

1 **The contribution of sex chromosome conflict to disrupted spermatogenesis in**  
2 **hybrid house mice**

3 Emily E. K. Kopania<sup>\*,1</sup>, Eleanor M. Watson<sup>‡</sup>, Claudia C. Rathje<sup>§</sup>, Benjamin M. Skinner<sup>‡</sup>,  
4 Peter J. I. Ellis<sup>§</sup>, Erica L. Larson<sup>†,2</sup>, Jeffrey M. Good<sup>\*,1,2</sup>

5  
6 <sup>\*</sup>Division of Biological Sciences, University of Montana, Missoula, MT 59812 USA

7 <sup>‡</sup>School of Life Sciences, University of Essex, Colchester, CO4 3SQ, UK

8 <sup>§</sup>School of Biosciences, University of Kent, Canterbury, CT2 7NJ, UK

9 <sup>†</sup>Department of Biological Sciences, University of Denver, Denver, CO 80208 USA

10 <sup>1</sup>Corresponding authors: [emily.kopania@umconnect.umt.edu](mailto:emily.kopania@umconnect.umt.edu),

11 [jeffrey.good@umontana.edu](mailto:jeffrey.good@umontana.edu)

12 <sup>2</sup>Co-senior authors

13 ORCID ID: 0000-0002-2710-2491 (EEKK)

14 ORCID ID: 0000-0002-1290-4360 (EMW)

15 ORCID ID: 0000-0002-7152-1167 (BMS)

16 ORCID ID: 0000-0001-9709-7934 (PJIE)

17 ORCID ID: 0000-0003-3006-645X (ELL)

18 ORCID ID: 0000-0003-0707-5374 (JMG)

19

20

21 **Running Title:** Sex Chromosomes and Hybrid Sterility

22 **Keywords:** speciation, intragenomic conflict, ampliconic genes, hybrid male sterility,  
23 testis expression, FACS, sex chromosomes

24 **Corresponding Authors:**

25 Emily E. K. Kopania

26 Division of Biological Sciences

27 University of Montana

28 Missoula, MT 59812

29 Email: emily.kopania@umconnect.umt.edu

30 Jeffrey M. Good

31 Division of Biological Sciences

32 University of Montana

33 Missoula, MT 59812

34 Phone: (406) 243 - 5122

35 Email: jeffrey.good@umontana.edu

36

37 **Abstract**

38 Incompatibilities on the sex chromosomes are important in the evolution of hybrid male  
39 sterility, but the evolutionary forces underlying this phenomenon are unclear. House  
40 mice (*Mus musculus*) lineages have provided powerful models for understanding the  
41 genetic basis of hybrid male sterility. X chromosome-autosome interactions cause  
42 strong incompatibilities in *Mus musculus* F1 hybrids, but variation in sterility phenotypes  
43 suggests a more complex genetic basis. Additionally, X-Y chromosome conflict has  
44 resulted in rapid expansions of ampliconic genes with dosage-dependent expression  
45 that is essential to spermatogenesis. Here we evaluated the contribution of X-Y lineage  
46 mismatch to male fertility and stage-specific gene expression in hybrid mice. We  
47 performed backcrosses between two house mouse subspecies to generate reciprocal  
48 Y-introgression strains and used these strains to test the effects of X-Y mismatch in  
49 hybrids. Our transcriptome analyses of sorted spermatid cells revealed widespread  
50 overexpression of the X chromosome in sterile F1 hybrids independent of Y  
51 chromosome subspecies origin. Thus, postmeiotic overexpression of the X  
52 chromosome in sterile F1 mouse hybrids is likely a downstream consequence of  
53 disrupted meiotic X-inactivation rather than X-Y gene copy number imbalance. Y-  
54 chromosome introgression did result in subfertility phenotypes and disrupted expression  
55 of several autosomal genes in mice with an otherwise non-hybrid genomic background,  
56 suggesting that Y-linked incompatibilities contribute to reproductive barriers, but likely  
57 not as a direct consequence of X-Y conflict. Collectively, these findings suggest that  
58 rapid sex chromosome gene family evolution driven by genomic conflict has not resulted  
59 in strong male reproductive barriers between these subspecies of house mice.

60

## 61 Introduction

62

63 Sex chromosomes are often involved in the evolution of reproductive isolation between  
64 animal species (Coyne and Orr 1989; Turelli and Orr 2000; Presgraves and Meiklejohn  
65 2021), with hybrid sterility or inviability arising more often in the heterogametic sex (i.e.,  
66 Haldane's Rule, Haldane 1922; Coyne and Orr 2004). Hybrid incompatibilities also tend  
67 to accumulate more rapidly on the X chromosome (Masly and Presgraves 2007), which  
68 is referred to as the large X-effect (Coyne and Orr 1989). Known as the two rules of  
69 speciation (Coyne and Orr 1989; Coyne and Orr 2004), these patterns have been  
70 supported across diverse taxa (Good *et al.* 2008a; Davis *et al.* 2015; Bi *et al.* 2019;  
71 Matute and Cooper 2021; Presgraves and Meiklejohn 2021) and undoubtedly drive the  
72 early stages of intrinsic reproductive isolation in many systems. Both Haldane's rule and  
73 the large-X effect appear particularly strong when considering hybrid male sterility in XY  
74 systems, suggesting an important role for X chromosome evolution in both speciation  
75 and the evolution of spermatogenesis. However, it remains unclear to what extent these  
76 general patterns reflect common evolutionary processes, functional mechanisms unique  
77 to sex chromosomes, or a mixture of both (Meiklejohn and Tao 2010).

78 Intrinsic reproductive barriers between nascent species often arise as an indirect  
79 consequence of rapid evolution within populations (Dobzhansky 1937; Coyne and Orr  
80 2004; Coughlan and Matute 2020), so the outsized contribution of sex chromosomes to  
81 male sterility may be an inevitable consequence of rapid molecular evolution on the X  
82 and Y chromosomes. For example, recurrent genomic conflict is thought to be rampant  
83 on the X and Y chromosomes because selfish genetic elements are more likely to arise  
84 on sex chromosomes (i.e., meiotic drive *sensu lato*; Frank 1991; Hurst and  
85 Pomiankowski 1991; Meiklejohn and Tao 2010; Lindholm *et al.* 2016). Hemizyosity of  
86 the X chromosome is also expected to promote more rapid adaptive molecular evolution  
87 relative to the autosomes across a broad range of conditions (i.e., the faster-X effect;  
88 Charlesworth *et al.* 1987; Vicoso and Charlesworth 2009). Note that hemizyosity on  
89 the X and Y chromosomes will also result in differential exposure of hybrid  
90 incompatibilities on the sex chromosomes in males if incompatibilities tend to be at least  
91 partially recessive (Turelli and Orr 1995; Turelli and Orr 2000). However, progress on

92 understanding how often these diverse evolutionary processes contribute to the  
93 evolution of hybrid male sterility has been hampered by a lack of data on the genetic  
94 underpinnings of reproductive isolation.

95 From a mechanistic perspective, the X and Y chromosomes are also subject to  
96 unique regulatory processes during mammalian spermatogenesis that are critical for  
97 normal male fertility and shape patterns of molecular evolution (Larson *et al.* 2018a).  
98 Both the X and Y chromosomes are packaged into condensed chromatin early in  
99 meiosis, resulting in transcriptional silencing of most sex-linked genes known as meiotic  
100 sex chromosome inactivation (MSCI; McKee and Handel 1993). Repressive chromatin  
101 persists through the postmeiotic stages (Namekawa *et al.* 2006), although many  
102 essential X- and Y-linked genes are highly expressed in postmeiotic, haploid round  
103 spermatids (Mueller *et al.* 2008; Sin and Namekawa 2013). Failure to broadly repress X-  
104 linked expression during these critical meiotic and postmeiotic stages can trigger  
105 spermatogenic disruption, reduced sperm production, and abnormal sperm morphology  
106 (Burgoyne *et al.* 2009; Turner 2015). Interestingly, sex chromosome repression during  
107 both stages appears prone to disruption in hybrid mammals (Mihola *et al.* 2009; Good *et al.*  
108 *al.* 2010; Campbell *et al.* 2013; Davis *et al.* 2015; Larson *et al.* 2017), which may reflect  
109 common regulatory pathways underlying the evolution of hybrid male sterility  
110 (Bhattacharyya *et al.* 2013; Larson *et al.* 2021). Understanding how these intermediate  
111 developmental sterility phenotypes relate to genomic conflict and the broader  
112 evolutionary dynamics of the sex chromosomes awaits more data.

113 House mice (*Mus musculus*) have emerged as predominant models for  
114 understanding both the basic molecular control of spermatogenesis and the evolution of  
115 hybrid male sterility in mammals (Phifer-Rixey and Nachman 2015). Closely related  
116 subspecies of mice, *Mus musculus musculus* and *M. m. domesticus* (hereafter,  
117 “*musculus*” and “*domesticus*”), readily hybridize in both the lab and along a natural  
118 hybrid zone in Europe (Janoušek *et al.* 2012). Hybrid male sterility is the strongest and  
119 likely primary reproductive barrier isolating these incipient species in nature  
120 (Vyskočilová, et al. 2005; Turner, et al. 2012) and in the lab (Good *et al.* 2008b;  
121 Vyskočilová *et al.* 2009) following Haldane’s rule (Haldane 1922; but see Suzuki and  
122 Nachman 2015). Male sterility is polymorphic with laboratory crosses yielding sterile,

123 subfertile, or fertile male hybrids depending on genotype and cross direction (Good *et al.*  
124 *et al.* 2008b; Balcova *et al.* 2016; Larson *et al.* 2018b; Widmayer *et al.* 2020); *musculus*<sup>♀</sup>  
125 × *domesticus*<sup>♂</sup> crosses usually result in sterile F1 males, while the reciprocal cross  
126 tends to be more fertile (Good *et al.* 2008b). This asymmetry is caused by epistatic  
127 incompatibilities that are exposed on the *musculus* X chromosome in hybrid males  
128 (Storchová *et al.* 2004; Good *et al.* 2008a; Turner and Harr 2014). House mice also  
129 remain the only mammalian system where the evolution of a specific gene, *Prdm9*, has  
130 been directly linked to the evolution of intrinsic reproductive barriers (Mihola *et al.* 2009;  
131 Bhattacharyya *et al.* 2013; Mukaj *et al.* 2020). *Prdm9* is an autosomal gene encoding a  
132 DNA-binding protein that directs double stranded breaks where meiotic recombination  
133 occurs (Grey *et al.* 2011). PRDM9 binding sites evolve rapidly (Oliver *et al.* 2009; Baker  
134 *et al.* 2015), leading to asymmetric binding in hybrid mice that triggers autosomal  
135 asynapsis and disruption of MSCI during early pachytene of Meiosis I (Mihola *et al.*  
136 2009; Davies *et al.* 2016). *Prdm9*-related sterility depends on *Prdm9* heterozygosity and  
137 epistatic interactions with other unlinked factors, including a major incompatibility locus,  
138 *Hstx2*, located near the middle the *musculus* X chromosome (Forejt *et al.* 2021). This  
139 same X-linked region also influences hybrid male sterility in backcrossed consomic  
140 models (i.e., presumably independent of *Prdm9*; Storchová *et al.* 2004; Good *et al.*  
141 2008a), and recombination rate variation between *M. m. musculus* and another  
142 subspecies, *M. m. castaneus* (Dumont and Payseur 2011).

143 This broad foundation on the genetics of hybrid male sterility provides an  
144 opportunity to further unravel the various evolutionary and mechanistic processes that  
145 contribute to the large X-effect in mice. *Prdm9*-related sterility plays a central role in the  
146 evolution of hybrid male sterility and the disruption of MSCI in F1 mouse hybrids (Forejt  
147 *et al.* 2021; Larson *et al.* 2021). However, X- and Y-linked hybrid sterility arises across a  
148 broader range of genetic architectures and phenotypes than can be easily ascribed to  
149 *Prdm9*-related interactions (Campbell *et al.* 2012; Campbell and Nachman 2014; Larson  
150 *et al.* 2018b; Larson *et al.* 2021). The mouse X and Y chromosomes also contain  
151 clusters of several high copy ampliconic genes (Mueller *et al.* 2008; Soh *et al.* 2014;  
152 Case *et al.* 2015; Morgan and Pardo-Manuel De Villena 2017; Larson *et al.* 2021) that  
153 appear to have evolved in response to intense intragenomic conflict (Cocquet *et al.*

154 2009; Ellis *et al.* 2011; Cocquet *et al.* 2012). These X- and Y-linked gene clusters are  
155 primarily expressed in postmeiotic cells with repressed sex chromatin (Namekawa *et al.*  
156 2006; Sin *et al.* 2012) and thus increases in copy number may help counteract  
157 repressive chromatin (Ellis *et al.* 2011; Mueller *et al.* 2013; Sin and Namekawa 2013).  
158 Conflict arises because the maintenance of repressive postmeiotic sex chromatin  
159 appears to be controlled by dosage dependent interactions between X-linked (*Slx* and  
160 *Slx1*) and Y-linked (*Sly*) gene families (Cocquet *et al.* 2012; Kruger *et al.* 2019).  
161 Experimental knockdowns of *Slx* and *Slx1* showed increased sex chromosome  
162 repression, abnormal sperm head morphology, and an excess of male offspring. In  
163 contrast, knockdowns of *Sly* showed sex chromosome overexpression, abnormal sperm  
164 head morphology, and an excess of female offspring (Cocquet *et al.* 2009; Cocquet *et*  
165 *al.* 2012) due to reduced motility of Y-bearing sperm (Rathje *et al.* 2019). CRISPR-  
166 based deletions have further shown that sex-ratio distortion is primarily mediated by  
167 *Slx1* versus *Sly* competition for the spindlin proteins (SPIN1, SSTY1/2; Kruger *et al.*  
168 2019).

169 Copy numbers of *Slx*, *Slx1*, and *Sly* genes have co-evolved in different mouse  
170 lineages (Ellis *et al.* 2011; Good 2012; Morgan and Pardo-Manuel De Villena 2017),  
171 such that hybrids could have copy number mismatch sufficient to generate dosage-  
172 based sterility phenotypes seen in genetic manipulation studies (Ellis *et al.* 2011). In  
173 support of this model, hybrid interactions between the *musculus* X and the *domesticus*  
174 Y have been shown to cause abnormal sperm head morphology (Campbell *et al.* 2012;  
175 Campbell and Nachman 2014), and male sterility is associated with extensive  
176 overexpression of the sex chromosomes in postmeiotic round spermatids in *musculus*<sup>♀</sup>  
177 × *domesticus*<sup>♂</sup> mice (Larson *et al.* 2017). These hybrids have proportionally higher  
178 numbers of *Slx* and *Slx1* relative to *Sly* copies compared to non-hybrids and show  
179 patterns qualitatively consistent with the overexpression phenotypes observed in *Sly*  
180 knockdown and *Slx/Slx1* duplication mice (Cocquet *et al.* 2012; Kruger *et al.* 2019).  
181 However, postmeiotic sex chromatin repression is thought to partially depend on  
182 repressive histone marks established during meiosis (Turner *et al.* 2006), and the same  
183 direction of the hybrid cross also shows disrupted MSCI in meiotic spermatocytes  
184 (Campbell *et al.* 2013; Larson *et al.* 2017). Thus, it remains unclear if the disruption of

185 repressive postmeiotic chromatin is a consequence of X-Y mismatch or primarily a  
186 downstream epigenetic effect of deleterious interactions between the *musculus* X  
187 chromosome and *Prdm9* during meiosis (Larson *et al.* 2021).

188 Here, we advance understanding of the basis of hybrid male sterility in this  
189 system using a reciprocal backcrossing scheme to generate mice with the Y  
190 chromosome of one *Mus musculus* subspecies on the genomic background of another  
191 (Figure 1A). We used these Y-consomic genetic models to perform two reciprocal cross  
192 experiments while controlling for the effects of inbreeding. First, we tested for the  
193 potential rescue of sterility phenotypes in hybrid males with F1 autosomal genotypes but  
194 with matching X and Y chromosomes from the same subspecies (Hybrid F1 XY Match;  
195 Figure 1B). This experiment allowed us to tease apart X-Y interactions (i.e., *Slx* and  
196 *Slx1* versus *Sly*) from X-autosomal interactions (i.e., *Prdm9*-related sterility). Second,  
197 we tested the effects of X-Y mismatch on different subspecific backgrounds (Non-hybrid  
198 XY Mismatch; Figure 1B). This experiment allowed us to test for incompatibilities  
199 exposed on introgressed Y chromosomes that occur independently of other hybrid  
200 interactions. We used genome sequencing to quantify X- and Y-linked gene copy  
201 numbers, quantified male reproductive phenotypes (testis weight and high-resolution  
202 sperm head morphology), and used Fluorescence-Activated Cell Sorting (FACS) to  
203 isolate cell populations enriched for either early meiotic leptotene-zygotene  
204 spermatocytes or postmeiotic round spermatids. We used these experiments to address  
205 three main questions: (i) Does X-Y mismatch cause abnormal male reproductive traits?  
206 (ii) Do differences in copy number predict differences in ampliconic gene family  
207 expression levels during late spermatogenesis? (iii) Is X-Y mismatch associated with  
208 disrupted gene expression during late spermatogenesis, particularly on the sex  
209 chromosomes?

210

## 211 **Materials and Methods**

212

### 213 **Mouse resources and experimental design**

214 We sought to test the effects of X-Y mismatch independent of the effects of X-  
215 autosomal incompatibilities and inbreeding. To do so, we conducted two experiments:

216 (1) a “Hybrid F1 XY Match” experiment to test if matching the subspecies origin of the X  
 217 and Y rescued expression and reproductive phenotypes on an otherwise F1 hybrid  
 218 autosomal background, and (2) a “Non-hybrid XY Mismatch” experiment to test if  
 219 introgressed X-Y subspecies origin mismatch disrupted expression and reproductive  
 220 phenotypes on a non-hybrid autosomal background. To breed mice for these  
 221 experiments, we first generated reciprocal consomic introgression strains with the Y  
 222 chromosome from one subspecies on the genetic background of the other by  
 223 backcrossing *musculus* (PWK) and *domesticus* (LEWES) for 10 generations, which we  
 224 refer to as *musculus*<sup>domY</sup> and *domesticus*<sup>musY</sup> (Figure 1A). We tested to ensure our Y-  
 225 introgression strains had copy number mismatch representative of that expected in  
 226 natural hybrids. We used publicly available whole genome sequence data to estimate  
 227 copy number in wild house mice (PRJEB9450 for *domesticus*, 14 males and 1 female,  
 228 Pezer *et al.* 2015; PRJEB11742 for *musculus*, 5 males and 11 females, Harr *et al.*  
 229 2016) and wild-derived inbred laboratory mouse strains representing *musculus*  
 230 (PWK/PhJ and CZECHII/EiJ) and *domesticus* (LEWES/EiJ and WSB/EiJ;  
 231 PRJNA732719; one male individual per strain; Larson *et al.* 2021). We then used these  
 232 Y-introgression strains to perform two experiments and test the effects of X-Y mismatch  
 233 on hybrid sterility independent of X-autosomal incompatibilities (Figure 1B).

234  
 235 **Experiment 1, Hybrid F1 XY Match:** To test the effects of X-autosomal F1  
 236 incompatibilities without the effect of sex chromosome mismatch, we crossed Y-  
 237 introgression males to females with the same autosomal and X chromosome  
 238 type as the male Y chromosome (LEWES or PWK). This generated mice with an  
 239 F1 hybrid autosomal background and X-autosomal mismatch but X and Y  
 240 chromosomes from the same subspecies. Throughout the text, we refer to these  
 241 mice as *mus*×*dom*<sup>musY</sup> and *dom*×*mus*<sup>domY</sup>. We compared these mice to standard  
 242 F1 hybrid mice with the same X chromosome and autosomal background but no  
 243 Y chromosome introgression (PWK<sup>♀</sup> × LEWES<sup>♂</sup>, hereafter “*mus*×*dom*” and  
 244 LEWES<sup>♀</sup> × PWK<sup>♂</sup>, hereafter “*dom*×*mus*”).

