-	-
VI	-	-	-	0.897 (0.309, 1.3)	1.15 (0.598, 1.4)	0.886 (0.0498, 0.981)
VII	-	-	-	0.84 (0.457, 1.29)	0.694 (0.319, 0.924)	0.829 (0.192, 1.15)
VIII	1.01 (0.877, 1.7)	1.13 (0.758, 1.6)	1.03 (0.664, 1.5)	1.27 (0.984, 1.86)	0.904 (0.423, 1.13)	0.77 (0.0893, 0.817)

Table 6: Parameter estimates for demographic models fit to Y chromosome SFS by ABC. Values shown are posterior medians with 50% HPDIs. Models refer to **Figure 3**. Units are as follows: for population sizes, thousands of individuals; for split times, thousands of generations; for growth rates, unitless exponential rate constants; for migration rates, units of $4N_em$, the effective number of migrants per generation.

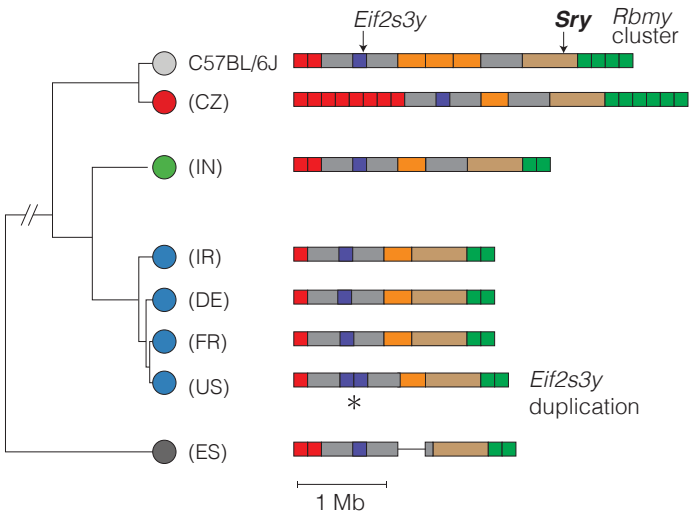

Figure 4: Schematic view of structural variation on the Y chromosome short arm (Yp), superposed on SNV-based phylogenetic tree. Copy-number variable regions are indicated with colored blocks and invariant regions with grey blocks. All CNVs shown overlap a segmental duplication in the reference sequence. Only two CNVs overlap protein-coding genes: a duplication in North American mice encompassing *Eif2s3y* (purple) and an expansion of the ampliconic *Rbmy* cluster (green) in *M. m. musculus*. Color scheme for *Mus* taxa follows **Figure 2**.

ampliconic unit, we find that the proportional composition of Yq is very similar across taxa (**Figure S3A**). However, the total size of Yq varies dramatically within *Mus*: from a median 19 Mb in *M. spretus* to 61 Mb in *M. m. domesticus* (**Figure S3B**). Size differences estimated from whole-genome sequence are supported by cytological observations that the Y chromosomes of wild-caught *M. m. musculus* appear much larger than those of *M. spicilegus* or *M. spretus* (Bulatova and Kotenkova, 1990; Yakimenko et al., 1990).

The intragenomic conflict hypothesis also implies selection at co-amplified regions on the X chromosome. This should reduce nucleotide diversity at sites closely linked to the co-amplified regions relative to sites further away. We calculated nucleotide diversity (θ_π) and Tajima's D in 100 kb windows across the X chromosome in same samples for which we estimated copy number on Yq. Notwithstanding the X-chromosome-wide deficit in nucleotide diversity relative to autosomes, we observed neither additional reduction in diversity in the vicinity of co-amplified regions nor a skew towards low-frequency variants (**Figure S4**). Tests for a linear relationship between diversity and distance from the nearest co-amplified region, or for an ordinal trend across bins of distance, were not significant in any population.

2.6 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 next 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 a 8 million year evolutionary transect across the murid rodents (Neme and Tautz, 2016) (**Figure 6A**). For genes on the auto-

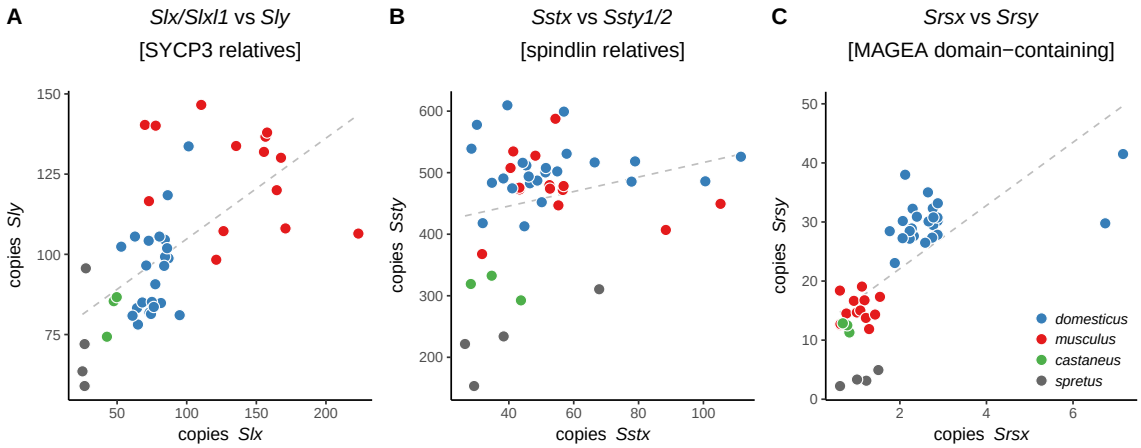

Figure 5: Approximate copy number of co-amplified gene families on X and Yq. Each dot represents a single individual. Grey dashed line is simple linear regression of Y-linked versus X-linked copy number.

somes and X chromosome, the great majority of expression variance lies between tissues rather than between (sub)species (**Figure 6B**). This is not the case for Y-linked genes: the first principal component separates the testis from the other two tissues, but the second principal component (11% of variance explained) separates species.

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 6C**). 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 6D**). 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. This divergence can be attributed to gene families acquired and amplified since the divergence of the sex chromosomes (*Rbm31y*, *H2al2y*, *Sly*, *Ssty1/2*, *Srsy* (Soh et al., 2014)), rather than to gene families present on the ancestral autosome (**Figure 6E**).

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. *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*.