245

246 **Experiment 2, Non-hybrid XY Mismatch:** To test the effects of X-Y mismatch  
 247 while controlling for inbreeding effects, we crossed Y-introgression males to  
 248 females from the same subspecies but a different strain from the genomic  
 249 background of the Y-introgression strain (CZECHII or WSB). This generated  
 250 mice with a non-hybrid (intrasubspecific) F1 autosomal background and  
 251 mismatched sex chromosomes (i.e., no X-autosomal mismatch), which we will  
 252 refer to as  $mus^{domY}$  and  $dom^{musY}$ . We compared these to intrasubspecific F1 mice  
 253 with the same autosomal background as these F1 Y-introgression mice, but  
 254 without sex chromosome mismatch (CZECHII<sup>♀</sup> × PWK<sup>♂</sup>, hereafter “*mus*” and  
 255 WSB<sup>♀</sup> × LEWES<sup>♂</sup>, hereafter “*dom*”). Note that these Non-hybrid XY Mismatch  
 256 mice had X chromosomes from different laboratory strains than the Hybrid F1 XY  
 257 Match mice of the same subspecies as a necessary consequence of breeding  
 258 mice with a heterozygous F1 background.

259  
 260 All mice from wild-derived inbred strains, Y-introgression strains, and experimental  
 261 crosses were maintained in breeding colonies at the University of Montana (UM)  
 262 Department of Laboratory Animal Resources (IACUC protocols 002-13, 050-15, and  
 263 062-18), which were initially purchased from The Jackson Laboratory, Bar Harbor, ME  
 264 in 2010. Replacement stock of LEWES/EiJ mice were ordered in 2013, and these mice  
 265 were used for the backcrosses to generate the  $dom^{musY}$  Y-introgression strains, as  
 266 dams in the *dom* intrasubspecific F1s, and as sires in the *dom* × *mus* and *dom* × *mus*<sup>*domY*</sup>  
 267 crosses.

## 269 **Whole genome sequencing and copy number estimation**

270 We sequenced whole genomes from one male mouse of each Y-introgression strain to  
 271 estimate ampliconic gene family copy numbers. We extracted DNA from mouse liver  
 272 using a Qiagen DNeasy kit and sent samples to Novogene (Novogene Corporation Inc.,  
 273 Sacramento, California) for library preparation and sequencing using Illumina HiSeq  
 274 paired-end 150bp. Libraries were prepared and sequenced twice to increase unique  
 275 coverage. We trimmed raw reads with Trimmomatic version 0.39 (Bolger *et al.* 2014).  
 276 We mapped reads to the mouse reference genome build GRCm38 using bwa mem

277 version 0.7.17 (Li and Durbin 2009) and used picard version 2.18.29 to fix mates  
278 and mark duplicates (Picard Toolkit). Data from the two sequencing runs were then  
279 merged for each sample.

280 To identify paralogs of ampliconic gene families, we extracted known X (*Slx*,  
281 *Slx11*, *Sstx*), Y (*Sly*, *Ssty1*, *Ssty2*), and autosomal (*Speer*, and  $\alpha$ -*takusan*) ampliconic  
282 gene sequences from the mouse reference GRCm38 using Ensembl annotation  
283 version 102 (Yates *et al.* 2019). We used the predicted gene *Gm5926* for *Sstx*  
284 because *Sstx* was not annotated in this version of Ensembl. For the autosomal gene  
285 families, we used the longest annotated genes in the gene family ( $\alpha$ 7-*takusan* and  
286 *Speer4f2*). We performed Ensembl BLAT searches with these sequences against  
287 the GRCm38 mouse reference, allowing up to 1000 hits. We then extracted all BLAT  
288 hits with greater than or equal to 97% sequence identity and an e-value of 0.0 and  
289 considered these filtered BLAT hits to be gene family paralogs for downstream copy  
290 number estimation.

291 We estimated copy numbers using a relative coverage approach similar to  
292 (Morgan and Pardo-Manuel De Villena 2017) and AmpliCoNE (Vegesna *et al.* 2020).  
293 For the relative coverage approach, we used Mosdepth v0.3.2 (Pedersen and  
294 Quinlan 2017) to estimate coverage across paralogous regions and divided this sum  
295 by half the genome-wide average coverage to account for hemizyosity of the sex  
296 chromosomes in males.

297 AmpliCoNE also estimates copy number based on relative coverage, while  
298 also controlling for GC content and only using informative regions based on repeat  
299 masking and mappability. AmpliCoNE was developed for estimating copy number on  
300 the assembly and annotation of the human Y, so we made some modifications to  
301 allow AmpliCoNE to work with the mouse sex chromosomes (Larson *et al.* 2021);  
302 <https://github.com/ekopania/modified-AmpliCoNE>). Specifically, we replaced  
303 AmpliCoNE's method for identifying informative sites with an approach more suitable  
304 for the mouse assembly. For each ampliconic gene family, we extracted all k-mers  
305 of length 101bp from the sequence of one gene representing the ampliconic family  
306 and mapped these back to the mouse reference genome using Bowtie2 and

307 allowing up to 500 multiple mapping hits. For each gene, we identified the most  
308 frequent number of times ( $m$ ) k-mers mapped to the mouse genome and kept only k-  
309 mers that mapped  $m$  times. We identified all locations where these k-mers mapped  
310 with 2 or fewer mismatches. We considered the start locations of these k-mer  
311 mapping hits to be “informative sites.”

312 A small amount of autosomal material (~0.1%) is expected to have introgressed  
313 along with the Y chromosome in our backcross experiments. To test this theoretical  
314 expectation and identify regions of introgression, we mapped whole genome sequence  
315 data from Y-introgression strains to both parental genomes using bwa mem v0.7.17-  
316 r1188 (Li and Durbin 2009) and called variants with GATK HaplotypeCaller v4.2.2.0. We  
317 then counted the number of variants in 100kb windows across the autosomes and  
318 identified regions where the number of variants when mapped to the maternal parent  
319 (autosomal background) genome exceeded the number of variants when mapped to the  
320 paternal parent (Y-introgression) genome. We repeated this analysis using whole  
321 genome sequence data from PWK and LEWES samples in our mouse colony. We  
322 excluded regions that had more variants when mapped to the opposite strain than when  
323 mapped to the same strain, as these are likely regions where genotype calls are  
324 unreliable due to assembly issues. After excluding these regions, 100kb windows with  
325 at least two more variants when mapped to the maternal parent compared to the  
326 paternal parent were considered introgressed in Y-introgression strains, reflecting the  
327 95<sup>th</sup> percentile of differences in the number of variants within a window.

328

### 329 **Reproductive phenotypes**

330 We phenotyped unmated male mice that were weaned at 21 days post-partum (dpp)  
331 into same-sex sibling groups and housed individually starting at 45 dpp to minimize  
332 effects of social dominance. Phenotypes were collected from at least six individuals for  
333 each cross type; sample sizes for each phenotype and cross type are in Table 1. We  
334 weighed paired testes and paired seminal vesicles and calculated their mass relative to  
335 body weight. We compared offspring sex ratios from Y-introgression mice by recording  
336 the number of offspring of each sex at weaning. We then tested for a significant  
337 difference from an even sex ratio using a Pearson’s chi-squared test in R, and did a

338 power analysis for this chi-squared test using the *pwr.chisq.test* function in the *pwr*  
339 package in R.

340 To quantify sperm morphology, we extracted sperm from each cross type from  
341 cauda epididymides diced in 1mL Dulbecco's PBS (Sigma) and incubated at 37°C for 10  
342 minutes. Sperm were fixed in 2% PFA, then dropped onto a slide with DAPI solution to  
343 stain the sperm nuclei. We imaged greater than 400 nuclei per genotype and analyzed  
344 the images using the Nuclear Morphology Analysis software (Skinner *et al.* 2019). We  
345 used two microscopes but performed clustering analysis on combined nuclei imaged  
346 from both microscopes to ensure that nuclei imaged on one scope were not clustering  
347 separately from those taken on the other microscope (Supplemental Material, Figure  
348 S1). The Nuclear Morphology Analysis software uses a Canny edge detection algorithm  
349 to detect objects (nuclei) within images, orients and aligns the nuclei, and uses a  
350 modification of the Zahn-Roskies transformation of the nucleus outlines to automatically  
351 detect landmarks. The software estimates area, perimeter, bounding height, bounding  
352 width, regularity, difference from median, and a consensus shape of the nuclei for each  
353 genotype. We tested for significant differences among cross types for each of these  
354 parameters using a Wilcoxon rank sum test in R. Using this automated morphology  
355 analysis software, we were able to analyze 5652 nuclei and detect subtle but significant  
356 differences that may not be measurable by eye or qualitative analysis.

357

### 358 **Testis sorting and RNA sequencing**

359 We collected testes from mice immediately following euthanization and isolated cells at  
360 different stages of spermatogenesis using Fluorescence-Activated Cell Sorting (FACS;  
361 Getun *et al.* 2011). The full FACS protocol is available on GitHub  
362 (<https://github.com/goodest-goodlab/good-protocols/tree/main/protocols/FACS>). Briefly,  
363 we decapsulated testes and washed them twice with 1mg/mL collagenase (Worthington  
364 Biochemical), 0.004mg/mL DNase I (Qiagen), and GBSS (Sigma), followed by  
365 disassociation with 1mg/mL trypsin (Worthington Biochemical) and 0.004mg/mL DNase  
366 I. We then inactivated trypsin with 0.16mg/mL fetal calf serum (Sigma). For each wash  
367 and disassociation step, we incubated and agitated samples at 33°C for 15 minutes on  
368 a SciGene Model 700 Microarray Oven at approximately 10rpm. We stained cells with

369 0.36mg/mL Hoechst 33324 (Invitrogen) and 0.002mg/mL propidium iodide and filtered  
370 with a 40µm cell filter. For Hybrid F1 XY Match, we sorted using a FACSAria Fusion  
371 flow cytometer, and for Non-hybrid XY Mismatch we sorted cells using a FACSAria Ilu  
372 cell sorter (BD Biosciences), both at the UM Center for Environmental Health Sciences  
373 Fluorescence Cytometry Core. We periodically added 0.004mg/mL DNase I as needed  
374 during sorting to prevent DNA clumps from clogging the sorter. We sorted cells into  
375 15µL beta-mercaptoethanol (Sigma) per 1mL of RLT lysis buffer (Qiagen) and kept  
376 samples on ice whenever they were not in the incubator or the cell sorter. We  
377 performed cell sorting on four individuals of each cross type and focused on two cell  
378 populations: early meiotic spermatocytes (leptotene/zygotene) and postmeiotic round  
379 spermatids. We extracted RNA using the Qiagen RNeasy Blood and Tissue Kit and  
380 checked RNA integrity with a TapeStation 2200 (Agilent). Only two samples had RNA  
381 integrity numbers (RIN) less than 8 (RIN = 7 and 7.1; Supplemental Material, Table S1).  
382 We prepared RNAseq libraries using the KAPA mRNA hyperprep kit and sequenced  
383 samples with Novogene (Illumina NovaSeq6000 PE 150). Samples were prepared and  
384 sequenced together, but Hybrid F1 XY Match mice and Non-hybrid XY Mismatch mice  
385 were sorted on different FACS machines, so to minimize experimental batch effects we  
386 analyzed these two experiments separately unless otherwise noted.

387

### 388 **Gene expression analyses**

389 We performed gene expression analyses on FACS expression data representing two  
390 cell populations: early meiosis (leptotene-zygotene, hereafter “early”) and postmeiosis  
391 (round spermatids, hereafter “late”). For the early cell type, a few samples did not group  
392 with others of the same cross type in multidimensional scaling (MDS) plots  
393 (Supplemental Material, Figure S2). These samples were likely contaminated with other  
394 cell types based on their relative expression levels of cell-type marker genes from *Mus*  
395 *musculus* testes single-cell RNAseq experiments (Supplemental Material, Figure S3;  
396 Green *et al.* 2018; Hunnicutt *et al.* 2021), and were therefore removed from expression  
397 analyses. Because sex chromosome ampliconic genes are primarily expressed in late  
398 spermatogenesis (Mueller *et al.* 2013; Larson *et al.* 2018a), and disrupted sex  
399 chromosome expression in hybrid males primarily occurs after the early cell type stage

400 (Larson *et al.* 2017), we focus on data from the late cell type in the main text and report  
401 results from the early cell type in the Supplemental Material.

402 We performed gene expression analyses using mice from both our Hybrid F1 XY  
403 Match and Non-hybrid XY Mismatch experiments, and reanalyzed expression data from  
404 (Larson *et al.* 2017), which generated spermatogenesis cell-type enriched gene  
405 expression data from the same F1 hybrid crosses (PWK<sup>♀</sup> × LEWES<sup>♂</sup> and LEWES<sup>♀</sup> ×  
406 PWK<sup>♂</sup>) and intrasubspecific F1 crosses (CZECHII<sup>♀</sup> × PWK<sup>♂</sup> and WSB<sup>♀</sup> × LEWES<sup>♂</sup>)  
407 used in this study.

408 We trimmed RNAseq reads using trimmomatic v0.39 (Bolger *et al.* 2014). One  
409 sample (PP.LL30.7MLZ) had about an order of magnitude more reads than any other  
410 sample (> 900 million raw reads), so we downsampled to the mean number of reads  
411 after trimming using fastq-sample version 0.8.3 and verified that reads were properly  
412 paired after downsampling using fastq\_pair (Edwards and Edwards 2019). We  
413 quantified reads using a kmer-based quasi-mapping approach implemented in salmon  
414 v1.4.0 (Patro *et al.* 2017) and a salmon index based on the mouse reference  
415 transcriptome version GRCm38. We then converted from transcript-level counts to  
416 gene-level counts using the R packages tximport 1.14.2 and EnsDb.Mmusculus.v79.  
417 We used EdgeR version 3.32.1 to normalize expression data. First, we filtered out  
418 genes with low expression by only including genes that had an FPKM > 1 in at least 4  
419 samples. Then, we normalized expression data following the recommendations in the  
420 tximport documentation.

421 We quantified expression levels of ampliconic gene families by calculating  
422 transcripts per million (TPM) for each gene separately then summing TPM values for all  
423 paralogs of a gene family (≥97% sequence identity). We used linear mixed-effect  
424 models to test if gene family expression level was significantly associated with copy  
425 number for *Slx*, *Slxl1*, *Sly*, *Ssty1*, *Ssty2*, and *α-takusan*. We compared disrupted  
426 expression levels on the autosomes, X chromosome, and Y chromosome by subtracting  
427 normalized FPKM values in control mice from normalized FPKM values in X-Y  
428 mismatch mice and control mice for every gene (Good *et al.* 2010). We then used a  
429 Mann-Whitney U test to compare the distribution of normalized FPKM differences  
430 among the chromosome types. To identify Differentially Expressed (DE) genes between

431 cross types, we used the likelihood ratio test approach with false-discovery rate (FDR)  
432 correction in EdgeR and visualized overlaps in DE genes among cross types using the  
433 R package UpSetR (Conway *et al.* 2017). We removed DE genes in autosomal regions  
434 we identified as putatively introgressed, because these genes may be DE due to  
435 introgressed autosomal variants rather than incompatibilities resulting from mismatching  
436 sex chromosomes. For ampliconic genes with high sequence similarity, some reads are  
437 expected to map multiply but will only be assigned to one member of the ampliconic  
438 gene family. Therefore, individual genes within gene families may sometimes be  
439 identified as DE, even though their paralogs are not, due to differences in read  
440 assignment across paralogs.

441 We further investigated genome-wide expression differences among cross types  
442 using weighted correlation network analyses (WGCNA; Langfelder and Horvath 2008).  
443 We identified correlated expression modules significantly associated with different cross  
444 types using a linear model and Tukey's honest significant difference (HSD) test. We  
445 used R version 4.0.3 for all statistical tests and to implement all R packages (R Core  
446 Team).

447

## 448 **Results**

449

### 450 **Copy Number Imbalance in Y-introgression Mice**

451 We first estimated ampliconic gene family copy numbers in wild mice, wild-derived  
452 inbred strains, and Y-introgression mice using whole genome sequencing. The samples  
453 that we sequenced had genome-wide average coverages of 10-15×, and samples with  
454 publicly available data all had coverage >5×. We found that *musculus* tended to have  
455 higher *Slx* and *Sly* copy numbers than *domesticus* (median *Slx* copy number in  
456 *musculus*: 62, in *domesticus*: 17, FDR-corrected Wilcoxon rank sum  $P < 0.01$ ; median  
457 *Sly* copy number in *musculus*: 226, in *domesticus*: 109, FDR-corrected Wilcoxon rank  
458 sum  $P < 0.01$ ), qualitatively consistent with previous studies (Ellis *et al.* 2011; Case *et*  
459 *al.* 2015; Morgan and Pardo-Manuel De Villena 2017; Figure 2A). *Slx11* copy numbers  
460 also tended to be higher in *musculus*, but there was high copy number variation for this  
461 gene family in *domesticus* with some samples reaching copy numbers as high as those

462 found in *musculus* (median *Slx1* copy number in *musculus*: 37, in *domesticus*: 31, FDR-  
463 corrected Wilcoxon rank sum  $P < 0.01$ ; Figure 2B). *Slx*, *Slx1*, and *Sly* copy numbers for  
464 wild-derived inbred strains were representative of those found in wild mice (Figures 2A  
465 and 2B; Supplemental Material, Table S3), consistent with previous results (Larson *et*  
466 *al.* 2021). Our Y-introgression mice retained copy numbers similar to those of pure  
467 strains with the same X and Y chromosome genotypes, so they had *Slx-Sly* and *Slx1-*  
468 *Sly* dosage imbalance similar to that expected in natural hybrids (Figures 2A and 2B;  
469 Supplemental Material, Table S3).

470 Additional ampliconic gene families showed copy number differences between  
471 *musculus* and *domesticus* that were also represented in our Y-introgression mice. *Sstx*  
472 had similar copy numbers in *musculus* and *domesticus*, but its two Y-linked homologs  
473 showed differences between subspecies, with *Ssty1* having more copies in *domesticus*  
474 and *Ssty2* having more copies in *musculus* (median *Sstx* copy number in *musculus*: 48,  
475 in *domesticus*: 39, FDR-corrected Wilcoxon rank sum  $P = 0.57$ ; median *Ssty1* copy  
476 number in *musculus*: 74, in *domesticus*: 139, FDR-corrected Wilcoxon rank sum  $P <$   
477  $0.01$ ; median *Ssty2* copy number in *musculus*: 145, in *domesticus*: 92, FDR-corrected  
478 Wilcoxon rank sum  $P < 0.01$ ; Figure 2C, 2D).

479 We also estimated copy number for  *$\alpha$ -takusan* and *Speer*, two autosomal  
480 ampliconic gene families thought to be regulated by sex chromosome ampliconic genes  
481 (Moretti *et al.* 2020). In both males and females,  *$\alpha$ -takusan* showed a high correlation in  
482 copy number with *Slx* ( $r = 0.95$ ; Pearson's correlation  $P < 0.001$ ), suggesting that it was  
483 co-amplified with the *Slx* gene family (Figure 2E). Note that correlation tests were  
484 performed without phylogenetic correction, because we wanted to test if gene families  
485 were co-amplified regardless of whether this was a result of shared evolutionary history.  
486 *Speer* copy number was more difficult to estimate using our approaches due to lower  
487 sequence similarity among *Speer* paralogs compared to other ampliconic gene families,  
488 but our estimates suggested that *Speer* may also have higher copy number in *musculus*  
489 relative to *domesticus* (Supplemental Material, Table S3). To verify our computational  
490 copy number estimates, we also performed digital droplet PCR (ddPCR) on a subset of  
491 *dom* samples using the *Slx1* primers from (Kruger *et al.* 2019). We found 15 *Slx1*  
492 copies with ddPCR, consistent with findings in (Kruger *et al.* 2019). While our

493 computational estimates are higher than this, we found similar results if we imposed a  
494 stricter cutoff for considering genes paralogs (98-99% sequence identity), likely  
495 reflecting a high specificity of the primers we used. We also found similar results using a  
496 different computational approach based on relative coverage (Supplemental Material,  
497 Table S3; Larson *et al.* 2021).

498

### 499 **Residual Autosomal Introgression in Y-introgression Strains**

500 We identified putative introgressed regions by mapping samples to both subspecies  
501 reference genomes, dividing the reference genome autosomal regions into 24,639  
502 100kb windows, and identifying SNPs in these windows. We found evidence for  
503 introgression in 105 windows in *domesticus*<sup>mus<sup>Y</sup></sup>, and 33 windows in *musculus*<sup>dom<sup>Y</sup></sup>,  
504 representing 0.43% and 0.13% of the autosomal windows that passed filtering,  
505 respectively (Supplemental Material, Table S4). Thus, the *domesticus*<sup>mus<sup>Y</sup></sup> strain had  
506 approximately four times more introgression than the theoretical expectation of 0.1%  
507 based on the number of backcross generations. The relatively large difference in  
508 percentages of introgression between the strains was primarily due to an ~7.6 Mbp  
509 introgressed region on chromosome 2 in *domesticus*<sup>mus<sup>Y</sup></sup> (Supplemental Material, Figure  
510 S4). This large introgressed region had an average difference of 958 SNPs, in contrast  
511 to the median difference of eight SNPs across all other putatively introgressed  
512 autosomal regions. Thus, the introgressed region on chromosome 2 in the  
513 *domesticus*<sup>mus<sup>Y</sup></sup> strain likely represents the only large track of autosomal introgression,  
514 with some evidence for additional, smaller amounts of introgression throughout the  
515 autosomes in both reciprocal Y-introgression strains.

516 Some of the putatively introgressed regions we identified may be prone to  
517 introgression more generally. The large area on chromosome 2 overlapped with a  
518 region with evidence for introgression from *musculus* into the *domesticus* wild-derived  
519 inbred strains STRA and STRB (Mukaj *et al.* 2020). We used the Mouse Phylogeny  
520 Viewer (Yang *et al.* 2011) to identify an additional nine mouse inbred strains with  
521 introgression from *musculus* into a *domesticus* background in this region (Supplemental  
522 Material, Figure S4C). In one area of the mouse hybrid zone, a SNP contained within  
523 this introgressed region showed evidence for excess of the *musculus* allele in mice with

524 primarily *domesticus* backgrounds, suggesting that introgression of this region from  
525 *musculus* into *domesticus* may have occurred in wild populations (Teeter *et al.* 2010).  
526 This region is also adjacent to *R2d2*, a copy number variant in mice that shows  
527 transmission ratio distortion in females heterozygous for the high copy number *R2d2*  
528 drive allele (Didion *et al.* 2016). We also identified 5 different 100kb windows near each  
529 other on chromosome 14 with evidence for introgression in *musculus*<sup>domY</sup> mice that  
530 overlap with a region in the *musculus* wild-derived strain PWD with evidence for  
531 introgression from *domesticus* (41.3-41.4Mb, 41.8-41.9Mb, 42.2-42.3Mb, 42.3-43.4Mb,  
532 and 44.2-44.3Mb; Mukaj *et al.* 2020).

533

### 534 **X-Y Mismatch Contributed to Male Sterility Phenotypes**

535 We next asked if X-Y mismatch was associated with male sterility phenotypes (Table 1).  
536 For Hybrid F1 XY Match, where we compared hybrid mice both with and without sex  
537 chromosome mismatch, hybrids with a *musculus*<sup>♀</sup> × *domesticus*<sup>♂</sup> background had lower  
538 relative testes mass than hybrids with the reciprocal *domesticus*<sup>♀</sup> × *musculus*<sup>♂</sup>  
539 background regardless of whether they had X-Y mismatch or not (Figure 3A). These  
540 results were consistent with previous studies showing more severe hybrid sterility in the  
541 *musculus*<sup>♀</sup> × *domesticus*<sup>♂</sup> direction of this cross (Good *et al.* 2008b; Good *et al.* 2010;  
542 Campbell *et al.* 2012; Larson *et al.* 2017). Although *domesticus*<sup>♀</sup> × *musculus*<sup>♂</sup> showed  
543 much less severe sterility phenotypes than the reciprocal F1 hybrid, we still considered  
544 these mice to be potentially subfertile because their relative testes mass and sperm  
545 morphology parameters were significantly different from those of either pure *dom* or  
546 pure *mus* (Figure 3, Table 1), and even subtle reductions in fertility may be important in  
547 nature, where sperm competition is high for house mice (Dean *et al.* 2006). For Hybrid  
548 F1 XY Match mice, *dom* × *mus*<sup>domY</sup> mice had higher relative testis mass than *dom* × *mus*  
549 mice, suggesting that X-Y match partially rescued relative testes mass in some mice  
550 with a hybrid autosomal background (Figure 3A). In the reciprocal direction, however, X-  
551 Y match had no significant effect on relative testes mass (Figure 3A). For Non-hybrid  
552 XY Mismatch, we found that mice with X-Y mismatch had reduced relative testis mass  
553 compared to control mice with the same non-hybrid X and autosomal background  
554 (Figure 3A). In summary, we found little effect of X-Y mismatch on testis mass in the

555 most sterile F1 cross (*musculus*<sup>♀</sup> × *domesticus*<sup>♂</sup>), where sterility is therefore likely due  
556 to X-autosomal or autosomal-autosomal incompatibilities (Campbell and Nachman  
557 2014). However, in the reciprocal and more fertile F1 direction X-Y mismatch seemed to  
558 have an important effect on testis mass. Furthermore, in the absence of any autosomal  
559 or X-autosomal incompatibilities, X-Y mismatch resulted in slightly but significantly  
560 decreased relative testis mass.

561 We saw severe sperm head abnormalities in our Hybrid F1 XY Match crosses  
562 with a *musculus*<sup>♀</sup> × *domesticus*<sup>♂</sup> background (*mus*×*dom* and *mus*×*dom*<sup>*mus*<sup>Y</sup></sup>). Sperm  
563 from both these cross types had significantly lower bounding height and bounding width  
564 compared to all other cross types (FDR-corrected Wilcoxon rank sum P << 0.0001;  
565 Table 1), largely due to their shortened hook and consistent with hybrid sterility in this  
566 direction of the cross (Figure 3B, 3C). This was also consistent with previous manual  
567 (categorical) observations of abnormal sperm head morphology in this cross type in  
568 other studies (Good *et al.* 2008a; Campbell and Nachman 2014; Larson *et al.* 2017;  
569 Larson *et al.* 2018b). The reciprocal *dom*×*mus* F1 hybrids had sperm with higher  
570 bounding height and bounding width compared to sperm from all other cross types,  
571 including the reference subspecies (FDR-corrected Wilcoxon rank sum P < 0.01; Table  
572 1; Figure 3B, 3C). This direction of the cross is generally considered more fertile but  
573 sometimes shows reduced fertility compared to non-hybrid mice (Larson *et al.* 2018b). It  
574 is possible that the larger overall size of these sperm may reflect abnormal nuclear  
575 packaging and could contribute to reduced fertility in *domesticus*<sup>♀</sup> × *musculus*<sup>♂</sup> F1 mice.  
576 When comparing X-Y match mice to F1 hybrids with abnormally small sperm heads,  
577 *mus*×*dom*<sup>*mus*<sup>Y</sup></sup> mice had significantly higher bounding width and bounding height than  
578 *mus*×*dom* mice (FDR-corrected Wilcoxon rank sum P < 0.01; Table 1; Figure 3B, 3C).  
579 These results suggest that X-Y match rescued some of the aberrant sperm head  
580 morphology associated with hybrid sterility in *musculus*<sup>♀</sup> × *domesticus*<sup>♂</sup> F1s, but the  
581 effects of X-Y match rescue were subtle, consistent with previous observations  
582 (Campbell and Nachman 2014). In the reciprocal cross direction, *dom*×*mus*<sup>*dom*<sup>Y</sup></sup> had  
583 lower bounding width and bounding height than the abnormally large *dom*×*mus* sperm  
584 heads (FDR-corrected Wilcoxon rank sum P << 0.0001; Table 1; Figure 3B, 3), so X-Y

585 match rescued some of the oversized sperm head morphology we observed in  
586 *dom*×*mus*.

587 In Non-hybrid XY Mismatch, we observed subtle effects of X-Y mismatch  
588 consistent with our Hybrid F1 XY Match observations. Sperm from *mus*<sup>*dom*Y</sup> mice had  
589 slightly lower bounding height and bounding width compared to sperm from *mus* (FDR-  
590 corrected Wilcoxon rank sum  $P < 0.01$ ; Table 1; Figure 3B, 3C), consistent with lower  
591 bounding height and bounding width in sperm from *mus*×*dom* mice that also had a *mus*  
592 X chromosome and *dom* Y chromosome. However, *mus*<sup>*dom*Y</sup> sperm were more similar in  
593 size to *mus* sperm than *mus*×*dom* sperm and qualitatively had a hook morphology more  
594 similar to that of fertile *mus* than sterile *mus*×*dom* mice, so the contribution of X-Y  
595 mismatch to sperm head morphology is small compared to the effect of X-autosomal  
596 interactions. In the reciprocal direction, *dom*<sup>*mus*Y</sup> mice had sperm with higher bounding  
597 height and bounding width compared to sperm from *dom* mice (FDR-corrected Wilcoxon  
598 rank sum  $P \ll 0.0001$ ; Table 1; Figure 3B, 3C), consistent with the higher bounding  
599 height and bounding width in *dom*×*mus* hybrids. Sperm from *dom*<sup>*mus*Y</sup> mice also had  
600 smaller areas (FDR-corrected Wilcoxon rank sum  $P \ll 0.0001$ ; Table 1; Supplemental  
601 Material, Figure S5), so the larger bounding height and bounding width are primarily the  
602 result of a slightly elongated hook rather than an overall increase in the sperm head  
603 size. Other sperm head morphology parameters, including area, perimeter, and  
604 differences from median, showed similar subtle differences or no differences among  
605 cross types (Table 1; Supplemental Material, Figures S1 and S5).

606 Genetic manipulation studies have shown offspring sex ratio skews under *Slx1*-  
607 *Sly* dosage imbalance, contributing to evidence for *Slx1*-*Sly* intragenomic conflict. Male  
608 mice with an excess of *Sly* relative to *Slx1* produce more male offspring, while mice  
609 with an excess of *Slx1* produce more female offspring (Cocquet *et al.* 2012; Kruger *et*  
610 *al.* 2019) due to reduced motility of Y-bearing sperm (Rathje *et al.* 2019). We asked if  
611 more subtle imbalances in relative copy numbers expected in natural hybrid mice also  
612 result in sex ratio skews and did not see a significant difference from a 50:50 sex ratio  
613 for offspring of X-Y mismatch mice (Supplemental Material, Table S5). A more extreme  
614 dosage imbalance than that seen in our X-Y mismatch experimental mice (and in  
615 natural hybrids) is probably required to produce a large sex ratio skew. However, it is

616 important to note that we had very little power to detect differences in sex ratio, with  
617 type II error probabilities over 0.8 (Supplemental Material, Table S5).

618

### 619 ***Slx*- and *Slx1*-*Sly* Dosage Imbalance Did Not Lead to Ampliconic Gene Family** 620 **Overexpression**

621 Copy number imbalance of *Slx* and *Slx1* relative to *Sly* is thought to disrupt expression  
622 of these gene families in late spermatogenesis, with particularly strong evidence for *Slx*  
623 and *Slx1* overexpression when *Sly* is knocked down (Cocquet *et al.* 2009; Cocquet *et*  
624 *al.* 2012) and *Slx1* overexpression when *Slx* and *Slx1* are duplicated (Kruger *et al.*  
625 2019). *Slx*, *Slx1*, and *Sly* appear to be involved in the regulation of sex chromatin which  
626 impacts the regulation of many genes during late spermatogenesis (Kruger *et al.* 2019).  
627 Therefore, we predicted that their misregulation may disrupt the expression of additional  
628 genes, including additional Y-linked ampliconic gene families *Ssty1/2* and the autosomal  
629 ampliconic gene family  $\alpha$ -*takusan* (Larson *et al.* 2017; Moretti *et al.* 2020). To test if *Slx*,  
630 *Slx1*, *Sly*, *Ssty1*, *Ssty2*, and  $\alpha$ -*takusan* expression was disrupted under less extreme  
631 copy number differences in hybrid mice, we compared ampliconic gene family  
632 expression levels in round spermatids among cross types. We did not directly quantify  
633 copy number for the mice that were FACS sorted, so we used our previous copy  
634 number estimates from pure strains sharing the same sex chromosomes as our  
635 experimental mice (Larson *et al.* 2021). For all six gene families, expression level was  
636 significantly associated with copy number based on a linear mixed-effects model with  
637 experiment as a random effect to control for batch effects (FDR-corrected  $P < 0.05$ ;  
638 Figure 4). However, for *Slx1*, this association was negative, suggesting that copy  
639 number was not the primary determinant of *Slx1* expression. This is interesting given  
640 that we found high overlap in the range of *Slx1* copy numbers in naturally occurring  
641 *musculus* and *domesticus* (Figure 2B), and the previous demonstration that *Slx1* plays  
642 a more direct role in sex ratio bias than *Slx* (Kruger *et al.* 2019). We then tested if X-Y  
643 mismatch had a significant effect on expression level using a linear mixed-effects model  
644 with both copy number and presence of X-Y mismatch as fixed effects and experiment  
645 as a random effect. We used an ANOVA to compare this model to a null model with  
646 copy number as the only fixed effect and experiment as a random effect. For all six

647 genes, X-Y mismatch was not significantly associated with ampliconic gene expression  
 648 levels (FDR-corrected ANOVA  $P > 0.05$ ). When we specified the direction of X-Y  
 649 mismatch (i.e., *musculus* X and *domesticus* Y, the direction with an excess of *Slx*  
 650 relative to *Sly*), only *Ssty2* expression was significantly associated with X-Y mismatch in  
 651 this direction (FDR-corrected ANOVA  $P > 0.05$ ).

652 We also tested if X-autosomal background was significantly associated with  
 653 expression levels using the same mixed-effects model approach. For *Slx*, *Slx11*, *Sly*,  
 654 *Ssty1*, and *Ssty2*, the sterile hybrid background (*musculus*<sup>♀</sup> × *domesticus*<sup>♂</sup>) was  
 655 significantly associated with expression levels after FDR-correction (*Slx* ANOVA  $P \ll$   
 656 0.0001; *Slx11*  $P < 0.001$ ; *Sly*  $P = 0.01$ ; *Ssty1*  $P < 0.001$ ; *Ssty2*  $P = 0.001$ ). We observed  
 657 overexpression of *Slx*, *Slx11*, *Sly*, *Ssty1*, and *Ssty2* relative to their copy numbers for  
 658 mice with *musculus*<sup>♀</sup> × *domesticus*<sup>♂</sup> backgrounds (*mus*×*dom* and *mus*×*dom*<sup>musY</sup>; Figure  
 659 4A-E), consistent with previous studies showing that these hybrid mice exhibit  
 660 widespread overexpression on the sex chromosomes (Good *et al.* 2010; Campbell *et al.*  
 661 2013; Larson *et al.* 2017). Both *mus*×*dom* and *mus*×*dom*<sup>musY</sup> mice in our study  
 662 overexpressed *Slx*, *Slx11*, and *Sly* (Figure 4A, 4B, and 4C), suggesting that matching X  
 663 and Y chromosomes from *musculus* did not rescue *Slx*, *Slx11*, or *Sly* upregulation, and  
 664 that the overexpression we observed likely results from X-autosomal incompatibilities  
 665 that disrupt MSCI rather than *Slx*- or *Slx11*-*Sly* dosage imbalance. Additionally, *mus*<sup>domY</sup>  
 666 mice from our Non-hybrid XY Mismatch also had a *musculus* X and *domesticus* Y, the  
 667 same X and Y chromosome combination found in sterile hybrids that results in an  
 668 excess of *Slx* and *Slx11* copies relative to *Sly* copies. If *Slx*- or *Slx11*-*Sly* dosage  
 669 imbalance contributed to *Slx*, *Slx11*, and *Sly* overexpression, we would expect *mus*<sup>domY</sup>  
 670 mice to have higher expression than *mus* controls. We observed the opposite effect,  
 671 with *mus*<sup>domY</sup> mice showing slightly lower *Slx*, *Slx11*, and *Sly* expression levels (Figure  
 672 4A, 4B, and 4C). This result provides further evidence that postmeiotic *Slx*, *Slx11*, and  
 673 *Sly* overexpression in sterile F1 hybrids is unlikely to be primarily due to *Slx*- or *Slx11*-*Sly*  
 674 dosage imbalance, and that X-Y mismatch in the absence of autosomal mismatch is not  
 675 sufficient to cause overexpression of *Slx*, *Slx11*, and *Sly*.

676 Given that *Slx*, *Slx11*, and *Sly* are thought to regulate the *α-takusan* ampliconic  
 677 family, we predicted that *α-takusan* expression levels would also be associated with a

678 *musculus*<sup>♀</sup> × *domesticus*<sup>♂</sup> background. Surprisingly, this association was not significant  
679 (ANOVA P = 0.40). Instead, we observed that *α-takusan* was overexpressed in all cross  
680 types with an F1 autosomal background regardless of cross direction (Figure 4F), and  
681 that expression was significantly associated with an F1 autosomal background (ANOVA  
682 P < 0.01). This suggests that *α-takusan* regulation likely involves autosomal loci in  
683 addition to SLX, SLXL1, SLY, SSTY1, and SSTY2 (Moretti *et al.* 2020).

684 Sex-linked ampliconic genes are primarily expressed during postmeiotic  
685 spermatogenesis, in mice and more generally across mammals (Cocquet *et al.* 2012;  
686 Mueller *et al.* 2013; Sin and Namekawa 2013). Our non-hybrid expression data  
687 supported this, with little to no expression of *Slx*, *Slxl1*, *Sly*, or *Ssty1/2* in early meiotic  
688 cells in our *mus* and *dom* samples. However, we did detect some meiotic expression of  
689 *Slx*, *Slxl1*, *Sly*, and *Ssty2* in mice with hybrid autosomal backgrounds, and expression  
690 levels of these gene families in early meiosis was significantly associated with F1  
691 autosomal background (ANOVA P < 0.05, Supplemental Material, Figure S6). X  
692 chromosome expression has been shown to be disrupted throughout spermatogenesis  
693 in F1 hybrids, although the effect was smaller during earlier spermatogenic stages  
694 (Larson *et al.* 2017). Our results suggest that disruption of early spermatogenesis  
695 regulatory networks may result in spurious expression of sex-linked ampliconic genes  
696 during early meiotic stages when they are normally silenced.

697

### 698 **X-Y Mismatch Was Not Associated with Sex Chromosome Overexpression in** 699 **Sterile F1 Hybrids**

700 Next we sought to differentiate if widespread postmeiotic overexpression in sterile  
701 hybrids was a direct result of sex chromosome mismatch, a continuation of disrupted  
702 meiotic sex chromosome inactivation (MSCI), or a combination of both (Larson *et al.*  
703 2017; Larson *et al.* 2021). We first reanalyzed data from (Larson *et al.* 2017) and  
704 repeated their result showing sex chromosome upregulation in late spermatogenesis in  
705 sterile F1 hybrids (*mus* × *dom*, Figure 5A and 5D). We then tested if upregulation was  
706 due to X-Y mismatch by comparing relative expression levels in F1 hybrids to those in  
707 our Hybrid F1 XY Match mice, which had sex chromosomes from the same subspecies.  
708 If X-Y mismatch contributed to sex chromosome upregulation in sterile hybrids, we

709 would expect to see some rescue from disrupted postmeiotic expression in these Hybrid  
 710 F1 XY Match mice, with *mus*×*dom*<sup>musY</sup> mice having lower expression on the X  
 711 chromosome relative to *mus*×*dom* F1s. Contrary to this prediction, the X chromosome  
 712 showed similar expression levels when comparing expression in these two cross types.  
 713 Therefore, restoring matching sex chromosomes did not rescue expression levels on  
 714 the *musculus* X chromosome from overexpression in hybrids (Figure 5B). We further  
 715 tested the effects of sex chromosome mismatch using our Non-hybrid XY Mismatch  
 716 mice, which had introgressed Y chromosomes on a non-hybrid autosomal background.  
 717 If mismatch between a *musculus* X chromosome and *domesticus* Y chromosome was  
 718 sufficient to induce postmeiotic sex chromosome overexpression, then we would expect  
 719 to see higher X chromosome expression in *mus*<sup>domY</sup> mice. Instead, we observed slight  
 720 under expression on the X chromosome compared to the autosomes in *mus*<sup>domY</sup> mice,  
 721 confirming that sex chromosome mismatch does not cause X chromosome  
 722 overexpression in late spermatogenesis (Figure 5C).