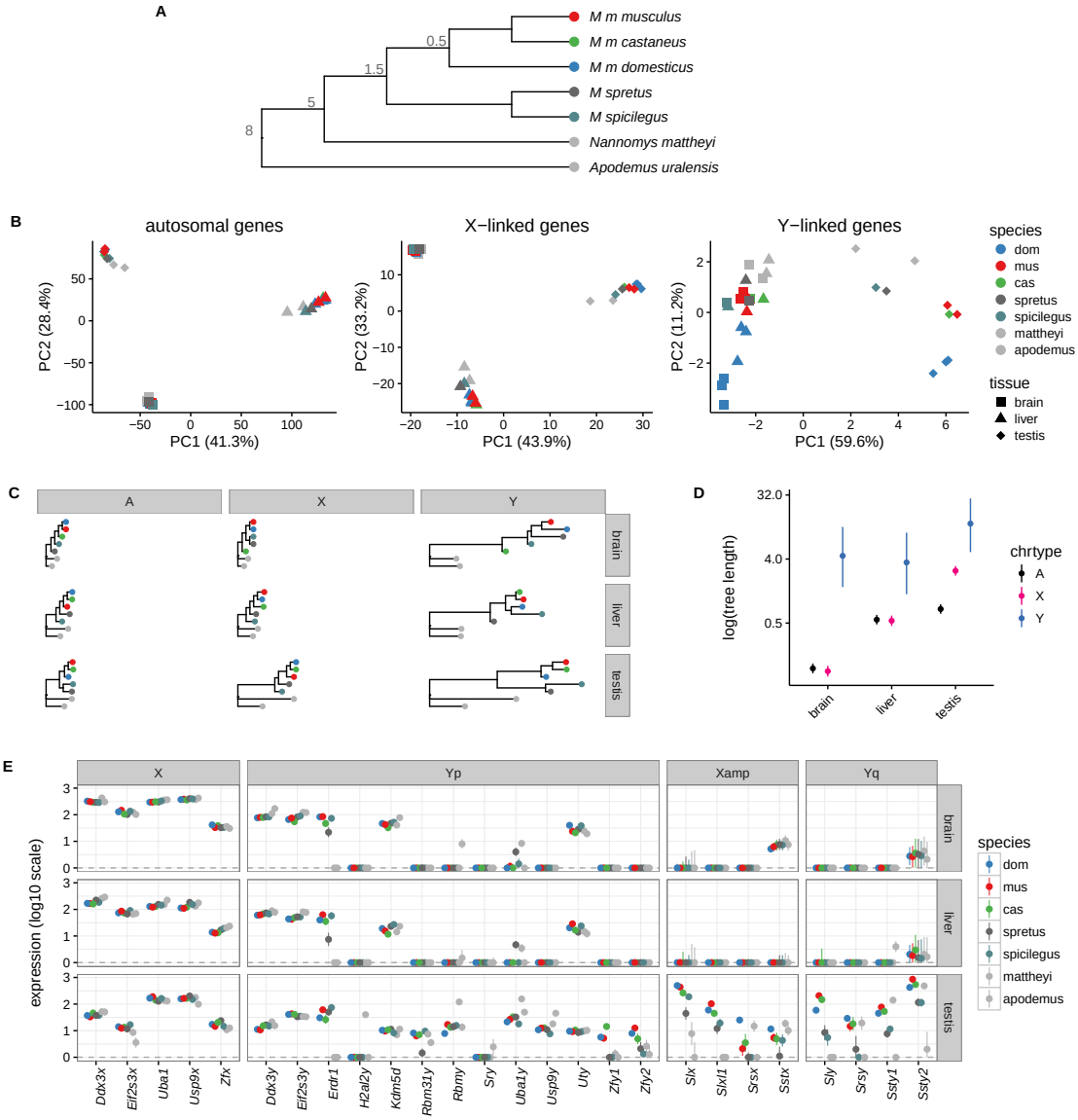

Figure 6: Divergence of sex-linked gene expression in murid rodents. (**A**) Schematic phylogeny of taxa in the multi-tissue expression dataset. Node labels are approximate divergence times (Mya); branch lengths not to scale. (**B**) Projection of samples onto the top two principal components of expression 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 values and standard errors for genes derived from the autosomal autosome pair (panels X, Yp); ampliconic genes on X; and ampliconic genes on Yq.

2.7 Sex-chromosome conflict in spermatids requires unknown loci

Although comparative analyses of gene expression in whole tissues are informative for broad-scale patterns, they cannot distinguish changes in cellular expression levels from allometric shifts in tissue composition. This confounding is particularly relevant to sex-linked genes, which are more likely to have restricted expression patterns and to be specialized for function in the germline in mammals (Mueller et al., 2008, 2013). Genes expressed later in spermatogenesis have been shown to have accelerated rates of sequence evolution in rodents (Good and Nachman, 2005; Ellegren and Parsch, 2007). Changes in the timing of spermatogenesis or the relative abundance of germline versus somatic tissue in the testis might underly apparent differences in expression of the spermatid-specific gene families implicated in sex-chromosome conflict.

To test hypotheses regarding sex-chromosome conflict we therefore took advantage of a recent study of gene expression in specific stages of spermatogenesis (**Figure 7A**; (Larson et al., 2016b,a)). Mitotic (spermatogonia), meiotic (spermatocytes) and postmeiotic (spermatids) cell populations were isolated from testes of reciprocal F_1 hybrids between the wild-derived inbred strains LEWES/EiJ and WSB/EiJ (from Delaware and Maryland, respectively; *M. m. domesticus* ancestry) and PWK/PhJ and CZECHII/EiJ (from the Czech Republic; primarily *M. m. musculus* ancestry). Males from the LEWES/EiJ×PWK/PhJ (*dom*×*mus*) cross are fertile but males from the reciprocal PWK/PhJ×LEWES/EiJ (*mus*×*dom*) cross are nearly sterile due to arrest in late pachytene followed by massive germ cell loss (Turner et al., 2006; Good et al., 2008; Bhattacharyya et al., 2013; Campbell et al., 2013). Using isolated cell populations allows us to estimate gene expression in all possible F_1 genotypes despite these dramatic differences in the cellular composition of the testis. We re-analyzed the RNA-seq data using an augmented transcript annotation which includes a comprehensive set of transcript models for co-amplified genes on Yq and the X chromosome in addition to transcript models in the public Ensembl annotation (see **Materials and methods**). Expression was estimated at the transcript level, and these estimates were aggregated to gene level for analysis.

Our working model posits a dosage imbalance between X-linked (*Slx*, *Slxl1*) and Y-linked (*Sly*) co-amplified gene families as the driver of post-meiotic sex-chromosome conflict in house mice. The model makes four predictions: (1) copy number differences between subspecies at *Slx* and *Sly* lead to differential expression in spermatids; (2) expression balance between *Slx* and *Sly* is therefore disrupted in hybrids relative to intra-subspecific crosses; (3) expression from the sex chromosomes is inversely related to the level of *Sly* expression (since SLY represses the sex chromatin (Cocquet et al., 2009)); and finally (4) offspring sex ratio is distorted in favor of males in hybrids with higher *Sly* expression, and in favor of females in hybrids with higher *Slx* expression.

As reported previously (Ellis et al., 2011), *Sly* (higher copy number in *musculus*) has 2.9-fold (95% CI $2.2 - 3.9$; $p = 5.6 \times 10^{-5}$) higher expression in F_1 spermatids with a *musculus* Y chromosome. Likewise, *Slxl1* (also higher copy number in *musculus*) has 1.8-fold ($1.4 - 2.5$; $p = 6.7 \times 10^{-3}$) higher expression in F_1 spermatids with a *musculus* X chromosome. (Expression differences for *Slx* did not reach significance.) Tests for X-Y interaction effects in spermatids were not significant for *Slx*, *Slxl1* or *Sly*. We estimated the X:Y expression ratio for each co-amplified gene family by F_1 genotype (**Figure 7B**). Expression of Y-linked copies (versus X-linked homologs) of *Sly* and *Srsy* is lower from the *domesticus* than the *musculus* Y chromosome, independent of X chromosome origin. Expression is shifted in favor of the Y in the *mus*×*dom* versus the *dom*×*dom* cross, but is within the range of the two intraspecific crosses. The pattern for *Ssty1* and *Ssty2* relative to *Sstx* is

more complex: expression balance for these genes appears to depend on an X-Y interaction.

Hybrid male sterility in the *mus*×*dom* cross is associated with massive over-expression of the X chromosome in spermatocytes due to failure of MSCI (Campbell et al., 2013; Larson et al., 2016a). To quantify this effect while accounting for baseline differences in expression between subspecies, we repeated the analyses presented in Larson et al. (2016a) and estimated expression contrasts between each inter-subspecific hybrid and the intrasubspecific cross corresponding to the hybrid's X chromosome (*i.e. mus*×*dom* versus *mus*×*mus*; **Figure 7C**). Since MSCI predicts PSCR (**Figure S5**) and *Sly* has lower expression from the *domesticus* Y, the general derangement of sex-chromosome expression in *mus*×*dom* is confounded with any sex-chromosome conflict. However, over-expression of the X (but not the Y) in *mus*×*dom* is mirrored by under-expression in *dom*×*mus* only in spermatids and not in spermatocytes (**Figure 7D**). (Fold-difference values are estimated separately and orthogonally in each hybrid, so this reciprocal pattern is not a consequence of model choice.)

3 Discussion

Because their repetitive sequence content is difficult to assemble by standard methods, Y chromosomes are a "last frontier" of even well-studied mammal genomes. Yet Y chromosomes are extremely variable in content and organization between species, and are valuable indicators of male-specific mutational and demographic processes. In this manuscript we exploit public whole-genome sequencing data to characterize the landscape of small sequence variants and large copy-number variants on the Y chromosome in natural populations of mice. We integrate sequence data with gene expression data from both wild mice and laboratory strains to demonstrate functional divergence of Y chromosome lineages. In particular, we attempt to evaluate support for the hypothesis that evolutionary trajectory of the mouse Y chromosome has been dictated by intragenomic conflict with the X chromosome, mediated by amplified gene families acquired since the divergence of the mammal sex chromosomes from the ancestral autosome pair.

3.1 Phylogeography of mouse Y chromosomes

We confirm and strengthen the long-standing observation that at least two Y haplogroups are present in classical laboratory strains and related outbred stocks (Bishop et al., 1985; Tucker et al., 1992). One haplogroup falls within the *M. m. musculus* clade almost certainly originated in Japan and represents part of the *M. m. molossinus* contribution to classical inbred strains (Yang et al., 2007, 2011). The last common ancestor of Y chromosomes in this haplogroup was recent: within the last $336 - 8,700$ years. The other two haplogroups are of *M. m. domesticus* origin. One is present in "Swiss" mice such as NOD/ShiLtJ and FVB/NJ and has closest affinity to Y chromosomes found in present-day northern and central Germany, while the other is found in American strains and is not clearly associated with a sampled European lineage (**Figure 2**).

Among Y chromosomes from wild mice, phylogenetic affinity mirrors geography. The same cannot be said for the mitochondria, which are both more genetically diverse within populations and less differentiated between them: see **Figure 2**. The correlation between geographic origin and phylogenetic distance is $\rho = 0.24$ (95% CI $0.21 - 0.27$) for the Y chromosome but only $\rho = 0.10$ (95% CI $0.08 - 0.13$) for the mitochondrial genome. We found one case of inter-specific introgression

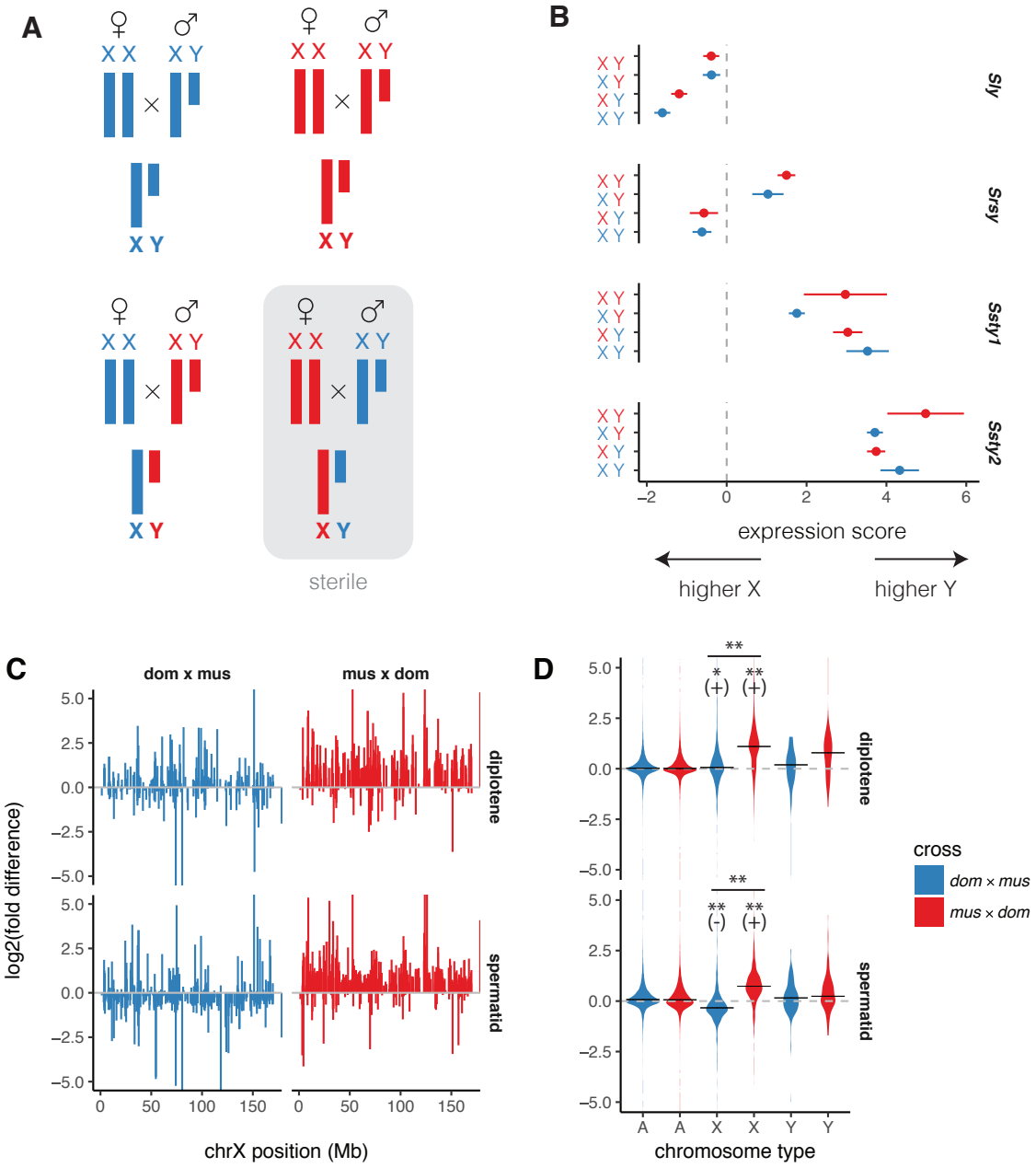

Figure 7: Disruption of sex-chromosome expression in hybrids. (**A**) Experimental design of Larson et al. (2016a): intra-subspecific (top panel) and reciprocal inter-subspecific (bottom panel) F_1 hybrids between wild-derived strains. *M. m. domesticus* shown in blue and *M. m. musculus* in red. (**B**) Expression score (see **Materials and methods**), measuring relative expression from X- and Y-linked members of co-amplified gene families in round spermatids of males from the crosses shown in panel A. (**C**) Fold-difference between expression in inter-subspecific hybrids and the intra-subspecific cross corresponding to to the hybrid's X chromosome, in diplotene spermatocytes (top) and round spermatids (bottom.) (D) Distribution of fold-difference values on autosomes (A), X chromosome (X) and Y chromosome (Y) in each cell type. Horizontal lines indicate group medians. $**(+)$, median logFC > 0 by Wilcoxon signed-rank test ($p < 10^{-5}$); $**(+)$, median logFC < 0. Panels C and D reproduce the analyses of Larson et al. (2016a).

involving a *M. spretus* female and a *M. m. domesticus* male. Taken together, these observations indicate that the degree of genetic mixing is greater for female than male lineages. Several explanations are possible. First, dispersal behavior may differ between sexes. There is little evidence to support the conjecture that female mice disperse more readily than males. If anything, the opposite is true (Maly et al., 1985; Rowe et al., 1987; Pocock et al., 2005), but females are generally more successful at integrating into a new group than males (Lidicker, 1976; Pocock et al., 2005). Second, genetic incompatibilities may accumulate more rapidly on the Y chromosome and serve as a barrier to gene flow. Studies of the *domesticus-musculus* hybrid zone in eastern Europe have consistently shown that allele-frequency clines are narrower and steeper for Y-linked than autosomal loci (Teeter et al., 2008; Macholán et al., 2008), and hybrid male sterility constitutes the primary reproductive barrier between mouse subspecies (Forejt and Iványi, 1974). However, it is seems unlikely that genetic incompatibilies would arise within $5,000 - 10,000$ generations between local populations of the same subspecies (*e.g* in France and Germany). Finally, the lack of apparent geographic mixing between male lineages may simply be a consequence of the shallow coalescent times of Y chromosomes.