723 We also found evidence that sex chromosome mismatch does not contribute to Y  
 724 chromosome overexpression in late spermatogenesis in sterile *musculus*<sup>♀</sup> ×  
 725 *domesticus*<sup>♂</sup> hybrids. The Y chromosome was upregulated in *mus*×*dom* sterile hybrids  
 726 relative to *dom*×*mus*<sup>domY</sup> mice. This could be due to rescue of *domesticus* Y  
 727 chromosome expression when paired with the *domesticus* X, but it could also be due to  
 728 overall lower sex chromosome expression in mice with a *domesticus*<sup>♀</sup> × *musculus*<sup>♂</sup>  
 729 background (Figure 5E). In Non-hybrid XY Mismatch, we saw that *mus*<sup>domY</sup> mice had  
 730 lower expression on the Y chromosome compared to *dom* controls, in contrast to the Y  
 731 chromosome overexpression observed in *mus*×*dom* hybrids (Figure 5F). Thus, X-Y  
 732 mismatch does appear to influence Y chromosome expression, but in the opposite  
 733 direction of that observed in sterile hybrids.

734 In the reciprocal cross (*domesticus*<sup>♀</sup> × *musculus*<sup>♂</sup> F1 hybrids), we found some  
 735 evidence that X-Y mismatch may contribute to disrupted expression of X-linked genes.  
 736 Here Y chromosome expression was not different from that on the autosomes (Figure  
 737 5G), but the X chromosome tended to be downregulated (Figure 5J; Larson *et al.* 2017).  
 738 There was no evidence that X-Y match restored normal X chromosome expression  
 739 levels in *dom*×*mus*<sup>domY</sup> (Hybrid F1 XY Match), with this cross type showing similar or

740 even slightly lower expression levels on the X chromosome relative to *dom*×*mus* hybrids  
 741 (Figure 5K). However, in Non-hybrid XY Mismatch we observed lower expression on the  
 742 X chromosome in *dom*<sup>musY</sup> mice relative to *dom* controls (Figure 5L). Therefore, a  
 743 *domesticus* X paired with a *musculus* Y can result in suppression of X-linked gene  
 744 expression even in the absence of autosomal incompatibilities.

745

## 746 **X-Y Mismatch Disrupted the Expression of Several Genes during Late**

### 747 **Spermatogenesis**

748 We also tested for effects of X-Y mismatch on individual genes by identifying  
 749 differentially expressed (DE) genes in X-Y mismatch mice compared to controls. In our  
 750 reanalysis, we identified many more overexpressed genes in sterile *mus*×*dom* hybrids  
 751 compared to *mus* and many more underexpressed genes in the reciprocal *dom*×*mus*  
 752 hybrids compared to *dom* on the X chromosome (Table 2), consistent with previous  
 753 results (Larson *et al.* 2017) and with our observations of overall expression differences  
 754 (Figure 5). We then asked if any of these X-linked DE genes were associated with X-Y  
 755 mismatch. If so, then we would expect our Hybrid F1 XY Match *mus*×*dom*<sup>musY</sup> to rescue  
 756 some of the disrupted X-linked expression, and thus manifest as DE genes in  
 757 comparisons between *mus*×*dom* and *mus*×*dom*<sup>musY</sup>. These genes should also overlap  
 758 with genes DE between *mus*×*dom* and *mus*. However, there were only two X-linked DE  
 759 genes in the *mus*×*dom* versus *mus*×*dom*<sup>musY</sup> comparison (Table 2), and only one was  
 760 also DE in the *mus*×*dom* versus *mus* comparison (Figure 6). This gene is a predicted  
 761 protein coding gene, *Gm10058*, that shares 97% sequence identity with *Slx* and is  
 762 therefore likely a paralog of this gene family. The other DE gene was *Btbd35f17*,  
 763 another ampliconic gene with a protein-protein binding domain that is specifically  
 764 expressed in male reproductive tissues (Smith *et al.* 2019). In Non-hybrid XY Mismatch,  
 765 we only observed one X-linked DE gene in *mus*<sup>domY</sup> compared to *mus*, and this gene  
 766 was not DE in any other comparisons. Taken together, both Hybrid F1 XY Match and  
 767 Non-hybrid XY Mismatch results suggest that almost all DE genes on the X  
 768 chromosome in sterile *musculus*<sup>♀</sup> × *domesticus*<sup>♂</sup> hybrids are disrupted due to X-  
 769 autosomal or autosomal-autosomal incompatibilities, rather than Y-linked  
 770 incompatibilities.

771 On the X chromosome, very few DE genes were shared across multiple  
772 comparisons. However, 57 DE genes were shared between the *mus*×*dom* versus *mus*  
773 and *dom*×*mus* versus *dom* comparisons. When we looked at DE genes separated by  
774 direction of expression difference, only eight were shared between these two  
775 comparisons (Supplemental Material, Figure S7), so most of the overlap represented  
776 genes overexpressed in *mus*×*dom* but underexpressed in *dom*×*mus*. This could  
777 indicate that similar regulatory networks are disrupted in reciprocal F1 hybrids, but in  
778 ways that disrupt gene expression levels in opposite directions.

779 In contrast to the X chromosome, more Y-linked DE genes were shared across  
780 comparisons (Figure 6). Sterile *mus*×*dom* hybrids had 17 Y-linked DE genes that  
781 showed a clear bias towards overexpression (Table 2). Of these 17 DE genes, 5 were  
782 shared with the Hybrid F1 XY Match comparison *mus*×*dom* versus *dom*×*mus*<sup>domY</sup>, so  
783 having *domesticus* X and Y chromosomes partially rescued expression levels on the Y  
784 chromosome in *dom*×*mus*<sup>domY</sup> mice. However, none of the 17 Y-linked genes DE in  
785 sterile hybrids were also DE in the Hybrid F1 XY Match comparison (*mus*<sup>domY</sup> versus  
786 *dom*), so it is unlikely that X-Y mismatch alone disrupts expression of these genes.  
787 Instead, there may be a complex interaction between X-Y mismatch and a hybrid  
788 autosomal background that disrupts Y chromosome expression. Consistent with this, we  
789 found the most Y-linked DE genes in comparisons between cross types with reciprocal  
790 hybrid autosomal backgrounds but the same Y chromosome (Table 2). Of these, 78 Y-  
791 linked DE genes were shared between these two comparisons (Figure 6), suggesting  
792 that reciprocal hybrid autosomal backgrounds may have resulted in disrupted  
793 expression for many of the same Y-linked genes, regardless of the subspecies origin of  
794 the Y chromosome.

795 We also found several autosomal genes that were DE between cross types with  
796 the same autosomal background but different sex chromosome combinations (Table 2).  
797 We excluded autosomal genes that overlapped with putatively introgressed regions, so  
798 the DE that we detected was unlikely to result from *cis*-regulatory effects of variants  
799 from the opposite subspecies that introgressed along with the Y chromosome. In Hybrid  
800 F1 XY Match, 104 autosomal genes were DE when comparing *mus*×*dom* to  
801 *dom*×*mus*<sup>domY</sup> and 494 autosomal genes were DE when comparing *dom*×*mus* to

802 *mus*×*dom*<sup>*mus*Y</sup> (Table 2). These comparisons involved reciprocal crosses with the same  
803 autosomal and Y chromosome genotypes, and so DE presumably resulted from X-  
804 autosomal incompatibilities. Although overexpression on the X chromosome tends to be  
805 the most notable expression pattern associated with X-autosomal incompatibilities,  
806 previous studies have shown disrupted postmeiotic autosomal expression in sterile  
807 hybrids as well (Larson *et al.* 2017). We detected only six (non-overlapping) DE genes  
808 in each comparison with different Y chromosomes but the same autosomal and X  
809 chromosome genotypes (*mus*×*dom* versus *mus*×*dom*<sup>*mus*Y</sup> and *dom*×*mus* versus  
810 *dom*×*mus*<sup>*dom*Y</sup>; Table 2).

811 In Non-hybrid XY Mismatch, we identified some autosomal DE genes in  
812 comparisons that had different Y chromosomes but the same autosomal and X  
813 backgrounds, suggesting that interactions involving the Y chromosome disrupted some  
814 autosomal expression, but the number of autosomal DE genes was not enriched  
815 relative to the number of X-linked DE genes (Fisher's Exact Test  $P > 0.05$ ; Table  
816 2). These autosomal DE genes tended to be underexpressed in the cross type with X-Y  
817 mismatch regardless of the direction of the cross (Table 2) and must result from direct  
818 interactions with the Y chromosome or indirect interactions with X-Y mediated  
819 expression changes. Only one autosomal gene, *Babam2*, was DE in both reciprocal  
820 comparisons. It is a member of the BRCA1-A complex, which is involved in DNA  
821 double-strand break repair (The Uniprot Consortium 2020).

822 Finally, we tested if DE genes tended to be in the same co-expression networks  
823 using weighted correlation network analysis (WGCNA). We found one module in Hybrid  
824 F1 XY Match associated with the *mus*×*dom* autosomal background, one module in Non-  
825 hybrid XY Mismatch associated with the *musculus* background, and one module in Non-  
826 hybrid XY Mismatch associated with the *domesticus* background (Figure 7A, B, D).  
827 These modules were significantly enriched for genes DE between cross types with  
828 different autosomal backgrounds (Table 3). There were also multiple modules enriched  
829 for DE genes despite not having a significant association with cross type (Table 3). For  
830 example, Module 5 was significantly enriched for DE genes in all pairwise comparisons  
831 in Hybrid F1 XY Match. Although we did not detect a significant cross type association  
832 for this module, there was a trend towards an autosomal background by sex

833 chromosome effect for this module, with *mus*×*dom* background cross types tending to  
834 have lower module membership in general, but with *mus*×*dom*<sup>*mus*Y</sup> mice tending to have  
835 higher module membership than *mus*×*dom* mice (Figure 7E). Another Hybrid F1 XY  
836 Match module showed a similar pattern (Module 3, Figure 7C) and was enriched for  
837 genes DE between *dom*×*mus* and *mus*×*dom*<sup>*mus*Y</sup> (Table 3). In Non-hybrid XY Mismatch,  
838 Module 5 was enriched for genes DE between *mus*<sup>*dom*Y</sup> and either subspecies (*mus* or  
839 *dom*; Table 3), and X-Y mismatch mice tended to have lower associations with this  
840 module (Figure 7). We likely did not have enough power to detect significant module  
841 associations with complex autosome by sex chromosome interactions given our sample  
842 size, especially because these effects on gene expression tended to be subtle and  
843 affect relatively few genes (Figure 5, Table 2). Despite low power, the fact that certain  
844 modules were enriched for DE genes suggests that groups of genes were disrupted in  
845 similar ways in X-Y mismatch mice, and that particular gene networks may be disrupted  
846 under X-Y mismatch. Additionally, we found a significant positive correlation in module  
847 eigengene values between Hybrid F1 XY Match and Non-hybrid XY Mismatch (Module  
848 5 in both experiments,  $r = 0.64$ ; FDR-corrected Pearson's correlation  $P < 0.001$ ;  
849 Supplemental Figure S8) and a significant overlap in genes (279 genes, FDR-corrected  
850 Fisher's Exact Test  $P < 0.001$ ), suggesting that these two modules represent genes with  
851 similar expression patterns between the two experiments. Interestingly, these modules  
852 trended towards a negative association with cross types that had a *musculus* X  
853 chromosome and *domesticus* Y chromosome (Figure 7E, 7F), and may represent genes  
854 with similar expression patterns under X-Y mismatch regardless of autosomal  
855 background. All DE genes and their module memberships are listed in Supplemental  
856 Material, Tables S6 and S7.

857

## 858 Discussion

859

860 The large X-effect and Haldane's rule are prevalent patterns observed in intrinsic hybrid  
861 incompatibilities across diverse taxa and suggest that sex chromosomes play a  
862 predominant role in speciation, but the evolutionary forces underlying rapid sex  
863 chromosome divergence that leads to hybrid incompatibilities remain unclear

864 (Presgraves and Meiklejohn 2021). One compelling hypothesis is that hybrid  
865 incompatibilities are a consequence of intragenomic conflict between sex chromosomes  
866 (Frank 1991; Hurst and Pomiankowski 1991; Lindholm *et al.* 2016). In this study, we  
867 showed that intragenomic conflict between the sex chromosomes may contribute to  
868 some hybrid incompatibilities in house mice, but not in a simple dosage-dependent  
869 manner, and with subtle effects relative to other components of F1 hybrid  
870 incompatibilities. Notably, we find that X-Y conflict does not appear to contribute to  
871 postmeiotic disruption of sex chromosome repression, a major regulatory phenotype  
872 associated with hybrid sterility in house mice (Larson *et al.* 2017). Below, we discuss  
873 the implications of our findings for the genetic basis of house mouse male hybrid sterility  
874 and the potential role of intragenomic conflict in speciation.

875

### 876 **Insights into the Genetic Basis of Mouse Male Hybrid Sterility**

877 Our results did not support the model of *S/x-* and *S/x1-S/y* dosage imbalance leading to  
878 X chromosome overexpression in mouse F1 hybrids. In Hybrid F1 XY Match, we  
879 showed that X-Y match on an F1 background did not restore postmeiotic X  
880 chromosome repression (Figure 5). In Non-hybrid XY Mismatch, we directly tested the  
881 effects of X-Y mismatch in the absence of X-autosomal mismatch on postmeiotic  
882 spermatogenesis gene expression. We found some evidence for disrupted expression  
883 in X-Y mismatch mice (Figure 5, Table 2), but the effects were relatively subtle and  
884 often in the opposite direction than expected based on genetic manipulation studies  
885 (Cocquet *et al.* 2012; Kruger *et al.* 2019) or disrupted expression in sterile F1 mice  
886 (Larson *et al.* 2017; Figures 4, 5, and 6).

887 Our results indicate that genetic manipulation studies, which performed nearly  
888 complete knockdowns or duplications, are not representative of the more subtle copy  
889 number differences expected to occur in natural hybrids (Cocquet *et al.* 2009; Cocquet  
890 *et al.* 2012; Kruger *et al.* 2019). Another important difference from genetic manipulation  
891 studies is that we used wild-derived inbred strains instead of the C57BL/6J classic  
892 laboratory mouse, which has a mostly *domesticus* background but some *musculus*  
893 introgression throughout, including the Y chromosome (Nagamine *et al.* 1992). Because  
894 C57BL/6J is mostly *domesticus* with a *musculus* Y chromosome, it has a similar genetic

895 composition as our wild-derived *dom*<sup>musY</sup> mice and therefore may show some of the  
896 same subtle disruptions to gene expression and sperm morphology that we observed  
897 compared to pure *domesticus* mice. We also introgressed the entire Y chromosome, so  
898 there should not have been dosage imbalances among ampliconic genes on the same  
899 sex chromosome. However, our Y-introgression mice also had imbalance between all  
900 Y-linked ampliconic genes and interacting genes on the X chromosome and autosomes,  
901 so it is unclear if introgressing the entire Y chromosome should cause larger or smaller  
902 effects on postmeiotic spermatogenesis expression.

903 SLX, SLXL1, and SLY proteins interact with other sex-linked and autosomal  
904 ampliconic genes, including *Ssty1/2*, *α-takusan*, and *Speer*, so additional gene families  
905 may be involved in intragenomic conflict with *Slx*, *Slx1/1*, and *Sly* (Kruger *et al.* 2019;  
906 Moretti *et al.* 2020). Our autosomal gene family expression results seem to further  
907 complicate understanding of the consequences of ampliconic gene conflict as we found  
908 that the *α-takusan* gene family is overexpressed in F1 hybrids regardless of cross  
909 direction or sex chromosome type (Figure 4F). Sex chromosome mismatch, however,  
910 did not disrupt *α-takusan* expression when the autosomal background was non-hybrid.  
911 This was somewhat puzzling because protein products of sex-linked ampliconic genes  
912 are thought to regulate *α-takusan* expression in late spermatogenesis, perhaps again  
913 indicating that copy number differences between subspecies are too subtle to generate  
914 strong regulatory phenotypes. Another surprising expression result was that *Slx1/1*  
915 expression levels were not correlated with *Slx1/1* copy numbers (Figure 4B). Other genes  
916 are likely involved in the regulation of *Slx1/1* (Moretti *et al.* 2020), and it is possible that  
917 the evolution of these *trans*-acting factors may play a more important role in determining  
918 overall *Slx1/1* expression levels than *Slx1/1* copy number *per se*.

919 On balance, our results suggest that differences in *Slx*- or *Slx1/1*-*Sly* dosage do  
920 not result in strong hybrid incompatibilities. We did not observe sex chromosome  
921 overexpression with an excess of *Slx* and *Slx1/1* copies or underexpression with an  
922 excess of *Sly* copies as predicted under the conflict model (Larson *et al.* 2017).  
923 Therefore, the primary mechanisms underlying postmeiotic X chromosome  
924 overexpression in sterile F1 hybrids likely do not involve X-Y interactions. Instead,

925 disrupted postmeiotic repression is likely a continuation of *Prdm9*-mediated MSCI  
926 disruption (Bhattacharyya *et al.* 2013; Bhattacharyya *et al.* 2014; Mukaj *et al.* 2020).

927         Although X-Y copy number imbalance is unlikely to explain disrupted postmeiotic  
928 repression in F1 hybrids, sex chromosome interactions may play a role in house mouse  
929 hybrid sterility. We showed that X-Y mismatch can lead to disrupted expression of  
930 ampliconic genes and other genes throughout the genome (Figure 4, Figure 6, Table 2),  
931 and some of these genes are essential for spermatogenesis. For example, *Taf7l*  
932 knockouts have abnormal sperm morphology (Cheng *et al.* 2007), *Prdx4* knockouts  
933 have reduced sperm counts (Iuchi *et al.* 2009), and both these genes were differentially  
934 expressed in *dom<sup>musY</sup>* mice. We also showed that hybrid interactions involving the Y-  
935 chromosome are associated with subfertility phenotypes (Table 1), consistent with  
936 previous studies (Campbell *et al.* 2012; Campbell and Nachman 2014). Here we have  
937 focused on interactions between the sex chromosomes because the ampliconic gene  
938 conflict model established a clear prediction for X-Y incompatibilities, but we could not  
939 distinguish X-Y incompatibilities from Y-autosomal incompatibilities in our experimental  
940 crosses. We note that several of our observations could result from Y-autosomal  
941 interactions. Indeed, introgression of the Y chromosome (Non-hybrid XY Mismatch)  
942 induced autosomal regulatory phenotypes.

943         We observed some autosomal regions that co-introgressed with the Y  
944 chromosome, and some of these regions have been shown to introgress in other mouse  
945 hybrids (Supplemental Material, Figure S4). These may be regions that are  
946 incompatible with the Y chromosome from the opposite subspecies, and therefore must  
947 co-introgress for mice to be viable or fertile. The large introgressed region we identified  
948 on chromosome 2 is adjacent to a multicopy gene, *R2d2*, involved in meiotic drive  
949 during female meiosis (Didion *et al.* 2016). *R2d2* has only been shown to act in females  
950 (Didion *et al.* 2016), but our crossing scheme only involved backcrossing hybrid males.  
951 We also generated Y-introgression mice using the LEWES/EiJ strain, which is fixed for  
952 the low copy number allele of *R2d2*, and PWK/PhJ, which also appears to have low  
953 *R2d2* copy number (Didion *et al.* 2016), so it is unlikely that this introgression is a direct  
954 result of *R2d2* drive as previously described. Nevertheless, the exact functions of *R2d2*

955 are unresolved, so this large region of introgression may be related to *R2d2*, but  
956 probably not through a direct meiotic drive mechanism.

957 Our results are likely important in the context of mouse speciation in nature. Mice  
958 sampled from the European hybrid zone are often advanced generation hybrids with  
959 complex patterns of ancestry from both *musculus* and *domesticus*, and true F1  
960 genotypes are exceptionally rare (Teeter *et al.* 2010; Turner *et al.* 2012). Therefore,  
961 understanding mechanisms of hybrid incompatibility in addition to F1 X-autosomal  
962 incompatibilities is essential for understanding the complex genetic basis of mouse  
963 speciation occurring in nature. The Non-hybrid XY Mismatch experiment demonstrated  
964 that disrupted gene expression phenotypes can occur in the absence of an F1  
965 autosomal background. Previous studies have shown that advanced intercrosses of  
966 hybrid mice show different sterility phenotypes than F1s (Campbell *et al.* 2012), and  
967 *Prdm9*-mediated hybrid sterility requires an F1 autosomal background, leading others to  
968 speculate that genetic incompatibilities underlying hybrid sterility may be different in  
969 later hybrid generations (Campbell and Nachman 2014; Mukaj *et al.* 2020). Our results  
970 show that Y chromosome introgression can contribute to reduced fertility (consistent  
971 with (Campbell *et al.* 2012) and some disrupted spermatogenesis gene expression in  
972 later generation hybrids with non-F1 autosomal backgrounds.

973

### 974 **What Is the Contribution of Sex Chromosome Conflict to Speciation?**

975 Several studies have proposed a link between intragenomic conflict and hybrid  
976 incompatibilities (Tao *et al.* 2001; Phadnis and Orr 2009; Wilkinson *et al.* 2014; Zanders  
977 *et al.* 2014; Case *et al.* 2015; Zhang *et al.* 2015; Larson *et al.* 2017), but it remains  
978 unknown how prevalent these systems are in natural populations or if intragenomic  
979 conflict is the primary selective force behind the evolution of these incompatibilities.  
980 While X-autosomal incompatibilities are known to play a central role in house mouse  
981 hybrid sterility, previous work has shown that house mouse speciation likely has a more  
982 complex genetic basis (Vyskočilová *et al.* 2005; Good *et al.* 2008b; Turner *et al.* 2012;  
983 Turner and Harr 2014; Larson *et al.* 2018b) and may involve sex chromosome  
984 intragenomic conflict (Ellis *et al.* 2011; Campbell *et al.* 2012; Larson *et al.* 2017). The  
985 exact mechanisms underlying reduced fertility associated with Y chromosome mismatch

986 is unknown, and it is still unclear what role, if any, sex chromosome intragenomic  
987 conflict may play (Ellis *et al.* 2011; Campbell *et al.* 2012; Larson *et al.* 2017).

988        Ampliconic genes are a common feature of mammalian sex chromosomes, and  
989 they tend to be expressed specifically during spermatogenesis (Li *et al.* 2013; Soh *et al.*  
990 2014; Skinner *et al.* 2016; Lucotte *et al.* 2018; Bellott *et al.* 2017; Hughes *et al.* 2020;  
991 reviewed in Larson *et al.* 2018a). Although difficult to quantify, evolution of ampliconic  
992 gene families involved in spermatogenesis is arguably one of the most rapidly evolving  
993 components of mammalian genomes (Mueller *et al.* 2013; Soh *et al.* 2014; Lucotte *et al.*  
994 2018; Cechova *et al.* 2020; Vegesna *et al.* 2020). Intragenomic conflict among sex  
995 chromosome ampliconic genes has been proposed as a mechanism through which  
996 hybrid incompatibilities have evolved in at least three mammalian groups (Davis *et al.*  
997 2015; Dutheil *et al.* 2015; Larson *et al.* 2018a; Kruger *et al.* 2019). In cats, loci  
998 associated with hybrid sterility tend to be in or near high copy number genes (Davis *et*  
999 *al.* 2015). In great apes, sex chromosome amplicon copy number can evolve rapidly  
1000 (Lucotte *et al.* 2018; Cechova *et al.* 2020), and ampliconic regions on the X  
1001 chromosome are thought to have experienced selective sweeps as a result of strong  
1002 selection pressures imposed by intragenomic conflict with the Y chromosome (Nam *et*  
1003 *al.* 2015). These regions also overlap sections of the modern human X chromosome  
1004 that lack Neandertal introgression, and therefore may represent regions involved in  
1005 genetic incompatibilities between modern humans and Neandertals (Dutheil *et al.*  
1006 2015). However, most of these connections remain speculative and the X chromosome  
1007 is clearly a hotspot of the evolution of hybrid incompatibilities (Masly and Presgraves  
1008 2007; Good *et al.* 2008a).

1009        Theoretical work introducing the idea that sex chromosome intragenomic conflict  
1010 could contribute to hybrid incompatibilities focused on this phenomenon as an  
1011 explanation for Haldane's rule and the large X-effect (Frank 1991; Hurst and  
1012 Pomiankowski 1991). However, genetic conflict between the sex chromosomes during  
1013 reproduction cannot explain some observations, such as the applicability of Haldane's  
1014 rule and the large X-effect to hybrid inviability or the important role of the X chromosome  
1015 in many incompatibilities that occur in homogametic hybrids (Coyne 1992). In this study,  
1016 we showed that X-Y conflict may have a small effect on male hybrid sterility, but *Prdm9*-

1017 mediated incompatibilities probably play the most important role in the observations  
1018 consistent with Haldane's rule and the large X-effect in house mice. Interactions among  
1019 *Prdm9*, *Hstx2*, and other autosomal and X-linked loci in hybrids result in failed or  
1020 delayed double strand break repair, which eventually leads to meiotic arrest and male  
1021 sterility (Forejt *et al.* 2021). The rapid divergence of *Prdm9* and its binding sites is likely  
1022 the result of PRDM9 haplotype selection, leading to biased gene conversion and  
1023 hotspot erosion (Baker *et al.* 2015). Thus, intragenomic conflict is unlikely to be the  
1024 primary underlying cause of house mouse hybrid male sterility.

1025         It remains unknown if the recurrent evolution of ampliconic genes is a  
1026 consequence of intragenomic conflict across mammals, but this is generally assumed to  
1027 be the case. If so, intragenomic conflict may be much more important in the evolution of  
1028 hybrid incompatibility loci than once thought (Johnson and Wu 1992; Coyne and Orr  
1029 2004). Some recent empirical studies support this hypothesis in both flies and mammals  
1030 (Presgraves and Meiklejohn 2021), however, our study did not provide direct support for  
1031 this hypothesis. X-Y mismatch likely contributes to hybrid male sterility and disrupted  
1032 expression, but in more complex ways than the *Slx*, *Slx1/1*, and *Sly* dosage-based  
1033 conflict model, and with relatively small effects on hybrid sterility. In particular, we note  
1034 that *Ssty1* and *Sly* show opposing copy number patterns between subspecies, such that  
1035 replacing a *musculus* Y with a *domesticus* Y simultaneously increases *Ssty1* while  
1036 decreasing *Sly*, and vice versa. It is possible that higher copy number of one gene can  
1037 compensate for reduced copy number of the other in regulating postmeiotic sex  
1038 chromatin. This work is thus distinct from previous work focused on deletions (that  
1039 reduce copy number of both genes) or RNA interference (that selectively targets one  
1040 gene).

1041         Further work is required to identify loci involved in X-Y or Y-autosomal  
1042 incompatibilities, but it is plausible that intragenomic conflict among ampliconic genes  
1043 still plays a role given that these genes are the primary sex chromosome genes  
1044 expressed in the postmeiotic stages during which spermatogenesis expression is highly  
1045 disrupted (Sin and Namekawa 2013; Larson *et al.* 2017). Copy number mismatch  
1046 between these gene families may play important roles in reproductive outcomes in  
1047 nature, as has been implied from slight sex ratio skews in regions of the hybrid zone

1048 (Macholán *et al.* 2008). Even subtle differences in fertility could have important effects  
1049 on fitness, especially given that sperm competition appears to be common in mice  
1050 (Dean *et al.* 2006). However, our work suggests that such effects do not manifest as a  
1051 major reproductive barrier between populations.

1052

### 1053 **Data Availability**

1054 Whole genome sequence data from Y-introgression strains and RNAseq data from  
1055 testes cell sort populations are publicly available through the National Center for  
1056 Biotechnology Information Sequence Read Archive under accession numbers  
1057 PRJNA816542 (whole genome) and PRJNA816886 (RNAseq). Raw phenotype data  
1058 are available in the Supplemental Material, Table S2. Scripts used to modify the  
1059 AmpliCoNE program for copy number estimation are publicly available at:

1060 <https://github.com/ekopania/modified-AmpliCoNE>. Scripts used for gene expression  
1061 analyses are available at:

1062 [https://github.com/ekopania/xy\\_mismatch\\_expression\\_analyses](https://github.com/ekopania/xy_mismatch_expression_analyses).

1063

### 1064 **Acknowledgements**

1065 We would like to thank Sara Keeble for assistance with animal husbandry, Pamela K.  
1066 Shaw and the UM Fluorescence Cytometry Core supported by an Institutional  
1067 Development Award from the NIGMS (P30GM103338 and S10-OD025019) and the UM  
1068 Lab Animal Resources staff. This study included research conducted in the University of  
1069 Montana Genomics Core, supported by a grant from the M. J. Murdock Charitable Trust  
1070 (to J.M.G.). Computational resources and support from the University of Montana's Griz  
1071 Shared Computing Cluster (GSCC), supported by grants from the National Science  
1072 Foundation (CC-2018112 and OAC-1925267, J.M.G. co-PI), contributed to this  
1073 research. We thank members of the Good Lab, Polly Campbell, Lila Fishman, Doug  
1074 Emlen, and Travis Wheeler for helpful discussions and comments during the  
1075 development of this study. We thank three anonymous reviewers for helpful comments  
1076 on an earlier version of this manuscript. This work was supported by grants from the  
1077 Eunice Kennedy Shriver National Institute of Child Health and Human Development of  
1078 the National Institutes of Health (R01-HD073439, R01-HD094787 to J.M.G.). E.E.K.K.

1079 was supported by the National Science Foundation Graduate Research Fellowship  
1080 Program (DGE-1313190), a grant from the National Science Foundation (DEB-1754096  
1081 to J.M.G.), and a Rosemary Grant Award for Graduate Student Research from the  
1082 Society for the Study of Evolution. E.L.L. was supported by the National Science  
1083 Foundation (DEB-2012041). C.C.R. was supported by the BBSRC (BB/N000463/1 to  
1084 P.E.). P.E. also acknowledges funding from the Leverhulme Trust (RPG-2019-414 194).  
1085 E.M.W. was supported by a faculty studentship from the University of Essex. B.M.S was  
1086 supported by UKRI (University of Essex). Any opinions, findings, and conclusions or  
1087 recommendations expressed in this material are those of the author(s) and do not  
1088 necessarily reflect the views of the National Science Foundation, the National Institutes  
1089 of Health, or the Society for the Study of Evolution.

1090

### 1091 **Author Contributions**

1092 J.M.G. and E.L.L. conceived and funded the project. E.E.K.K., E.L.L., and J.M.G.  
1093 designed the experiments. E.L.L. and E.E.K.K. did the mouse husbandry and breeding.  
1094 E.E.K.K. performed the mouse dissections, cell sorts, and sequencing library  
1095 preparation. E.M.W., C.C.R., B.M.S., and P.J.I.E. performed and analyzed the sperm  
1096 morphology assays. E.E.K.K. analyzed the data. E.E.K.K., E.L.L., and J.M.G. wrote the  
1097 manuscript with input from all authors.

1098

### 1099 **Conflicts of Interest**

1100 The authors declare no conflicts of interest.

1101

### 1102 **References**

- 1103 Baker, C. L., S. Kajita, M. Walker, R. L. Saxl, N. Raghupathy *et al.*, 2015 PRDM9 Drives  
1104 Evolutionary Erosion of Hotspots in *Mus musculus* through Haplotype-Specific Initiation  
1105 of Meiotic Recombination. *PLOS Genetics* 11: e1004916.  
1106 Balcova, M., B. Faltusova, V. Gergelits, T. Bhattacharyya, O. Mihola *et al.*, 2016 Hybrid Sterility  
1107 Locus on Chromosome X Controls Meiotic Recombination Rate in Mouse. *PLOS*  
1108 *Genetics* 12: e1005906.  
1109 Bellott, D. W., H. Skaletsky, T.-J. Cho, L. Brown, D. Locke *et al.*, 2017 Avian W and mammalian  
1110 Y chromosomes convergently retained dosage-sensitive regulators. *Nat Genet* 49: 387-  
1111 394.

- 1112 Bhattacharyya, T., S. Gregorova, O. Mihola, M. Anger, J. Sebestova *et al.*, 2013 Mechanistic  
1113 basis of infertility of mouse intersubspecific hybrids. *Proceedings of the National*  
1114 *Academy of Sciences of the United States of America* 110: E468-E477.
- 1115 Bhattacharyya, T., R. Reifova, S. Gregorova, P. Simecek, V. Gergelits *et al.*, 2014 X  
1116 Chromosome Control of Meiotic Chromosome Synapsis in Mouse Inter-Subspecific  
1117 Hybrids. *PLOS Genetics* 10: e1004088.
- 1118 Bi, Y., X. Ren, R. Li, Q. Ding, D. Xie *et al.*, 2019 Specific Interactions Between Autosome and X  
1119 Chromosomes Cause Hybrid Male Sterility in *Caenorhabditis* Species. *Genetics* 212:  
1120 801-813.
- 1121 Bolger, A. M., M. Lohse and B. Usadel, 2014 Trimmomatic: a flexible trimmer for Illumina  
1122 sequence data. *Bioinformatics* 30: 2114-2120.
- 1123 Burgoyne, P. S., S. K. Mahadevaiah and J. M. A. Turner, 2009 The consequences of asynapsis  
1124 for mammalian meiosis. *Nat Rev Genet* 10: 207-216.
- 1125 Campbell, P., J. M. Good, M. D. Dean, P. K. Tucker and M. W. Nachman, 2012 The  
1126 Contribution of the Y Chromosome to Hybrid Male Sterility in House Mice. *Genetics* 191:  
1127 1271-1281.
- 1128 Campbell, P., J. M. Good and M. W. Nachman, 2013 Meiotic Sex Chromosome Inactivation Is  
1129 Disrupted in Sterile Hybrid Male House Mice. *Genetics* 193: 819-828.
- 1130 Campbell, P., and M. W. Nachman, 2014 X–Y Interactions Underlie Sperm Head Abnormality in  
1131 Hybrid Male House Mice. *Genetics* 196: 1231-1240.
- 1132 Case, L. K., E. H. Wall, E. E. Osmanski, J. A. Dragon, N. Saligrama *et al.*, 2015 Copy number  
1133 variation in Y chromosome multicopy genes is linked to a paternal parent-of-origin effect  
1134 on CNS autoimmune disease in female offspring. *Genome Biology* 16: 28.
- 1135 Cechova, M., R. Vegesna, M. Tomaszewicz, R. S. Harris, D. Chen *et al.*, 2020 Dynamic  
1136 evolution of great ape Y chromosomes. *Proceedings of the National Academy of*  
1137 *Sciences*: 202001749.
- 1138 Charlesworth, B., J. A. Coyne and N. H. Barton, 1987 The Relative Rates of Evolution of Sex  
1139 Chromosomes and Autosomes. *American Naturalist* 130: 113-146.
- 1140 Cheng, Y., M. G. Buffone, M. Kouadio, M. Goodheart, D. C. Page *et al.*, 2007 Abnormal Sperm  
1141 in Mice Lacking the *Taf7l* Gene. *Molecular and Cellular Biology* 27: 2582-2589.
- 1142 Cocquet, J., P. J. Ellis, S. K. Mahadevaiah, N. A. Affara, D. Vaiman *et al.*, 2012 A genetic basis  
1143 for a postmeiotic X versus Y chromosome intragenomic conflict in the mouse. *PLoS*  
1144 *Genet* 8: e1002900.
- 1145 Cocquet, J., P. J. Ellis, Y. Yamauchi, S. K. Mahadevaiah, N. A. Affara *et al.*, 2009 The multicopy  
1146 gene *Sly* represses the sex chromosomes in the male mouse germline after meiosis.  
1147 *PLoS Biol* 7: e1000244.
- 1148 Conway, J. R., A. Lex and N. Gehlenborg, 2017 UpSetR: an R package for the visualization of  
1149 intersecting sets and their properties. *Bioinformatics* 33: 2938-2940.
- 1150 Coughlan, J. M., and D. R. Matute, 2020 The importance of intrinsic postzygotic barriers  
1151 throughout the speciation process. *Philosophical Transactions of the Royal Society B:*  
1152 *Biological Sciences* 375: 20190533.
- 1153 Coyne, J. A., 1992 *Genetics and Speciation*. *Nature* 355: 511-515.
- 1154 Coyne, J. A., and H. A. Orr, 1989 Two rules of speciation, pp. 180–207 in *Speciation and its*  
1155 *consequences*, edited by D. Otte and J. Endler. Sinauer Associates, Sunderland  
1156 (Massachusetts).
- 1157 Coyne, J. A., and H. A. Orr, 2004 *Speciation*. W.H. Freeman.
- 1158 Davies, B., E. Hatton, N. Altemose, J. G. Hussin, F. Pratto *et al.*, 2016 Re-engineering the zinc  
1159 fingers of PRDM9 reverses hybrid sterility in mice. *Nature* 530: 171-176.
- 1160 Davis, B. W., C. M. Seabury, W. A. Brashear, G. Li, M. Roelke-Parker *et al.*, 2015 Mechanisms  
1161 Underlying Mammalian Hybrid Sterility in Two Feline Interspecies Models. *Molecular*  
1162 *Biology and Evolution* 32: 2534-2546.

1163 Dean, M. D., K. G. Ardlie and M. W. Nachman, 2006 The frequency of multiple paternity  
1164 suggests that sperm competition is common in house mice (*Mus domesticus*). *Molecular*  
1165 *Ecology* 15: 4141-4151.

1166 Didion, J. P., A. P. Morgan, L. Yadgary, T. A. Bell, R. C. McMullan *et al.*, 2016 *R2d2* Drives  
1167 Selfish Sweeps in the House Mouse. *Molecular Biology and Evolution* 33: 1381-1395.

1168 Dobzhansky, T., 1937 Genetic Nature of Species Differences. *The American Naturalist* 71: 404-  
1169 420.

1170 Dumont, B. L., and B. A. Payseur, 2011 Genetic Analysis of Genome-Scale Recombination  
1171 Rate Evolution in House Mice. *PLOS Genetics* 7: e1002116.

1172 Dutheil, J. Y., K. Munch, K. Nam, T. Mailund and M. H. Schierup, 2015 Strong Selective  
1173 Sweeps on the X Chromosome in the Human-Chimpanzee Ancestor Explain Its Low  
1174 Divergence. *PLOS Genetics* 11: e1005451.

1175 Edwards, J. A., and R. A. Edwards, 2019 Fastq-pair: efficient synchronization of paired-end  
1176 fastq files. *bioRxiv*: 552885.

1177 Ellis, P. J., J. Bacon and N. A. Affara, 2011 Association of Sly with sex-linked gene amplification  
1178 during mouse evolution: a side effect of genomic conflict in spermatids? *Hum Mol Genet*  
1179 20: 3010-3021.

1180 Forejt, J., P. Jansa and E. Parvanov, 2021 Hybrid sterility genes in mice (*Mus musculus*): a  
1181 peculiar case of PRDM9 incompatibility. *Trends in Genetics*.

1182 Frank, S. A., 1991 Divergence of Meiotic Drive-Suppression Systems as an Explanation for  
1183 Sex- Biased Hybrid Sterility and Inviability. *Evolution* 45: 262-267.

1184 Getun, I. V., B. Torres and P. R. J. Bois, 2011 Flow Cytometry Purification of Mouse Meiotic  
1185 Cells. *Journal of Visualized Experiments: JoVE*: 2602.

1186 Good, J. M., 2012 The Conflict within and the Escalating War between the Sex Chromosomes.  
1187 *PLoS Genet* 8: e1002955.

1188 Good, J. M., M. D. Dean and M. W. Nachman, 2008a A Complex Genetic Basis to X-Linked  
1189 Hybrid Male Sterility Between Two Species of House Mice. *Genetics* 179: 2213-2228.

1190 Good, J. M., T. Giger, M. D. Dean and M. W. Nachman, 2010 Widespread Over-Expression of  
1191 the X Chromosome in Sterile F<sub>1</sub> Hybrid Mice. *PLoS Genet* 6: e1001148.

1192 Good, J. M., M. A. Handel and M. W. Nachman, 2008b Asymmetry and polymorphism of hybrid  
1193 male sterility during the early stages of speciation in house mice. *Evolution* 62: 50-65.