3.2 What explains the deficit of Y-linked sequence variation?

We have shown that genetic diversity on the sex chromosomes of house mice is markedly reduced compared to the autosomes. This pattern could be a consequence of natural selection on the sex chromosomes — such as selection due to intragenomic conflict between the X and Y — or it could simply reflect neutral demographic forces. Formally discriminating the effects of non-equilibrium demography versus selection is very difficult, and neutral and non-neutral scenarios need not be mutually exclusive.

In the absence of selection and assuming equal mutation rates at all loci, genetic diversity is proportional to the (effective) number of chromosomes in the population (Wright, 1931). Expected diversity on the Y chromosome, which is hemizygous and only passed through the male germline, is therefore only one-fourth that of the autosomes, which are diploid and passed through both sexes, in a population with sex ratio at parity (Charlesworth et al., 1987). Departures from these ratios can be a signal of (1) unequal sex ratio; (2) sex differences in mutation rate; (3) population size changes; or (4) selection (Webster and Wilson Sayres, 2016).

Our data are not consistent with a skewed sex ratio. An excess of males versus females in the population increases Y:A and decreases X:A relative to the neutral expectation, and an excess of females has the reverse effect (Charlesworth et al., 1987); but we observe marked reduction in both X:A and Y:A (**Table 3**). Our estimate of Y:X (0.53 ± 0.05 in *M. m. domesticus* and *M. m. musculus*) is significantly greater than the expected value of $\frac{1}{3}$. This discrepancy can be explained in part by the quite strong reduction in diversity across the entire X chromosome relative to the autosomes (**Table 3**). Differences in germline mutation rate also likely contribute: in mammals, the mutation rate is generally higher in males than females (although the details of the relationship α depend on life history (Ségurel et al., 2014)). The Y:X diversity ratio we observe is consistent with $\alpha \approx 3$ (Webster and Wilson Sayres, 2016), in good agreement with the empirical estimate of $\alpha = 2.78$ (Drost and Lee, 1995).

A bottleneck is the most parsimonious neutral explanation for the deficit of nucleotide diversity we observe across both sex chromosomes in European populations and is consistent with established features of the demographic history of house

mice. Sex chromosomes, because of their smaller effective population sizes and consequently shallower coalescent times, are more sensitive to the effects of population growth and contraction than autosomes. Both X:A and Y:A decline during a bottleneck (Pool and Nielsen, 2007), and subsequent recovery in population size leaves behind an excess of low-frequency alleles (Tajima, 1989). This is exactly what we observe in *M. m. domesticus* and *M. m. musculus* for both sex chromosomes (**Table 2**). The demographic models we fit to Y chromosome SFS via ABC support a strong bottleneck in *M. m. domesticus* and *M. m. musculus*, the populations with the greatest reductions in X:A and Y:A diversity (**Figure 3**). The timing of the inferred bottleneck ($19,700$ generations in the past) is consistent with fossil and genetic evidence that Eurasian mammal populations experienced a sharp contraction around the time of the last glacial maximum $10,000 - 25,000$ years ago (Auffray et al., 1990; Hofreiter et al., 2004; Macholán et al., 2012). Populations in central Asia — the ancestral range of house mice — are generally considered members of the *castaneus* subspecies but retain more genetic and diversity than European populations represented in this study (Rajabi-Maham et al., 2012). The consensus of several decades of genetic studies of wild mice is that *M. m. musculus* and *M. m. domesticus* shed much of their ancestral diversity due to the combined effects of climate and migration out of central Asia and into the Levant and Europe, and that large census population sizes in the present are the result of expansions tied to commensalism with humans (Boursot et al., 1993; Din et al., 1996; Geraldes et al., 2008).

Reduced (on the X) or absent (on the Y) recombination also makes sex chromosomes especially sensitive to background selection (Hudson and Kaplan, 1995). Intragenomic conflict between the sex chromosomes should also reduce diversity on both X and Y. If conflict exerts selective pressure on co-amplified gene families on the X, we expect sequence diversity to be lower near these clusters than far away — a pattern we fail to detect. Population size changes, background selection and intragenomic conflict are all three likely to contribute to the reduced diversity on the sex chromosomes, but it seems unlikely that intragenomic conflict dominates.

3.3 Mutational mechanisms on the Y chromosome

The Y chromosome provides a direct view of the mutational spectrum of the male germline. Using this fact we estimate the male-specific point mutation rate in mouse to be $5.4 \times 10^{-9} - 8.1 \times 10^{-9}$ bp^{-1} generation^{-1}. We show, for the first time, that structural variation on Yq is abundant in wild populations: the Yq has more than tripled in size in less than 2 My between the divergence of *M. spretus* and *M. m. domesticus* (**Figure S3**). A high mutation rate is almost certainly required to generate such variation. Clusters of duplicated sequences are often assumed to be especially mutable because they are prone to non-allelic homologous recombination (Carvalho and Lupski, 2016), but this is trivially not the case for the male-specific portion of the Y chromosome, which has no homologous partner with which to pair or recombine. Structural variation on the Y must therefore arise via errors of replication during mitosis or by intrachromosomal recombination. In humans and other great apes, exchange between duplicated sequences on opposite arms of the metacentric primate Y chromosome appears to be common (Rozen et al., 2003; Repping et al., 2006; Hallast et al., 2013).

3.4 Evidence of X-Y 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). This has led several authors to conclude that the evolution of the mouse sex chromosomes is driven by recurrent intragenomic conflicts whose principal actors are members of these "co-amplified" gene families. Consistent with this prediction, one pair of co-amplified families, *Slx* and *Sly*, have opposing actions on the post-meiotic sex chromatin and weakly distort the sex ratio in favor of their own chromosome in laboratory crosses (Hendriksen et al., 1995; Cocquet et al., 2009, 2012).

In this manuscript we attempt to synthesize available evidence for sex-chromosome conflict. Sequence data in wild mice and expression in the germline provide compelling support for the conflict hypothesis. The expression patterns of ampliconic genes on Yq are diverging much more rapidly than autosomal genes within the murid rodents (**Figure 6**). We find that the copy number of *Slx/Slxl1* and *Sly* are indeed correlated across natural populations of house mice (**Figure 5**). In the context of *M. m. musculus* and *M. m. domesticus*, expression of X-linked genes in inter-subspecific hybrids "mismatched" for X- and Y-linked copy number is also consistent with sex-chromosome conflict. In the hybrid with higher *Sly* expression, X-linked gene expression is reduced relative to the conspecific parent; in the hybrid with lower *Sly* expression, X-linked gene expression in spermatids is increased (**Figure 7D**). Transcriptional control of the sex chromosomes in spermatids has been predicted to depend on the balance between *Slx* and *Sly* Cocquet et al. (2012). The data presented here, however, contradict that simple prediction: the balance between *Slx* and *Sly* expression in spermatids differs between both two intra-subspecific hybrids with normal fertility and and no X:Y copy number mismatch (**Figure 7B**). Sex-chromosome conflict in house mice must therefore depend on other loci in addition to *Slx* and *Sly*.

It is less clear whether molecular conflict in spermatids has meaningful effects on male reproductive success or offspring sex ratio. Sex-ratio distortion has been observed in the offspring of males with X:Y copy-number mismatch in some experiments (Cocquet et al., 2009; Case et al., 2015) but not in others (Turner et al., 2012; Albrechtová et al., 2012). We cannot rule out the possibility that detrimental effects of X:Y imbalance on sperm morphology or function (Campbell and Nachman, 2014) reduce reproductive success under natural breeding conditions, especially in the context of intense sperm competition (Firman and Simmons, 2008). However, among 450 recombinant lines from the Collaborative Cross population, in which X- and Y-chromosome haplotypes from all three house mouse subspecies are segregating, none of more than a dozen testis histology or sperm function traits map to the Y (John Shorter, Deborah O'Brien and Fernando Pardo-Manuel de Villena, in preparation).

We note that Y-associated sex-ratio distortion in laboratory crosses has only been observed (to our knowledge) under two conditions. The first condition is profound reduction in transcription from Yq, either by experimental knockdown (Cocquet et al., 2009, 2012) or as a result of large deletions encompassing more than two-thirds of Yq (Burgoyne et al., 1992; Conway et al., 1994; Touré et al., 2004). A smaller deletion of approximately half of Yq was not associated with subfertility or sex-ratio distortion (Fischer et al., 2016). The second condition is pairing of a *M. m. domesticus* Y chromosome with the X chromosome of the classical inbred strain C57BL/6J (Case et al., 2015). Evolutionary inferences from classical inbred strains

are complicated by the fact that they are admixed (Yang et al., 2007) and, as domesticated animals, have been under strong selection for fertility and husbandry traits.

A critical evaluation of the X-Y intragenomic conflict hypothesis in house mice must also account for the prominent role of the X chromosome in reproductive isolation between subspecies (Larson et al., 2016a). Because of the upstream disruption of MSCI in *M. m. musculus*-*M. m. domesticus* hybrids, this cross is not the ideal one for studying the effect of Y-linked factors on fertility. Examination of other genotypes which differ in copy number at X- and Y-ampliconic genes but have intact MSCI would provide useful insight. Hybrid zones besides the well-studied European zone — such as one between *M. m. domesticus* and *M. m. castaneus* in New Zealand (McCormick et al., 2014) — might provide convenient natural experiments.

3.5 Concluding remarks

Our analyses provide an expanded view of the evolution of the Y chromosome and the demographic history of male lineages in the genus *Mus*. Both the distribution of X- and Y-linked copy number and gene expression in spermatids are consistent with ongoing intragenomic conflict between the sex chromosomes. However, the mechanistic basis for this antagonism cannot be reduced to a simple interaction between the *Slx/Slxl1* and *Sly* gene families. The phenotypic consequences of sex-chromosome conflict *per se* are subtle at best. Furthermore, although sex-chromosome conflict could plausibly reduce genetic diversity on both the X and Y chromosomes, other factors — including population size changes and background selection — are probably more important.

Many open questions remain with respect to the origin and evolution of mouse-specific Y-linked gene families. Although we have documented large variation in copy number of ampliconic, mouse-specific gene families on Yq in natural mouse populations, we can say little about their higher-order organization. Nor can we determine how many gene copies in each family retain coding potential. Addressing these questions will require alternative sequencing technologies that provide longer reads and long-range physical linkage information. Which of the many copies of ampliconic gene families on X and Y are functionally equivalent, and the consequences of sequence and structural variation of particular copies for male reproductive traits, are important avenues of future study.

4 Materials and methods

4.1 Alignment and variant-calling

Whole-genome sequencing 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 in callable regions were identified using `freebayes` v1.0.2 (Garrison and Marth, 2012). Variants were called in all samples jointly. Reads with MQ < 10, basecall quality < 13 and > 9 mismatches (to the reference sequence) were excluded. Candidate variant sites were required to have read depth > 3 and at most 3 alleles. The raw call set was filtered to have quality score > 30 and per-sample depth > 3, all heterozygous genotypes were treated as missing to reflect the haploid nature of the Y.

Filtered variants were normalized to their atomic SNV or indel representation using `vcflib`. Functional consequences were predicted using `SnpEff` (Cingolani et al., 2012) using the most recent annotation database available (GRCm38/Ensembl 82).

4.2 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.

4.3 Estimation of site frequency spectra

Site frequency spectra (SFS) for the Y chromosome were calculated from genotype likelihoods at callable sites using `ANGSD` v0.910-133-g68dd0f2 (Korneliussen et al., 2014). Genotype likelihoods for the Y chromosome were calculated under the `GATK` haploid model after applying base alignment quality (BAQ) recalibration with the recommended settings for bwa alignments (`-baq 1 -c 50`). Only reads with MQ > 20 and bases with call quality > 13 were considered. Sites were filtered to have per-individual coverage consistent with the presence of a single haploid copy ($3 <$ depth < 40), and to be non-missing in at least 3 individuals per population. Site-wise allele frequencies were computed within each population separately, and the joint SFS across the three populations was estimated from these frequencies. The consensus genotype over 5 *M. spretus* males was used as the ancestral sequence to polarize alleles as ancestral or derived. For estimating uncertainties in diversity statistics, 100 bootstrap replicates were obtained for the joint SFS.

SFS for the X chromosome were estimated using the same parameters but with the consensus haploid genotype from a single *M. caroli* female as the ancestral sequence. For the mitochondria, different filtering criteria were used ($10 < \text{depth} < 1000$) to reflect differences in expected coverage for this organellar genome. *M. caroli* was again used as the ancestral sequence. For estimating the autosomal SFS we used representative sequence from chromosome 1 and used a diploid rather than haploid model for genotype likelihoods.

Some inconsistencies may arise due to the use of different outgroup species, at different evolutionary distances, for the autosomes, X, Y and mitochondria. We unfortunately did not have access to whole-genome sequence from a male more divergent than *M. spretus* to use as an outgroup for the Y. However, because hybrid offspring of a *M. musculus* dam and a *M. spretus* sire are generally sterile (Macholán et al., 2012), there is little change of introgression of a *M. spretus* Y chromosome into *M. musculus*. Nor did we find evidence for incomplete lineage sorting of Y chromosomes between *M. spretus* and *M. musculus* in our dataset.

4.4 Diversity statistics

Diversity statistics and neutrality tests were calculated from joint SFS using the `R` package `sfsr` (`http://github.com/andrewparkermorgan/sfsr`). Hudson-Kreitman-Aguade (HKA) tests were performed with `sfsr` and p-values obtained from the χ^2 distribution with a single degree of freedom as suggested in (Hudson et al., 1987). (Results were checked against the `HKA` software from Jody Hey, in which significance thresholds are set via coalescent simulations; all significant tests were significant under both methods.)

4.5 Demographic inference

Possible demographic scenarios for male lineages in *M. musculus* were explored using approximate Bayesian computation (ABC). All scenarios modelled three populations (corresponding to *M. m. domesticus*, *M. m. musculus* and *M. m. castaneus*) derived from a single ancestral population. The order of population splits was (castaneus,(musculus,domesticus)) — reflecting the phylogeny in **Figure 2** — and was kept fixed across all scenarios. Eight scenarios were tested: (I) constant population size, no migration; (II) recent bottleneck shared by *M. m. domesticus* and *M. m. musculus*; (III) recent bottleneck, followed by exponential growth; (IV) distant bottleneck, followed by exponential growth; (V) constant population size, with migration; (VI) exponential growth at independent rates, no migration; (VII) recent bottleneck, with migration; (VIII) exponential growth, with migration.

Briefly, $100,000$ simulations were performed for each model using parameter values drawn from uninformative or weakly-informative prior distributions. Fifteen summary statistics were calculated from the joint SFS generated by each simulation: number of segregating sites in each population (3); Tajima's D, Fu and Li's D and F in each population (9); and F_{st} between all population pairs (3). 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. Posterior distributions were computed via kernel smoothing over the parameter values of the retained simulations using the `R` package `abc` (Csilléry et al., 2012).