1194 Green, C. D., Q. Ma, G. L. Manske, A. N. Shami, X. Zheng *et al.*, 2018 A Comprehensive  
1195 Roadmap of Murine Spermatogenesis Defined by Single-Cell RNA-Seq. *Developmental*  
1196 *Cell* 46: 651-667.e610.

1197 Grey, C., P. Barthès, G. Chauveau-Le Friec, F. Langa, F. Baudat *et al.*, 2011 Mouse PRDM9  
1198 DNA-Binding Specificity Determines Sites of Histone H3 Lysine 4 Trimethylation for  
1199 Initiation of Meiotic Recombination. *PLOS Biology* 9: e1001176.

1200 Haldane, J. B. S., 1922 Sex ratio and unisexual sterility in hybrid animals. *Journal of Genetics*  
1201 12: 101-109.

1202 Harr, B., E. Karakoc, R. Neme, M. Teschke, C. Pfeifle *et al.*, 2016 Genomic resources for wild  
1203 populations of the house mouse, *Mus musculus* and its close relative *Mus spretus*. *Sci*  
1204 *Data* 3: 160075.

1205 Hughes, J. F., H. Skaletsky, T. Pyntikova, N. Koutseva, T. Raudsepp *et al.*, 2020 Sequence  
1206 analysis in *Bos taurus* reveals pervasiveness of X–Y arms races in mammalian lineages.  
1207 *Genome Research* 30: 1716-1726.

1208 Hunnicutt, K. E., J. M. Good and E. L. Larson, 2021 Unraveling patterns of disrupted gene  
1209 expression across a complex tissue. *Evolution* 76: 275-291.

1210 Hurst, L. D., and A. Pomiankowski, 1991 Causes of sex ratio bias may account for unisexual  
1211 sterility in hybrids: a new explanation of Haldane's rule and related phenomena.  
1212 *Genetics* 128: 841-858.

- 1213 Iuchi, Y., F. Okada, S. Tsunoda, N. Kibe, N. Shirasawa *et al.*, 2009 Peroxiredoxin 4 knockout  
 1214 results in elevated spermatogenic cell death via oxidative stress. *Biochemical Journal*  
 1215 419: 149-158.
- 1216 Janoušek, V., L. Wang, K. E. N. Luzynski, P. Dufková, M. Vyskočilová Martina *et al.*, 2012  
 1217 Genome-wide architecture of reproductive isolation in a naturally occurring hybrid zone  
 1218 between *Mus musculus musculus* and *M. m. domesticus*. *Molecular Ecology* 21: 3032-  
 1219 3047.
- 1220 Johnson, N. A., and C. I. Wu, 1992 An empirical test of the meiotic drive models of hybrid  
 1221 sterility: sex-ratio data from hybrids between *Drosophila simulans* and *Drosophila*  
 1222 *sechellia*. *Genetics* 130: 507-511.
- 1223 Kruger, A. N., M. A. Brogley, J. L. Huizinga, J. M. Kidd, D. G. de Rooij *et al.*, 2019 A  
 1224 Neofunctionalized X-Linked Ampliconic Gene Family Is Essential for Male Fertility and  
 1225 Equal Sex Ratio in Mice. *Current Biology* 29: 3699-3706.e3695.
- 1226 Langfelder, P., and S. Horvath, 2008 WGCNA: an R package for weighted correlation network  
 1227 analysis. *BMC Bioinformatics* 9: 559.
- 1228 Larson, E. L., S. Keeble, D. Vanderpool, M. D. Dean and J. M. Good, 2017 The Composite  
 1229 Regulatory Basis of the Large X-Effect in Mouse Speciation. *Molecular Biology and*  
 1230 *Evolution* 34: 282-295.
- 1231 Larson, E. L., E. E. K. Kopania and J. M. Good, 2018a Spermatogenesis and the Evolution of  
 1232 Mammalian Sex Chromosomes. *Trends Genet* 34: 722-732.
- 1233 Larson, E. L., E. E. K. Kopania, K. E. Hunnicutt, D. Vanderpool, S. Keeble *et al.*, 2021 Stage-  
 1234 specific disruption of X chromosome expression during spermatogenesis in sterile house  
 1235 mouse hybrids. *G3 Genes|Genomes|Genetics* 12: jkab407.
- 1236 Larson, E. L., D. Vanderpool, B. A. J. Sarver, C. Callahan, S. Keeble *et al.*, 2018b The Evolution  
 1237 of Polymorphic Hybrid Incompatibilities in House Mice. *Genetics* 209: 845–859.
- 1238 Li, G., B. W. Davis, T. Raudsepp, A. J. Pearks Wilkerson, V. C. Mason *et al.*, 2013 Comparative  
 1239 analysis of mammalian Y chromosomes illuminates ancestral structure and lineage-  
 1240 specific evolution. *Genome Research* 23: 1486-1495.
- 1241 Li, H., and R. Durbin, 2009 Fast and accurate short read alignment with Burrows–Wheeler  
 1242 transform. *Bioinformatics* 25: 1754-1760.
- 1243 Lindholm, A. K., K. A. Dyer, R. C. Firman, L. Fishman, W. Forstmeier *et al.*, 2016 The Ecology  
 1244 and Evolutionary Dynamics of Meiotic Drive. *Trends in Ecology & Evolution* 31: 315-326.
- 1245 Lucotte, E. A., L. Skov, J. M. Jensen, M. C. Macià, K. Munch *et al.*, 2018 Dynamic Copy  
 1246 Number Evolution of X- and Y-Linked Ampliconic Genes in Human Populations.  
 1247 *Genetics* 209: 907-920.
- 1248 Macholán, M., S. J. Baird, P. Munclinger, P. Dufková, B. Bímová *et al.*, 2008 Genetic conflict  
 1249 outweighs heterogametic incompatibility in the mouse hybrid zone? *BMC Evolutionary*  
 1250 *Biology* 8: 271.
- 1251 Masly, J. P., and D. C. Presgraves, 2007 High-Resolution Genome-Wide Dissection of the Two  
 1252 Rules of Speciation in *Drosophila*. *PLOS Biology* 5: e243.
- 1253 Matute, D. R., and B. S. Cooper, 2021 Comparative studies on speciation: 30 years since  
 1254 Coyne and Orr. *Evolution* 75: 764-778.
- 1255 Meiklejohn, C. D., and Y. Tao, 2010 Genetic conflict and sex chromosome evolution. *Trends in*  
 1256 *Ecology & Evolution* 25: 215-223.
- 1257 Mihola, O., Z. Trachtulec, C. Vlcek, J. C. Schimenti and J. Forejt, 2009 A Mouse Speciation  
 1258 Gene Encodes a Meiotic Histone H3 Methyltransferase. *Science* 323: 373-375.
- 1259 Moretti, C., M. Blanco, C. Ialy-Radio, M.-E. Serrentino, C. Gobé *et al.*, 2020 Battle of the Sex  
 1260 Chromosomes: Competition between X and Y Chromosome-Encoded Proteins for  
 1261 Partner Interaction and Chromatin Occupancy Drives Multicopy Gene Expression and  
 1262 Evolution in Muroid Rodents. *Molecular Biology and Evolution* 37: 3453–3468.

1263 Morgan, A. P., and F. Pardo-Manuel de Villena, 2017 Sequence and Structural Diversity of  
1264 Mouse Y Chromosomes. *Molecular Biology and Evolution* 34: 3186-3204.

1265 Mueller, J. L., S. K. Mahadevaiah, P. J. Park, P. E. Warburton, D. C. Page *et al.*, 2008 The  
1266 mouse X chromosome is enriched for multicopy testis genes showing postmeiotic  
1267 expression. *Nat Genet* 40: 794-799.

1268 Mueller, J. L., H. Skaletsky, L. G. Brown, S. Zaghlul, S. Rock *et al.*, 2013 Independent  
1269 specialization of the human and mouse X chromosomes for the male germ line. *Nat*  
1270 *Genet* 45: 1083-1087.

1271 Mukaj, A., J. Piálek, V. Fotopulosova, A. P. Morgan, L. Odenthal-Hesse *et al.*, 2020 *Prdm9*  
1272 Intersubspecific Interactions in Hybrid Male Sterility of House Mouse. *Molecular Biology*  
1273 *and Evolution* 37: 3423-3438.

1274 Nagamine, C. M., Y. Nishioka, K. Moriwaki, P. Boursot, F. Bonhomme *et al.*, 1992 The  
1275 musculus-type Y Chromosome of the laboratory mouse is of Asian origin. *Mammalian*  
1276 *Genome* 3: 84-91.

1277 Nam, K., K. Munch, A. Hobolth, J. Y. Dutheil, K. R. Veeramah *et al.*, 2015 Extreme selective  
1278 sweeps independently targeted the X chromosomes of the great apes. *Proc Natl Acad*  
1279 *Sci U S A* 112: 6413-6418.

1280 Namekawa, S. H., P. J. Park, L.-F. Zhang, J. E. Shima, J. R. McCarrey *et al.*, 2006 Postmeiotic  
1281 Sex Chromatin in the Male Germline of Mice. *Current Biology* 16: 660-667.

1282 Oliver, P. L., L. Goodstadt, J. J. Bayes, Z. Birtle, K. C. Roach *et al.*, 2009 Accelerated Evolution  
1283 of the *Prdm9* Speciation Gene across Diverse Metazoan Taxa. *PLOS Genetics* 5:  
1284 e1000753.

1285 Patro, R., G. Duggal, M. I. Love, R. A. Irizarry and C. Kingsford, 2017 Salmon: fast and bias-  
1286 aware quantification of transcript expression using dual-phase inference. *Nature*  
1287 *methods* 14: 417-419.

1288 Pedersen, B. S., and A. R. Quinlan, 2017 Mosdepth: quick coverage calculation for genomes  
1289 and exomes. *Bioinformatics* 34: 867-868.

1290 Pezer, Ž., B. Harr, M. Teschke, H. Babiker and D. Tautz, 2015 Divergence patterns of genic  
1291 copy number variation in natural populations of the house mouse (*Mus musculus*  
1292 *domesticus*) reveal three conserved genes with major population-specific expansions.  
1293 *Genome Research* 25: 1114-1124.

1294 Phadnis, N., and H. A. Orr, 2009 A Single Gene Causes Both Male Sterility and Segregation  
1295 Distortion in *Drosophila* Hybrids. *Science* 323: 376-379.

1296 Phifer-Rixey, M., and M. W. Nachman, 2015 Insights into mammalian biology from the wild  
1297 house mouse *Mus musculus*. *Elife* 4: e05959.

1298 Presgraves, D. C., and C. D. Meiklejohn, 2021 Hybrid Sterility, Genetic Conflict and Complex  
1299 Speciation: Lessons From the *Drosophila simulans* Clade Species. *Frontiers in Genetics*  
1300 12.

1301 Rathje, C. C., E. E. P. Johnson, D. Drage, C. Patinioti, G. Silvestri *et al.*, 2019 Differential  
1302 Sperm Motility Mediates the Sex Ratio Drive Shaping Mouse Sex Chromosome  
1303 Evolution. *Current Biology* 29: 3692-3698.e3694.

1304 Sin, H. S., Y. Ichijima, E. Koh, M. Namiki and S. H. Namekawa, 2012 Human postmeiotic sex  
1305 chromatin and its impact on sex chromosome evolution. *Genome Res* 22: 827-836.

1306 Sin, H. S., and S. H. Namekawa, 2013 The great escape: Active genes on inactive sex  
1307 chromosomes and their evolutionary implications. *Epigenetics* 8: 887-892.

1308 Skinner, B. M., C. C. Rathje, J. Bacon, E. E. P. Johnson, E. L. Larson *et al.*, 2019 A high-  
1309 throughput method for unbiased quantitation and categorization of nuclear morphology.  
1310 *Biology of Reproduction* 100: 1250-1260.

1311 Skinner, B. M., C. A. Sargent, C. Churcher, T. Hunt, J. Herrero *et al.*, 2016 The pig X and Y  
1312 Chromosomes: structure, sequence, and evolution. *Genome Research* 26: 130-139.

- 1313 Smith, C. M., T. F. Hayamizu, J. H. Finger, S. M. Bello, I. J. McCright *et al.*, 2019 The mouse  
1314 Gene Expression Database (GXD): 2019 update. *Nucleic acids research* 47: D774-  
1315 D779.
- 1316 Soh, Y. Q., J. Alfoldi, T. Pyntikova, L. G. Brown, T. Graves *et al.*, 2014 Sequencing the mouse Y  
1317 chromosome reveals convergent gene acquisition and amplification on both sex  
1318 chromosomes. *Cell* 159: 800-813.
- 1319 Storchová, R., S. Gregorová, D. Buckiová, V. Kyselová, P. Divina *et al.*, 2004 Genetic analysis  
1320 of X-linked hybrid sterility in the house mouse. *Mammalian Genome* 15: 515-524.
- 1321 Suzuki, T. A., and M. W. Nachman, 2015 Speciation and reduced hybrid female fertility in house  
1322 mice. *Evolution* 69: 2468-2481.
- 1323 Tao, Y., D. L. Hartl and C. C. Laurie, 2001 Sex-ratio segregation distortion associated with  
1324 reproductive isolation in *Drosophila*. *Proceedings of the National Academy of Sciences*  
1325 98: 13183-13188.
- 1326 Teeter, K. C., L. M. Thibodeau, Z. Gompert, C. A. Buerkle, M. W. Nachman *et al.*, 2010 The  
1327 variable genomic architecture of isolation between hybridizing species of house mice.  
1328 *Evolution* 64: 472-485.
- 1329 The UniProt Consortium, 2020 UniProt: the universal protein knowledgebase in 2021. *Nucleic  
1330 Acids Research* 49: D480-D489.
- 1331 Turelli, M., and H. A. Orr, 1995 The dominance theory of Haldane's rule. *Genetics* 140: 389-  
1332 402.
- 1333 Turelli, M., and H. A. Orr, 2000 Dominance, Epistasis and the Genetics of Postzygotic Isolation.  
1334 *Genetics* 154: 1663-1679.
- 1335 Turner, J. M. A., 2015 Meiotic Silencing in Mammals. *Annual Review of Genetics* 49: 395-412.
- 1336 Turner, J. M. A., S. K. Mahadevaiah, P. J. I. Ellis, M. J. Mitchell and P. S. Burgoyne, 2006  
1337 Pachytene Asynapsis Drives Meiotic Sex Chromosome Inactivation and Leads to  
1338 Substantial Postmeiotic Repression in Spermatids. *Developmental Cell* 10: 521-529.
- 1339 Turner, L. M., and B. Harr, 2014 Genome-wide mapping in a house mouse hybrid zone reveals  
1340 hybrid sterility loci and Dobzhansky-Muller interactions. *eLife* 3: e02504.
- 1341 Turner, L. M., D. J. Schwahn and B. Harr, 2012 Reduced Male Fertility Is Common but Highly  
1342 Variable in Form and Severity in a Natural House Mouse Hybrid Zone. *Evolution* 66:  
1343 443-458.
- 1344 Vegesna, R., M. Tomaszewicz, O. A. Ryder, R. Campos-Sánchez, P. Medvedev *et al.*, 2020  
1345 Ampliconic genes on the great ape Y chromosomes: Rapid evolution of copy number but  
1346 conservation of expression levels. *Genome Biology and Evolution* 12: 842-859.
- 1347 Vicoso, B., and B. Charlesworth, 2009 Effective population size and the faster-X effect: An  
1348 extended model. *Evolution* 63: 2413-2426.
- 1349 Vyskočilová, M., G. Pražanová and J. Piálek, 2009 Polymorphism in hybrid male sterility in wild-  
1350 derived *Mus musculus musculus* strains on proximal chromosome 17. *Mammalian  
1351 Genome* 20: 83.
- 1352 Vyskočilová, M., Z. Trachtulec, J. Forejt and J. Piálek, 2005 Does geography matter in hybrid  
1353 sterility in house mice? *Biological Journal of the Linnean Society* 84: 663-674.
- 1354 Widmayer, S. J., M. A. Handel and D. L. Aylor, 2020 Age and Genetic Background Modify  
1355 Hybrid Male Sterility in House Mice. *Genetics* 216: 585-597.
- 1356 Wilkinson, G. S., S. J. Christianson, C. L. Brand, G. Ru and W. Shell, 2014 Haldane's Rule Is  
1357 Linked to Extraordinary Sex Ratios and Sperm Length in Stalk-Eyed Flies. *Genetics* 198:  
1358 1167-1181.
- 1359 Yang, H., J. R. Wang, J. P. Didion, R. J. Buus, T. A. Bell *et al.*, 2011 Subspecific origin and  
1360 haplotype diversity in the laboratory mouse. *Nat Genet* 43: 648-655.
- 1361 Yates, A. D., P. Achuthan, W. Akanni, J. Allen, J. Allen *et al.*, 2019 Ensembl 2020. *Nucleic  
1362 Acids Research* 48: D682-D688.

- 1363 Zanders, S. E., M. T. Eickbush, J. S. Yu, J.-W. Kang, K. R. Fowler *et al.*, 2014 Genome  
1364 rearrangements and pervasive meiotic drive cause hybrid infertility in fission yeast. *eLife*  
1365 3: e02630.
- 1366 Zhang, L., T. Sun, F. Woldesellassie, H. Xiao and Y. Tao, 2015 Sex Ratio Meiotic Drive as a  
1367 Plausible Evolutionary Mechanism for Hybrid Male Sterility. *PLOS Genetics* 11:  
1368 e1005073.
- 1369  
1370

1371 **Tables**

1372 **Table 1: Reproductive phenotypes for experimental X-Y mismatch mice and**  
 1373 **controls.** Median values are presented +/- 1 standard error. Sample sizes are in  
 1374 parentheses. For sperm morphology parameters (bounding height, bounding width,  
 1375 area, perimeter, difference from median [a measure of the variability of nuclear shapes  
 1376 within the sample]), sample sizes indicate the number of sperm heads observed, and  
 1377 variance is depicted in violin plots (Figure 3; Supplemental Material, Figure S5). Gray  
 1378 boxes indicate significant differences (FDR-corrected Wilcoxon rank sum test  $P < 0.05$ )  
 1379 between X-Y mismatch cross types and control cross types with the same autosomal  
 1380 background. ( $\ddagger$ ) Indicates phenotypes with significant differences (FDR-corrected  
 1381 pairwise Wilcoxon rank sum test  $P < 0.05$ ) between *mus* $\times$ *dom* F1 hybrids and both  
 1382 parental subspecies (*mus* and *dom*). (\*) Indicates phenotypes with significant  
 1383 differences (FDR-correct pairwise Wilcoxon rank sum test  $P < 0.05$ ) between *dom* $\times$ *mus*  
 1384 F1 hybrids and both parental subspecies (*mus* and *dom*). Testes and seminal vesicle  
 1385 weights are both paired. SV = seminal vesicle

Phenotype	Hybrid F1 XY Match				Non-hybrid XY Mismatch			
	<i>dom</i> $\times$ <i>mus</i>	<i>dom</i> $\times$ <i>mus</i> <sup><i>dom</i><i>Y</i></sup>	<i>mus</i> $\times$ <i>dom</i>	<i>mus</i> $\times$ <i>dom</i> <sup><i>mus</i><i>Y</i></sup>	<i>mus</i>	<i>mus</i> <sup><i>dom</i><i>Y</i></sup>	<i>dom</i>	<i>dom</i> <sup><i>mus</i><i>Y</i></sup>
Body mass (g)	20 +/- 0.3 (24)	19.6 +/- 0.3 (7)	17.9 +/- 0.4 (24)	18 +/- 0.4 (12)	19 +/- 0.4 (23)	18.4 +/- 0.3 (47)	18 +/- 0.2 (67)	19 +/- 0.5 (21)
Testes mass (mg) <sup>‡</sup>	186.4 +/- 3 (24)	200.7 +/- 2 (6)	123.9 +/- 2 (23)	125.6 +/- 3 (12)	193.2 +/- 5 (23)	172.7 +/- 2 (47)	209.1 +/- 3 (67)	189.3 +/- 6 (21)
Relative testes mass (mg/g) <sup>*‡</sup>	9.1 +/- 0.1 (24)	10.4 +/- 0.2 (6)	7.2 +/- 0.1 (23)	6.9 +/- 0.1 (12)	10.2 +/- 0.2 (23)	9.2 +/- 0.1 (47)	11.7 +/- 0.1 (67)	10.1 +/- 0.2 (21)
Relative SV mass (mg/g)	6.6 +/- 0.2 (23)	7.3 +/- 0.6 (6)	5.2 +/- 0.3 (24)	5.3 +/- 0.3 (12)	6 +/- 0.3 (23)	6.7 +/- 0.2 (47)	5.2 +/- 0.2 (65)	5.9 +/- 0.3 (21)
Bounding height <sup>‡*</sup>	8.39 (1583)	8.14 (650)	7.46 (870)	7.52 (847)	8.21 (391)	8.02 (401)	8.11 (467)	8.23 (443)
Bounding width <sup>‡*</sup>	5.58 (1583)	5.07 (650)	4.02 (870)	4.09 (847)	5.02 (391)	4.87 (401)	4.9 (467)	5.11 (443)
Area <sup>‡*</sup>	24.5 (1583)	21.6 (650)	20.1 (870)	20.1 (847)	22.1 (391)	20 (401)	21.3 (467)	20.4 (443)
Perimeter <sup>‡*</sup>	23.8 (1583)	22.7 (650)	19.8 (870)	20.2 (847)	22.7 (391)	21.9 (401)	22.3 (467)	23 (443)

Difference from median $\pm$ *	6.22 (1583)	8.67 (650)	8.22 (870)	10.8 (847)	8.88 (391)	5.77 (401)	5.86 (467)	6.72 (443)
--------------------------------	----------------	---------------	---------------	---------------	---------------	---------------	---------------	---------------

1386

1387 **Table 2: Number of differentially expressed genes in round spermatids for**  
 1388 **different cross type comparisons.** “Higher” indicates higher expression (i.e.,  
 1389 overexpressed) in the cross type with X-Y mismatch (F1 hybrids in Larson et al. 2017  
 1390 and Hybrid F1 XY Match, Y-introgression F1 crosses in Non-hybrid XY Mismatch).  
 1391 “Lower” indicates lower expression (i.e., underexpressed) in the cross type with X-Y  
 1392 mismatch. For comparisons in the “Other Contrasts” category, “higher” indicates higher  
 1393 expression in the first cross type listed (*mus* or *mus* $\times$ *dom*). Gray boxes indicate  
 1394 chromosomes that are from the same subspecies in the two cross types being  
 1395 compared. Reciprocal F1s were considered as having the same autosomal  
 1396 backgrounds. Autosomal DE genes overlapping with putatively introgressed regions  
 1397 were excluded from comparisons involving Y-introgression mice.