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

4.6 Analyses of gene expression

Multi-tissue dataset of murid rodents. Neme and Tautz (Neme and Tautz, 2016) measured gene expression in whole testis from wild-derived outbred mice from several species (**Figure 6A**) 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.

Reciprocal F_1 hybrids. 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. Expression contrasts were estimated using the empirical Bayes procedure implemented in the `R` package `limma` (Ritchie et al., 2015).

To compare the relative expression levels of X- and Y-linked members of co-amplified gene families, we defined the "expression ratio" $z = \frac{y}{(x+y)}$ and transformed it to a log-odds "expression score" R:

$$R = \log \frac{z}{1-z}$$

514 The standard error of this quantity was calculated within each cross and within each gene family by the delta method
515 (Oehlert, 1992) as implemented in the `deltamethod()` function of `R` package `msm`.

516 Acknowledgments

517 The authors thank Jeff Good, Erica Larson, Michael Nachman, Megan Phifer-Rixey, Jacob Mueller, Alyssa Kruger and Marty
518 Ferris for insightful suggestions. This work was supported by the National Institutes of Health (F30MH103925 to APM).

519 References

520 Achaz G. 2008. Testing for neutrality in samples with sequencing errors. *Genetics*. 179:1409–1424.

521 Albrechtová J, Albrecht T, Baird SJE, Macholán M, Rudolfsen G, Munclinger P, Tucker PK, Piálek J. 2012. Sperm-related
522 phenotypes implicated in both maintenance and breakdown of a natural species barrier in the house mouse. *Proc. Biol. Sci.*
523 279:4803–4810.

524 Auffray JC, Vanlerberghe F, Britton-Davidian J. 1990. The house mouse progression in Eurasia: a palaeontological and
525 archaeozoological approach. *Biological Journal of the Linnean Society*. 41:13–25.

526 Beaumont MA, Zhang W, Balding DJ. 2002. Approximate Bayesian computation in population genetics. *Genetics*. 162:2025–
527 2035.

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

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

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

534 Bhattacharyya T, Gregorova S, Mihola O, Anger M, Sebestova J, Denny P, Simecek P, Forejt J. 2013. Mechanistic basis of
535 infertility of mouse intersubspecific hybrids. *PNAS*. 110:E468–E477.

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

538 Boursot P, Auffray JC, Britton-Davidian J, Bonhomme F. 1993. The Evolution of House Mice. *Annual Review of Ecology and*
539 *Systematics*. 24:119–152.

540 Brawand D, Soumillon M, Necsulea A, et al. (18 co-authors). 2011. The evolution of gene expression levels in mammalian
541 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.

Bulatova N, Kotenkova E. 1990. Variants of the Y-chromosome in sympatric taxa of Mus in southern USSR. *Bolletino di zoologia*. 57:357–360.

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.

Carvalho CMB, Lupski JR. 2016. Mechanisms underlying structural variant formation in genomic disorders. *Nat. Rev. Genet.* 17:224–238.

Case LK, Wall EH, Osmanski EE, Dragon JA, Saligrama N, Zachary JF, Lemos B, Blankenhorn EP, Teuscher C. 2015. Copy number variation in Y chromosome multicopy genes is linked to a paternal parent-of-origin effect on CNS autoimmune disease in female offspring. *Genome Biology*. 16:28.

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

Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly*. 6:80–92.

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.

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.

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.

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, Parsch J. 2007. The evolution of sex-biased genes and sex-biased gene expression. *Nat Rev Genet*. 8:689–698.

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.

Firman RC, Simmons LW. 2008. Polyandry, sperm competition, and reproductive success in mice. *Behavioral Ecology*. 19:695–702.

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

Fu YX, Li WH. 1993. Statistical tests of neutrality of mutations. *Genetics*. 133:693–709.

Garrison E, Marth G. 2012. Haplotype-based variant detection from short-read sequencing. *arXiv:1207.3907 [q-bio]*. ArXiv: 1207.3907.

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.

Good JM, Nachman MW. 2005. Rates of Protein Evolution Are Positively Correlated with Developmental Timing of Expression During Mouse Spermatogenesis. *Mol Biol Evol*. 22:1044–1052.

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

Hallast P, Balaresque P, Bowden GR, Ballereau S, Jobling MA. 2013. Recombination Dynamics of a Human Y-Chromosomal Palindrome: Rapid GC-Biased Gene Conversion, Multi-kilobase Conversion Tracts, and Rare Inversions. *PLOS Genet*. 9:e1003666.

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.

Hofreiter M, Serre D, Rohland N, Rabeder G, Nagel D, Conard N, Münzel S, Pääbo S. 2004. Lack of phylogeography in European mammals before the last glaciation. *Proc Natl Acad Sci U S A*. 101:12963–12968.

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

Hudson RR, Kreitman M, Aguadé M. 1987. A test of neutral molecular evolution based on nucleotide data. *Genetics*. 116:153–159.

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

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.

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

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.

628 Lidicker WZ. 1976. Social Behaviour and Density Regulation in House Mice Living in Large Enclosures. *Journal of Animal*
629 *Ecology*. 45:677–697.

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

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

634 Macholán M, Baird SJE, Munclinger P, Piálek J, editors. 2012. Evolution of the House Mouse. New York: Cambridge Univer-
635 sity Press, 1 edition edition.

636 Maly MS, Knuth BA, Barrett GW. 1985. Effects of Resource Partitioning on Dispersal Behavior of Feral House Mice. *Journal*
637 *of Mammalogy*. 66:148–153.

638 McCormick H, Cursons R, Wilkins RJ, King CM. 2014. Location of a contact zone between Mus musculus domesticus and M.
639 m. domesticus with M. m. castaneus mtDNA in southern New Zealand. *Mammalian Biology - Zeitschrift für Säugetierkunde*.
640 79:297–305.

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

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

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

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

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

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

654 Nishioka Y, Lamothe E. 1986. Isolation and characterization of a mouse Y chromosomal repetitive sequence. *Genetics*. 113:417–
655 432.

656 Oehlert GW. 1992. A Note on the Delta Method. *Am Stat*. 46:27–29.

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.

Pocock MJO, Hauffe HC, Searle JB. 2005. Dispersal in house mice. *Biological Journal of the Linnean Society*. 84:565–583.

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.

Rajabi-Maham H, Orth A, Siahsarvie R, Boursot P, Darvish J, Bonhomme F. 2012. The south-eastern house mouse Mus musculus castaneus (Rodentia: Muridae) is a polytypic subspecies. *Biol J Linn Soc Lond*. 107:295–306.

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.

Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43:e47.

Rowe FP, Quy RJ, Swinney T. 1987. Recolonization of the buildings on a farm by house mice. *Acta Theriologica*. 32:3–19.

Rozen S, Skaletsky H, Marszalek JD, Minx PJ, Cordum HS, Waterston RH, Wilson RK, Page DC. 2003. Abundant gene conversion between arms of palindromes in human and ape Y chromosomes. *Nature*. 423:873–876.

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

Ségurel L, Wyman MJ, Przeworski M. 2014. Determinants of mutation rate variation in the human germline. *Annu Rev Genomics Hum Genet*. 15:47–70.

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.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*. 123:585–595.

Tavaré S, Balding DJ, Griffiths RC, Donnelly P. 1997. Inferring Coalescence Times From DNA Sequence Data. *Genetics*. 145:505–518.

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.

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.

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 JMA, Mahadevaiah SK, Ellis PJI, Mitchell MJ, Burgoyne PS. 2006. Pachytene Asynapsis Drives Meiotic Sex Chromosome Inactivation and Leads to Substantial Postmeiotic Repression in Spermatids. *Developmental Cell*. 10:521–529.

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.

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.

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

Wright S. 1931. Evolution in Mendelian Populations. *Genetics*. 16:97–159.

Yakimenko LV, Korobitsyna KV, Frisman LV, Muntianu AI. 1990. Cytogenetic and biochemical comparison of Mus musculus and Mus hortolanus. *Experientia*. 46:1075–1077.

Yang H, Bell TA, Churchill GA, Pardo-Manuel de Villena F. 2007. On the subspecific origin of the laboratory mouse. *Nat Genet*. 39:1100–1107.

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.

5 Supplementary material

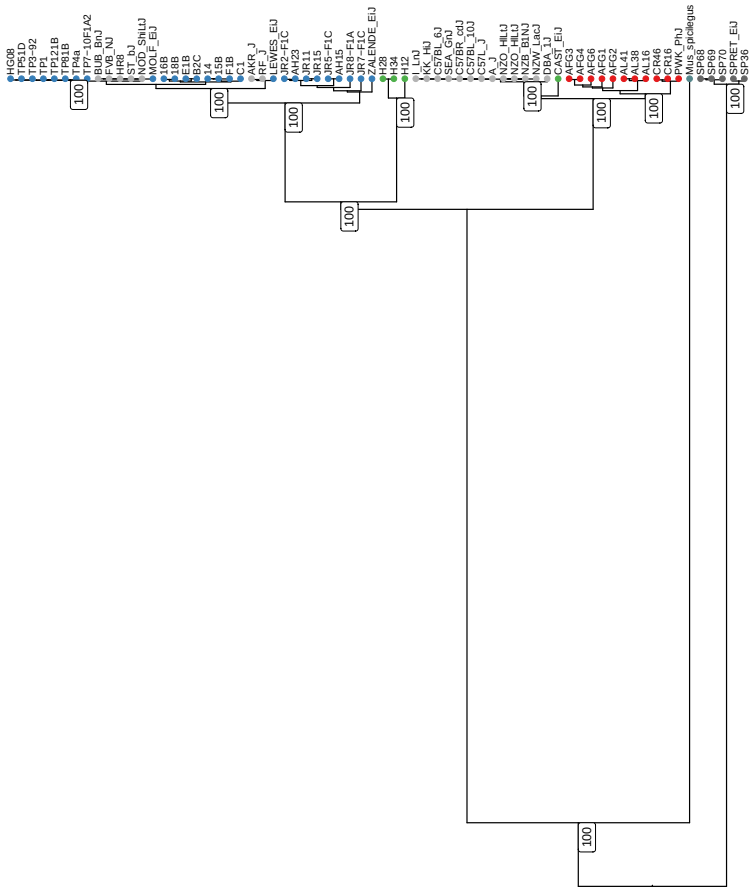

Figure S1: BEAST tree for Y chromosomes showing individual sample names and posterior support (as percentage) for key nodes. Samples are colored by subspecies following the same scheme as elsewhere in the manuscript.

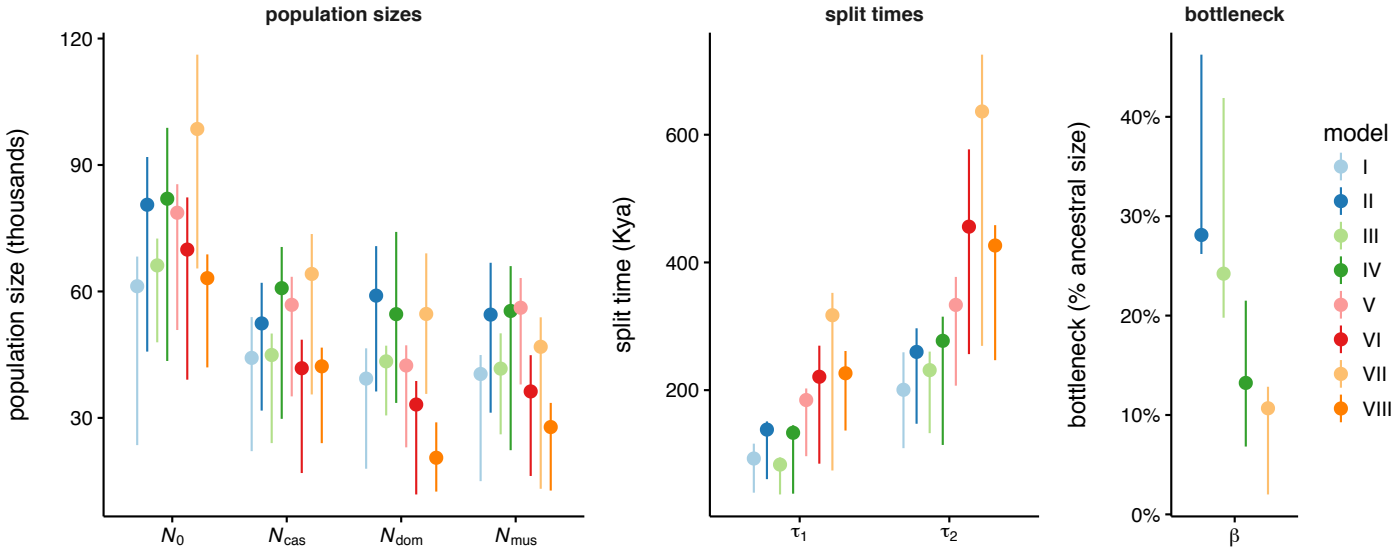

Figure S2: Marginal posterior distributions of key demographic parameters, shown as posterior median and 50% HPDI. Notation follows **Figure 3**.