		Autosomes		X Chromosome		Y Chromosome	
		Higher	Lower	Higher	Lower	Higher	Lower
Larson et al. 2017	<i>mus</i> $\times$ <i>dom</i> vs <i>mus</i>	1518	1476	252	13	109	66
	<i>mus</i> $\times$ <i>dom</i> vs <i>dom</i>	1357	1241	190	55	15	2
	<i>dom</i> $\times$ <i>mus</i> vs <i>mus</i>	1360	1009	62	73	6	8
	<i>dom</i> $\times$ <i>mus</i> vs <i>dom</i>	1237	878	27	73	69	86
Hybrid F1 XY Match	<i>mus</i> $\times$ <i>dom</i> vs <i>mus</i> $\times$ <i>dom</i> <sup><i>mus</i><i>Y</i></sup>	3	3	2	0	74	70
	<i>mus</i> $\times$ <i>dom</i> vs <i>dom</i> $\times$ <i>mus</i> <sup><i>dom</i><i>Y</i></sup>	21	83	38	96	68	84
	<i>dom</i> $\times$ <i>mus</i> vs <i>mus</i> $\times$ <i>dom</i> <sup><i>mus</i><i>Y</i></sup>	372	122	44	101	76	85
	<i>dom</i> $\times$ <i>mus</i> vs <i>dom</i> $\times$ <i>mus</i> <sup><i>dom</i><i>Y</i></sup>	2	4	1	0	71	66
Non-hybrid XY Mismatch	<i>mus</i> <sup><i>dom</i><i>Y</i></sup> vs <i>mus</i>	13	34	1	0	52	63
	<i>mus</i> <sup><i>dom</i><i>Y</i></sup> vs <i>dom</i>	1820	2269	28	179	3	69
	<i>dom</i> <sup><i>mus</i><i>Y</i></sup> vs <i>mus</i>	1634	1679	70	55	10	7
	<i>dom</i> <sup><i>mus</i><i>Y</i></sup> vs <i>dom</i>	13	63	0	10	14	70
Other Contrasts	<i>mus</i> vs <i>dom</i>	1536	1774	38	139	46	96
	<i>mus</i> $\times$ <i>dom</i> vs	42	19	85	36	5	1

	<i>dom</i> × <i>mus</i>						
--	-------------------------	--	--	--	--	--	--

1398

1399

**Table 3: Number of differentially expressed genes in each WGCNA module.** Rows

1400

indicate WGCNA modules and columns indicate comparisons between cross types

1401

used to identify differentially expressed (DE) genes. Module associations with cross

1402

types are based on linear models with post-hoc Tukey tests. Shaded boxes indicate a

1403

significant enrichment for DE genes based on a hypergeometric test with FDR-

1404

correction ( $P < 0.05$ ). Note that there is not necessarily a relationship between Hybrid

1405

F1 XY Match and Non-hybrid XY Mismatch modules with the same module number.

	Module	Significant cross type associations	Number of DE genes in module			
			<i>mus</i> × <i>dom</i> vs <i>mus</i> × <i>dom</i> <sup><i>musY</i></sup>	<i>mus</i> × <i>dom</i> vs <i>dom</i> × <i>mus</i> <sup><i>domY</i></sup>	<i>dom</i> × <i>mus</i> vs <i>mus</i> × <i>dom</i> <sup><i>musY</i></sup>	<i>dom</i> × <i>mus</i> vs <i>dom</i> × <i>mus</i> <sup><i>domY</i></sup>
Hybrid F1 XY Match	1	none	0	13	9	0
	2	none	1	19	35	2
	3	none	1	10	170	1
	4	none	4	3	11	2
	5	none	7	21	155	5
	6	<i>mus</i> × <i>dom</i> background	2	87	102	1
Non-hybrid XY Mismatch	1	<i>mus</i> background	0	1039	972	40
	2	none	0	168	133	4
	3	<i>dom</i> background	9	913	970	4
	4	none	4	91	358	2
	5	none	23	532	28	9
	6	none	3	329	17	5
	7	none	1	220	55	1
	8	none	1	22	77	8
	9	none	0	106	3	0
	10	none	1	1	104	1
	11	none	0	111	24	0
	12	none	1	3	6	0
	13	none	0	1	0	0
	14	none	1	1	12	1
	15	none	0	0	0	0
	16	none	0	0	5	0
	17	none	0	18	0	1
	18	none	0	0	0	0

1406

1407 **Figure Legends**

1408

1409 **Figure 1:** Experimental design. (A) Backcrosses used to generate Y-introgression  
 1410 mouse strains. We performed 10 generations of backcrosses in reciprocal directions to  
 1411 generate mice with a *Mus musculus domesticus* (*domesticus*) genetic background and  
 1412 *Mus musculus musculus* (*musculus*) Y chromosome (*domesticus*<sup>musY</sup>) and mice with a  
 1413 *musculus* genetic background and *domesticus* Y chromosome (*musculus*<sup>domY</sup>). The thin  
 1414 horizontal line on the autosomes represents residual autosomal introgression, which is  
 1415 theoretically expected to represent about 0.1% of the autosomes. (B) Crosses were  
 1416 performed with Y-introgression mice to produce two types of experimental F1 mice. For  
 1417 Hybrid F1 XY Match, we crossed Y-introgression males to females from the other  
 1418 subspecies to generate F1 mice with hybrid autosomes but matched sex chromosomes.  
 1419 For Non-hybrid XY Mismatch, we crossed Y-introgression males to females from a  
 1420 different strain but the same subspecies to generate F1 mice with X-Y mismatch and  
 1421 non-hybrid autosomes. Autos = autosomes, X = X chromosome, Y = Y chromosome.

1422

1423 **Figure 2:** Copy number estimates for ampliconic gene families in wild mice, wild-  
 1424 derived inbred strains, and Y-introgression strains. Copy number was estimated using a  
 1425 97% identity cutoff for paralogs. (A-D) show copy numbers in male mice, with Y  
 1426 chromosome genes on the y-axis and their X chromosome homologs on the x-axis. (E)  
 1427 includes both males and females and shows haploid copy number for the autosomal  
 1428 gene family *α-takusan* on the y-axis and haploid copy number for the X-linked family *S/x*  
 1429 on the x-axis. Note that (A) and (B) show the same information on the y-axis and (C)  
 1430 and (D) show the same information on the x-axis to compare copy numbers for  
 1431 ampliconic gene families that have two different homologous gene families on the  
 1432 opposite sex chromosome. Correlations and p-values are based on a Pearson's  
 1433 correlation test. P-values were FDR-corrected for multiple tests.

1434

1435 **Figure 3:** (A) Relative testes mass (mg/g), (B) sperm nucleus bounding width (μm), and  
 1436 (C) sperm nucleus bounding height (μm) by cross type. Letters above each violin plot  
 1437 indicate significant differences (FDR-corrected P < 0.05) based on a Welch's t-test  
 1438 (relative testes mass) or Wilcoxon rank-sum test (bounding width and height). Sample

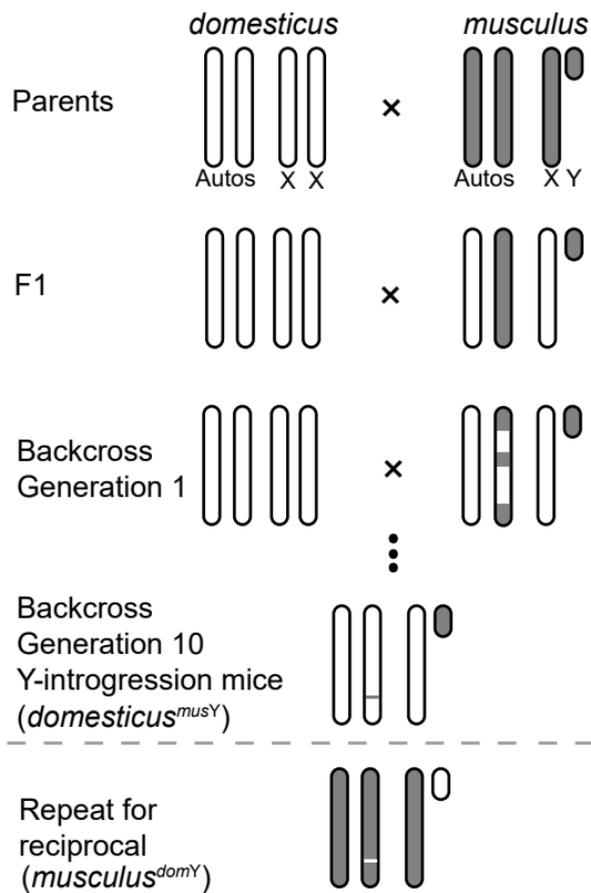
1439 size for each cross type is indicated below each violin plot. Bounding width and height  
1440 sample sizes indicate the number of sperm nuclei observed. Representative sperm  
1441 nuclei morphologies for each cross type are depicted above each violin plot in (B).

1442  
1443 **Figure 4:** Normalized expression levels of *Slx* (A), *Slx11* (B), *Sly* (C), *Ssty1* (D), *Ssty2*  
1444 (E), and  *$\alpha$ -takusan* (F) ampliconic gene families in different cross types plotted against  
1445 their copy numbers. Copy number estimates are based on estimates from wild-derived  
1446 strains used in experimental and control crosses (see Figure 2). Cross types with the  
1447 same sex chromosome and therefore same copy number estimate are jittered slightly  
1448 along the x-axis for clarity. Expression level was calculated by summing transcripts-per  
1449 million (TPM) for each paralog of the gene family with at least 97% sequence identity to  
1450 the ampliconic gene. Points represent values for individual samples, and lines indicate  
1451 median and standard deviation for each cross type.

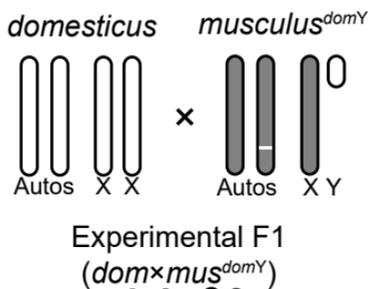
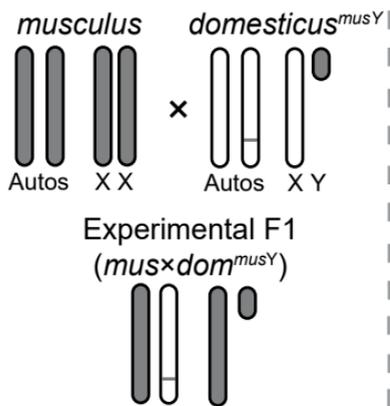
1452  
1453 **Figure 5:** Histograms of relative expression levels between experimental cross types  
1454 and control mice. (A-C) Contrasts that all have a *musculus* X chromosome, (D-F)  
1455 contrasts with a *domesticus* Y chromosome (G-I) contrasts with a *musculus* Y  
1456 chromosome, and (J-L) contrasts with a *domesticus* X chromosome. (A-F) represent  
1457 sex chromosome mismatch present in sterile hybrids (*musculus* X and *domesticus* Y),  
1458 while (G-L) represent sex chromosome mismatch present in more fertile hybrids  
1459 (*domesticus* X and *musculus* Y). The first column (A, D, G, and J) shows data  
1460 reanalyzed from (Larson *et al.* 2017). The second column (B, E, H, K) tests if gene  
1461 expression levels are rescued when the sex chromosomes are matched but on a hybrid  
1462 autosomal background (Hybrid F1 XY Match). The third column (C, F, I, L) tests for  
1463 disrupted expression due to sex chromosome mismatch alone, on a non-hybrid  
1464 autosomal background (Non-hybrid XY Mismatch). The y-axis shows the difference in  
1465 normalized expression levels between the two cross types being compared. The x-axis  
1466 shows the proportion of genes in each expression difference bin. Black bars represent  
1467 the autosomes, purple bars represent the X chromosome, and green bars represent the  
1468 Y chromosome. Letters indicate significant differences in median expression differences  
1469 among the chromosome types based on a Mann-Whitney U test (FDR-corrected  $P <$   
1470 0.05).

1471  
1472 **Figure 6:** Upset plots showing the number of differentially expressed (DE) genes in  
1473 each cross type comparison, and genes that are DE across multiple comparisons. (A)  
1474 DE genes on the X chromosome. (B) DE genes on the Y chromosome. Bars  
1475 corresponding to multiple dots connected by lines indicate genes that are DE across  
1476 multiple comparisons. Bars corresponding to single dots indicate genes that are DE in  
1477 only one comparison. Blue dots indicate comparisons on the *domesticus* X  
1478 chromosome (A) or *domesticus* Y chromosome (B), and red dots indicate comparisons  
1479 on the *musculus* X chromosome (A) or *musculus* Y chromosome (B). Genes that were  
1480 DE in opposite directions across multiple comparisons of the same sex chromosome  
1481 were excluded.

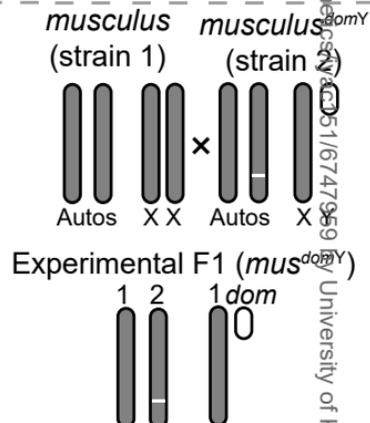
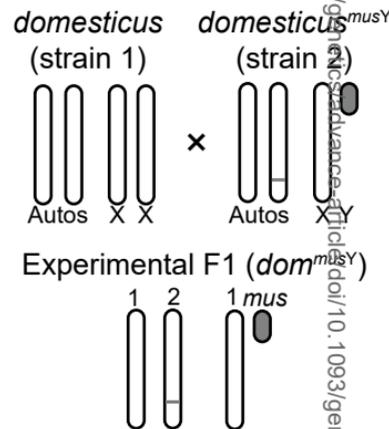
1482  
1483 **Figure 7:** Example WGCNA module eigengene values plotted by cross type. Note that  
1484 WGCNA was performed separately for each experiment, so there is not necessarily a  
1485 relationship between Hybrid F1 XY Match and Non-hybrid XY Mismatch modules with  
1486 the same number. Modules that were significantly associated with cross types are also  
1487 labeled based on these associations (A, B, and D). Other modules shown were not  
1488 significantly associated with a cross type but trended towards an association with X-  
1489 autosomal background by Y chromosome type interaction and were enriched for DE  
1490 genes in at least one comparison (C, E, and F; Table 3). Letters indicate significant  
1491 differences in module association based on linear models with post-hoc Tukey tests ( $P$   
1492  $< 0.05$ ).

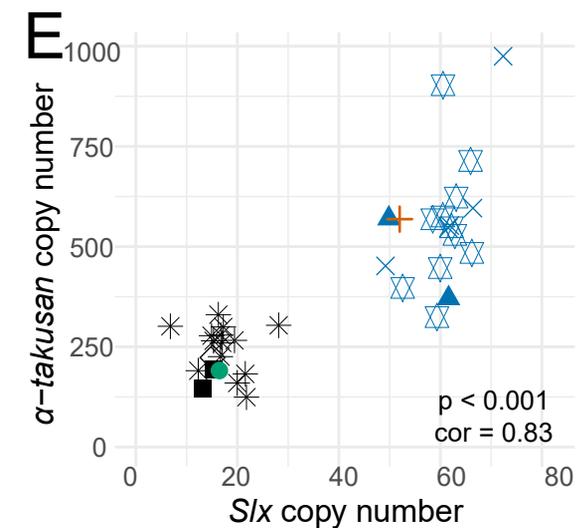
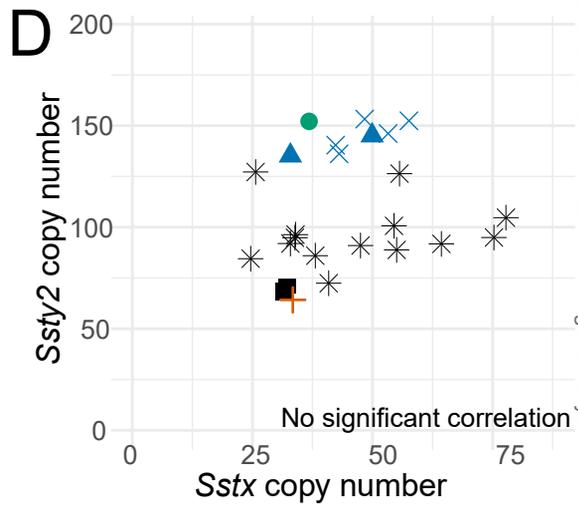
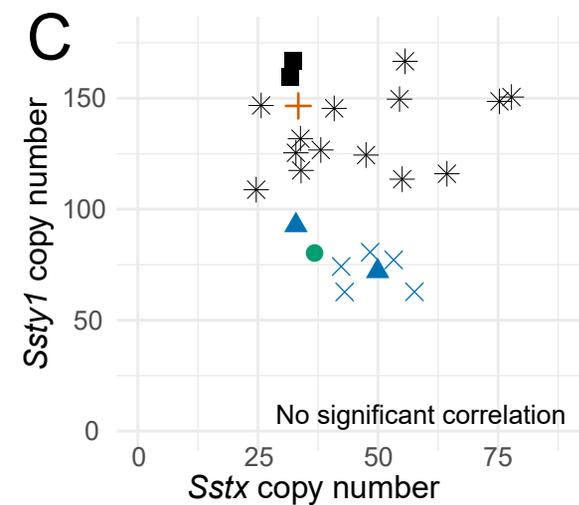
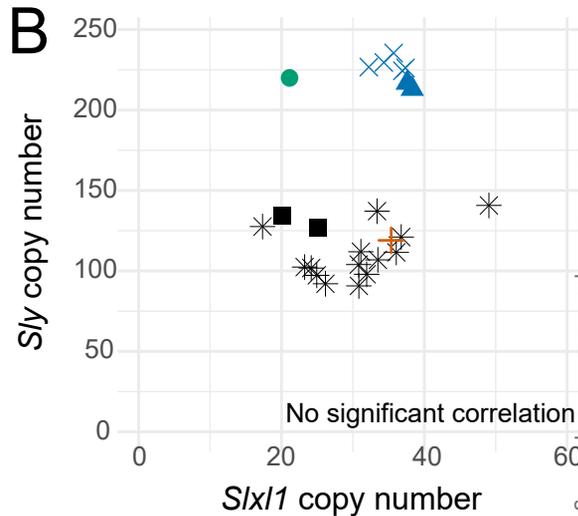
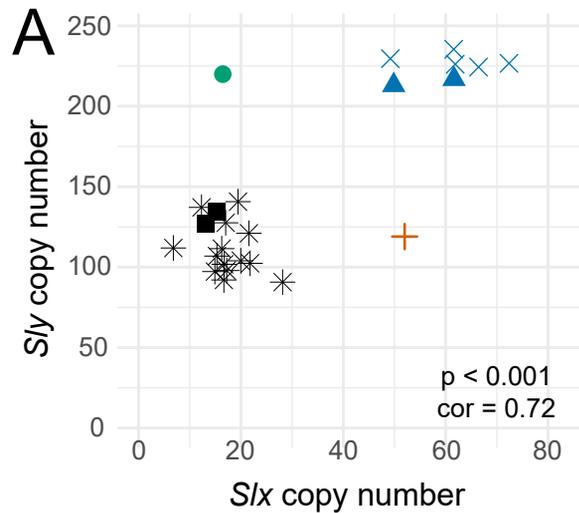
**A****Generation of Y-introgression Strains****B****Hybrid F1 XY Match:**

X-Y match on F1 hybrid autosomal background

**Non-hybrid XY Mismatch:**

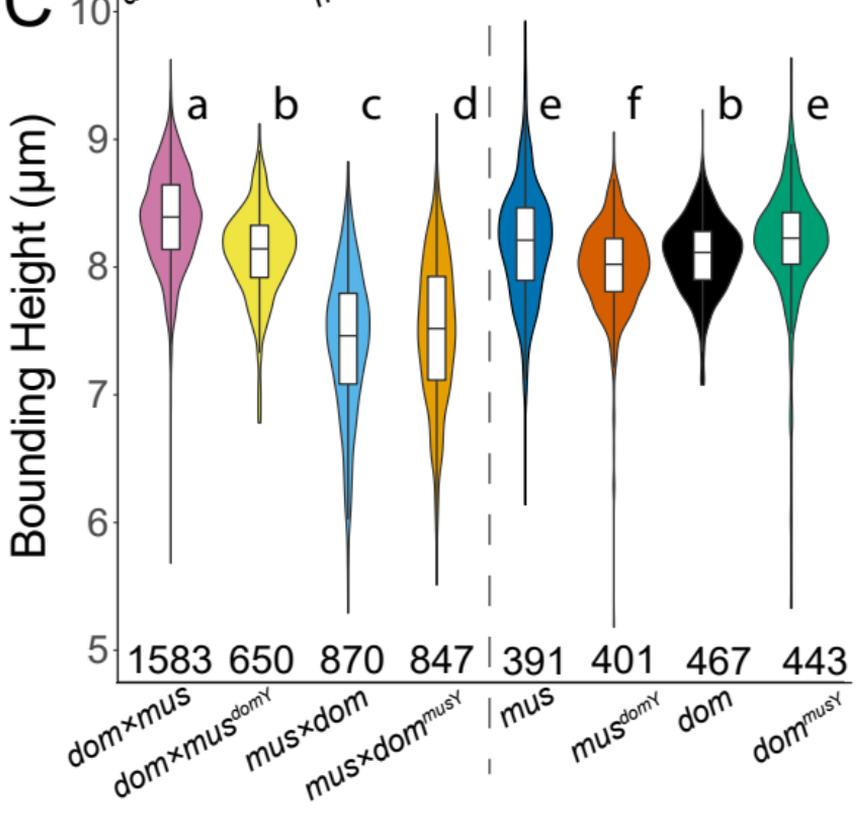
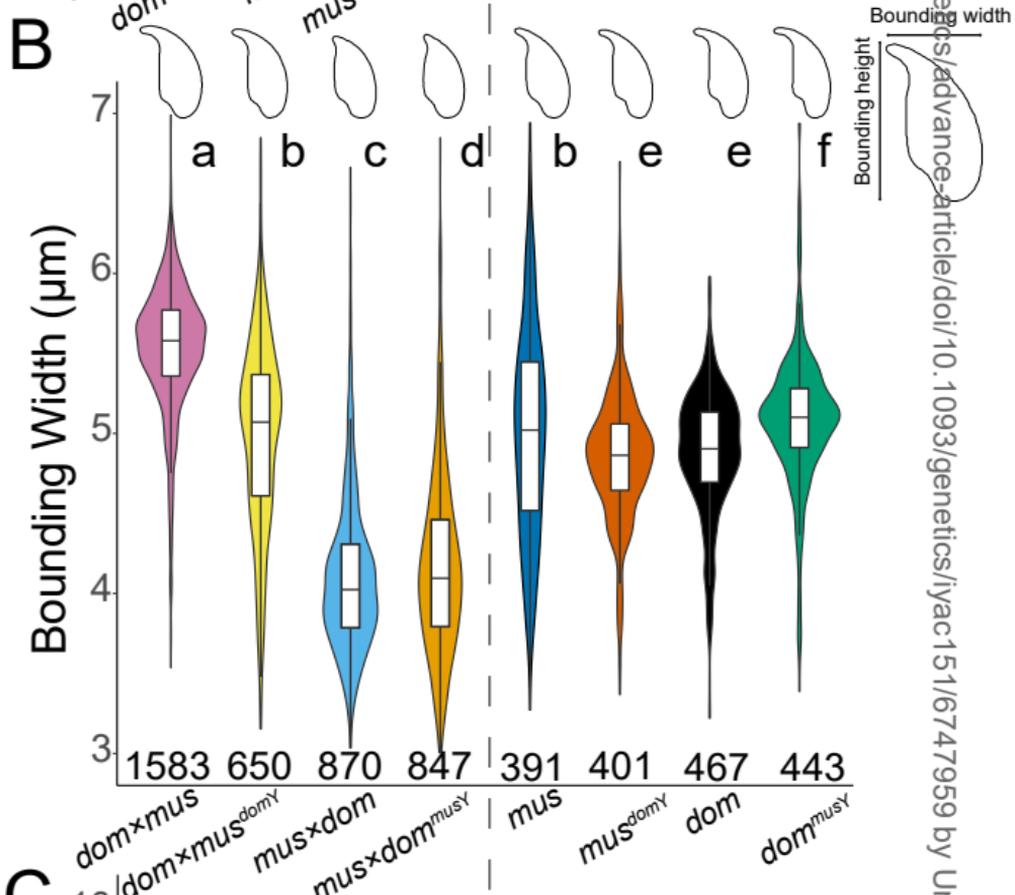
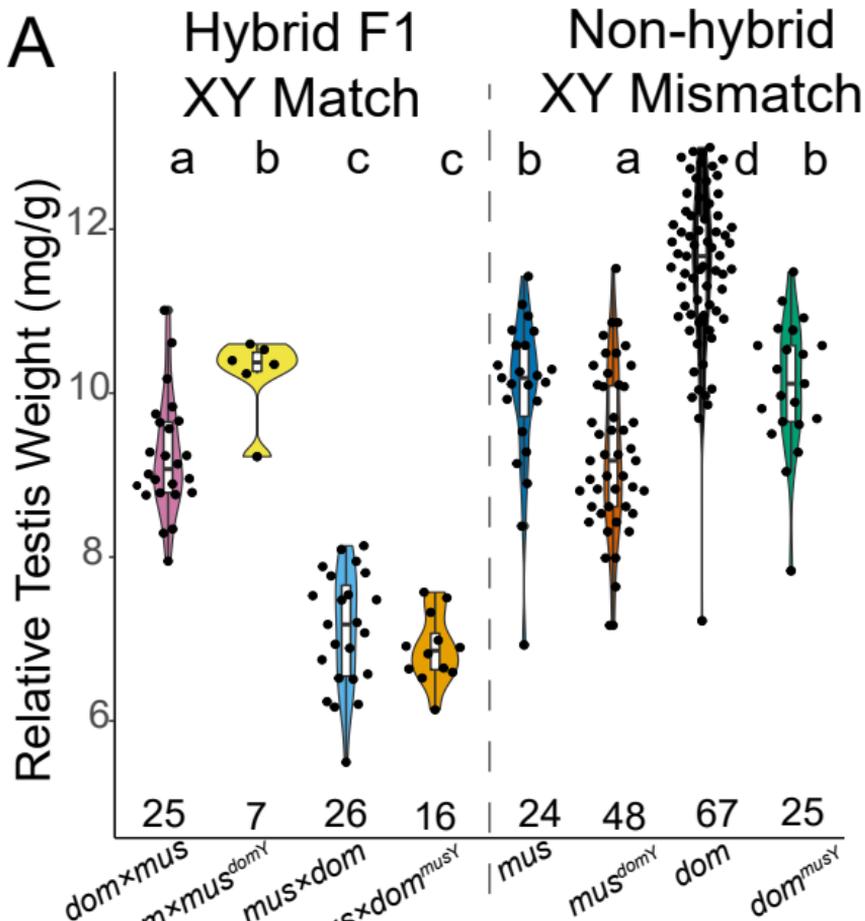
X-Y mismatch on non-hybrid autosomal background

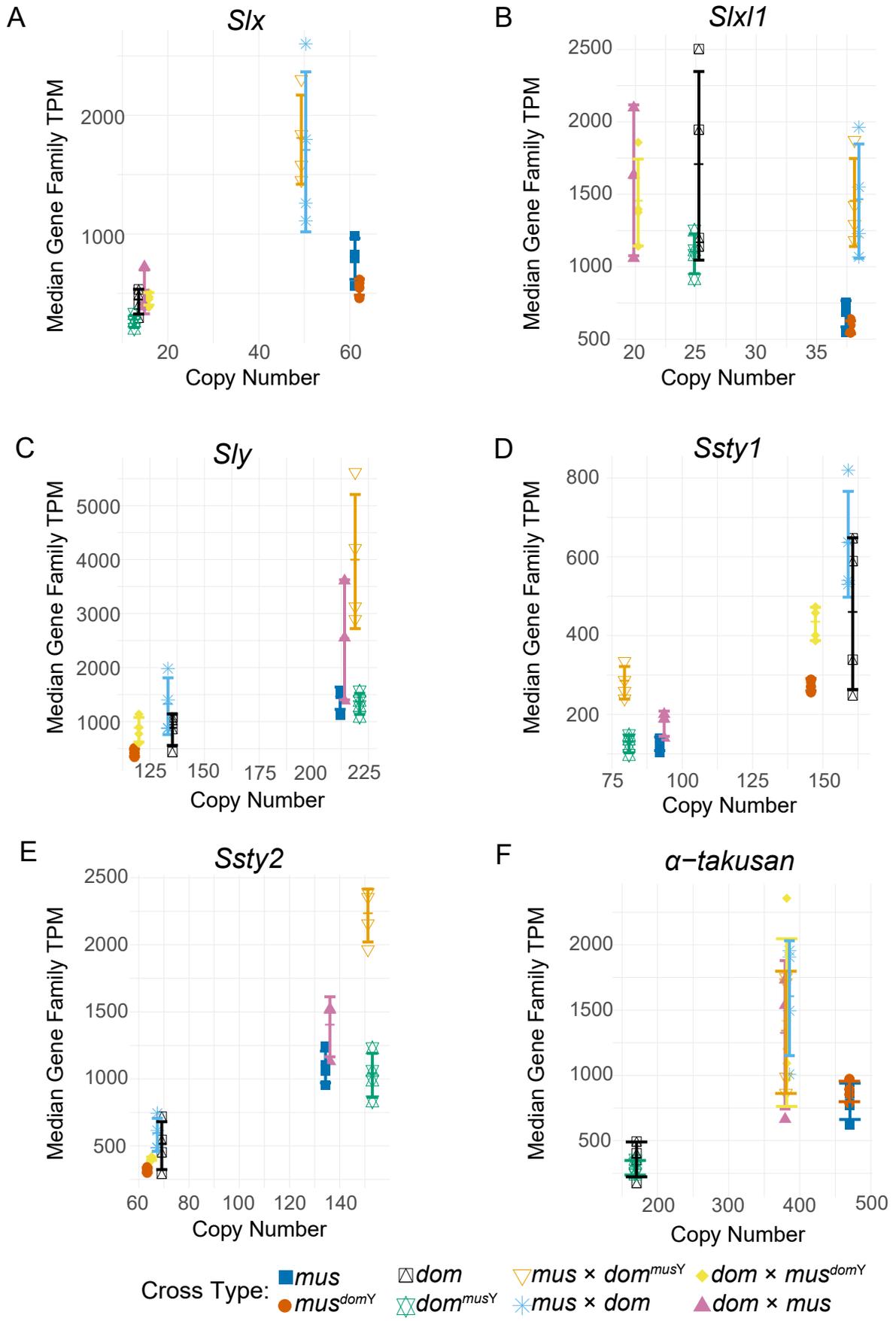




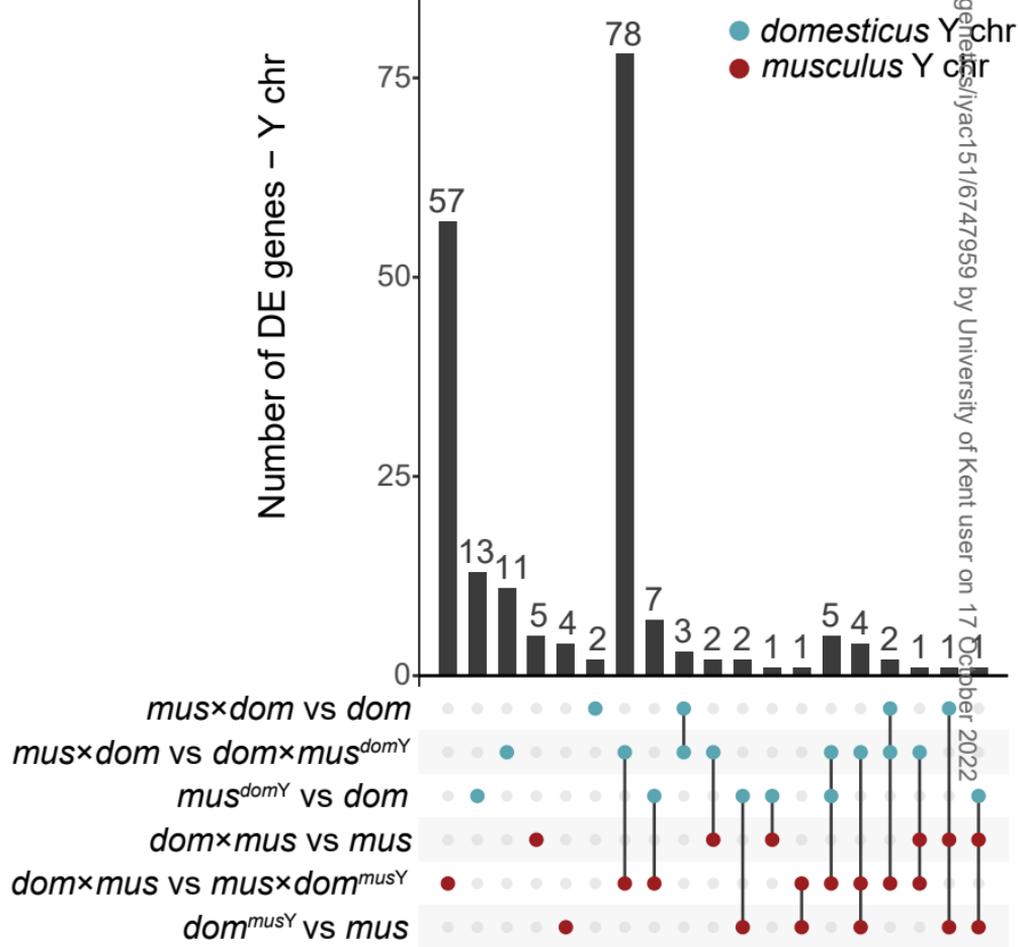
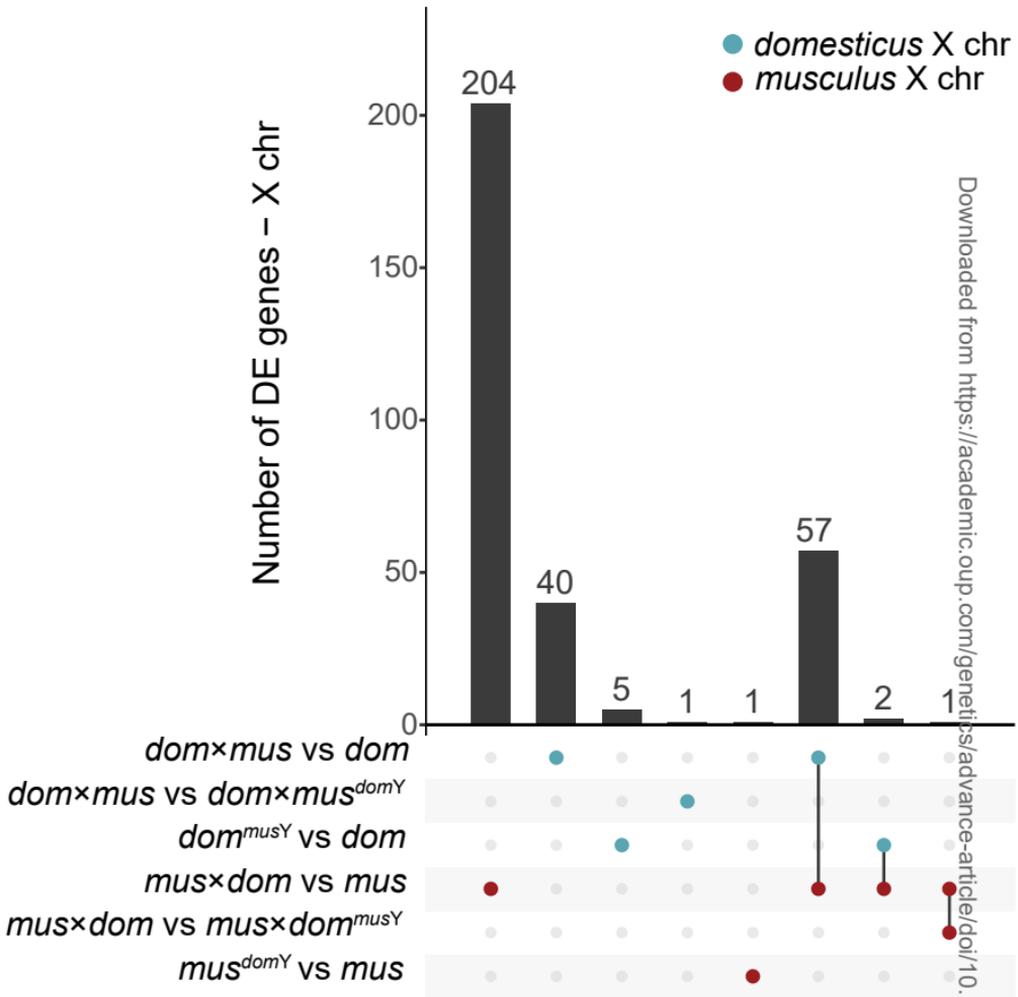
Subspecies/cross type, source, and sex

- \* *domesticus*, natural population, M
- ◇ *domesticus*, natural population, F
- *domesticus*, wild-derived strain, M
- *domesticus*<sup>mus<sup>Y</sup></sup>, Y-introgression strain, M
- × *musculus*, natural population, M
- ◇ *musculus*, natural population, F
- ▲ *musculus*, wild-derived strain, M
- + *musculus*<sup>dom<sup>Y</sup></sup>, Y-introgression strain, M









# Hybrid F1 XY Match

# Non-hybrid XY Mismatch