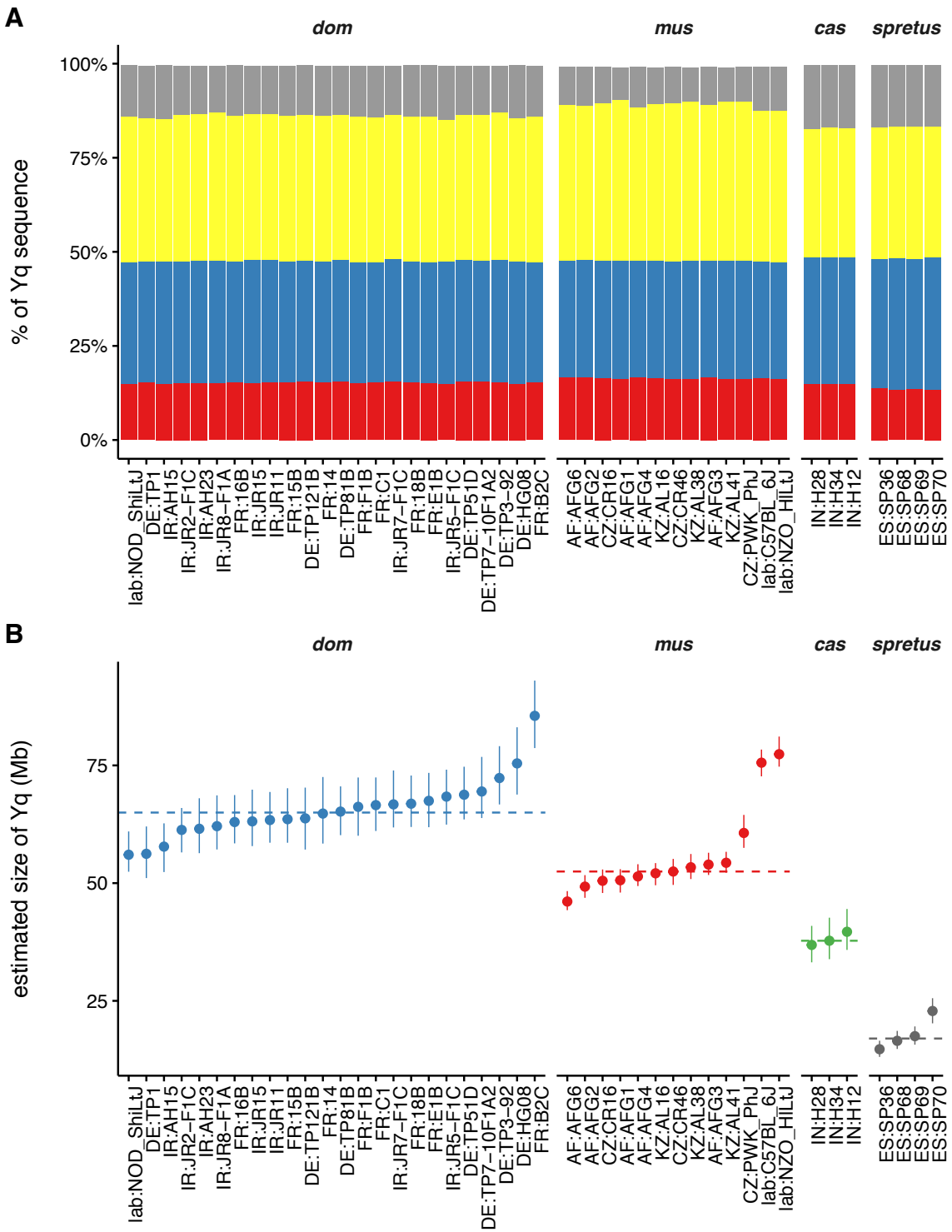

Figure S3: Structural variation on the Y chromosome long arm (Yq). (**A**) Proportional composition of Yq from wild mice and selected laboratory strains of all three subspecies plus *M. spretus*, according to "red," "yellow" and "blue" and "other" sequence families defined in (Soh et al., 2014). Each column corresponds to a single sample; sample names are prefaced by country of origin. (**B**) Estimated total size of Yq for the samples shown in panel A. Dashed lines indicate within-subspecies median.

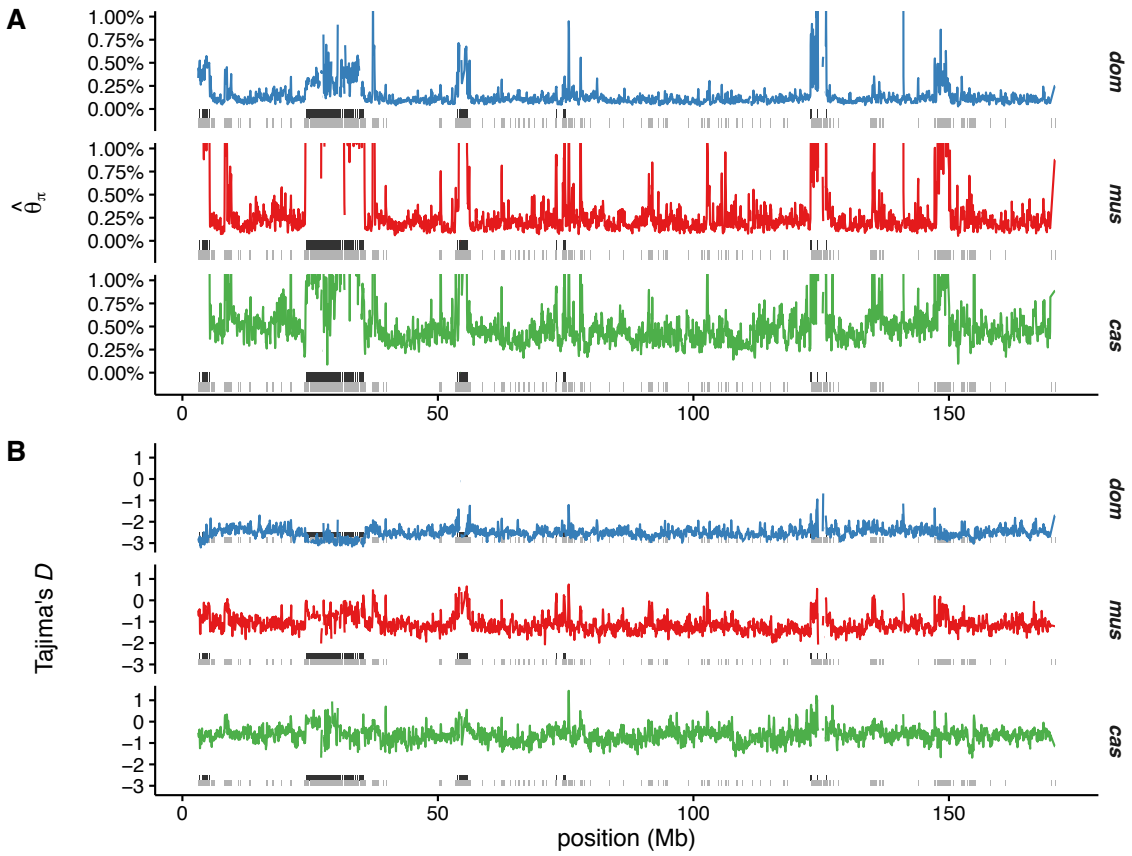

Figure S4: Sequence diversity across the X chromosome. (**A**) Within-population sequence diversity across the X chromosome, measured by Tajima's pairwise estimator θ_π. Dark grey boxes below the x-axis show locations of co-amplified regions; light grey boxes show all segmental duplications > 1 kb in size. Large vertical deviations are likely artifacts associated with paralogous variation associated with segmental duplications. (**B**) As above, but showing Tajima's D.

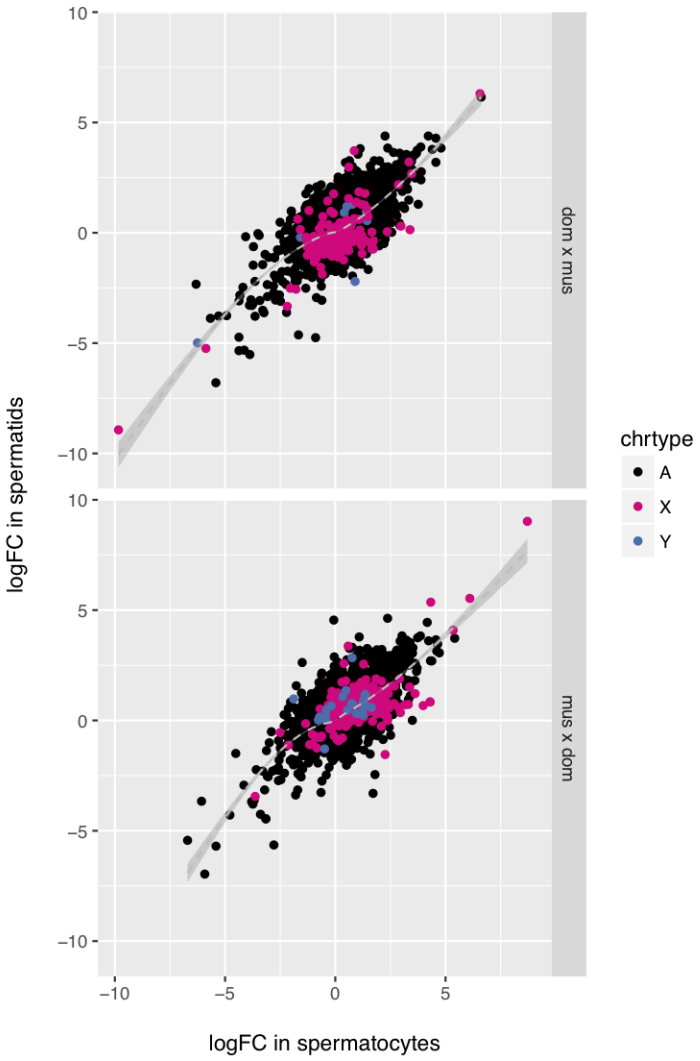

Figure S5: Differential expression in spermatocytes predicts differential expression in spermatids. Fold-differences are for the hybrid genotype versus the intra-subspecific cross corresponding to the hybrid's X chromosome (for X-linked genes) or Y chromosome (for Y-linked genes), and versus the mean of the intrasubspecific crosses for autosomes.