The Evolution of Polymorphic Hybrid Incompatibilities in House Mice

Erica L. Larson,*,† Dan Vanderpool,* Brice A. J. Sarver,* Colin Callahan,*,[1] Sara Keeble,*,‡ Lorraine L. Provencio,‡ Michael D. Kessler,‡ Vanessa Stewart,* Erin Nordquist,* Matthew D. Dean,‡ and Jeffrey M. Good*,[2]

*Division of Biological Sciences, University of Montana, Missoula, Montana 59812, †Department of Biological Sciences, University of Denver, Denver, CO 80210 and ‡Molecular and Computational Biology Section, Department of Biological Sciences, University of Southern California, Los Angeles, California 90089

ORCID IDs: 0000-0003-3006-645X (E.L.L.); 0000-0002-7318-3122 (S.K.); 0000-0003-0707-5374 (J.M.G.)

ABSTRACT Resolving the mechanistic and genetic bases of reproductive barriers between species is essential to understanding the evolutionary forces that shape speciation. Intrinsic hybrid incompatibilities are often treated as fixed between species, yet there can be considerable variation in the strength of reproductive isolation between populations. The extent and causes of this variation remain poorly understood in most systems. We investigated the genetic basis of variable hybrid male sterility (HMS) between two recently diverged subspecies of house mice, *Mus musculus domesticus* and *Mus musculus musculus*. We found that polymorphic HMS has a surprisingly complex genetic basis, with contributions from at least five autosomal loci segregating between two closely related wild-derived strains of *M. m. musculus*. One of the HMS-linked regions on chromosome 4 also showed extensive introgression among inbred laboratory strains and transmission ratio distortion (TRD) in hybrid crosses. Using additional crosses and whole genome sequencing of sperm pools, we showed that TRD was limited to hybrid crosses and was not due to differences in sperm motility between *M. m. musculus* strains. Based on these results, we argue that TRD likely reflects additional incompatibilities that reduce hybrid embryonic viability. In some common inbred strains of mice, selection against deleterious interactions appears to have unexpectedly driven introgression at loci involved in epistatic hybrid incompatibilities. The highly variable genetic basis to F1 hybrid incompatibilities between closely related mouse lineages argues that a thorough dissection of reproductive isolation will require much more extensive sampling of natural variation than has been commonly utilized in mice and other model systems.

KEYWORDS polymorphism; hybrid male sterility; transmission ratio distortion; QTL mapping; introgression

THE evolution of intrinsic hybrid incompatibilities, whereby divergent genomic regions interact negatively in hybrid genomes, is one of the most commonly studied models of speciation [*i.e.*, Bateson–Dobzhansky–Muller incompatibilities or BDMIs, Bateson 1909; Dobzhansky 1937; Muller 1942; reviewed in Maheshwari and Barbash 2011]. Although often viewed as fixed epistatic barriers to gene flow between species, many incompatible alleles are polymorphic within populations, leading to variation in the overall strength of reproductive isolation between populations (Gordon 1927; Patterson and Stone 1952; Forejt and Ivanyi 1974; Reed and Markow 2004; Good *et al.* 2008b; Scopece *et al.* 2010; Cutter 2012). Relatively few studies have examined the genetic basis of variation in BDMIs (Christie and Macnair 1987; Vyskočilová *et al.* 2005; Wright *et al.* 2013; Sweigart and Flagel 2015; Case *et al.* 2016) and the evolutionary forces underlying variable incompatibilities remain unexplored in most species.

Hybrid incompatible alleles can arise through any evolutionary process that contributes to genetic divergence (*e.g.*, genetic drift, natural or sexual selection). However, reproductive isolation is expected to evolve more quickly when divergence is driven by selection. Consistent with this, several incompatibility

doi: https://doi.org/10.1534/genetics.118.300840
Manuscript received February 20, 2018; accepted for publication April 23, 2018; published Early Online April 24, 2018.
Supplemental material available at Figshare: https://doi.org/10.25386/genetics.6149399.
[1]Present address: Center for Epigenetics, Johns Hopkins University School of Medicine, Baltimore, MD 21205.
[2]Corresponding author: Division of Biological Sciences, University of Montana, 32 Campus Drive, HS104, Missoula, MT 59812. E-mail: jeffrey.good@umontana.edu

genes show signatures of positive selection or have diverged through antagonistic coevolutionary dynamics (Johnson 2010; Presgraves 2010; Maheshwari and Barbash 2011). Polymorphism is an inevitable phase in the fixation of an allele, but directional selection should fix alleles relatively quickly. Thus, it should be rare that incompatibilities are sampled while polymorphic if positive directional selection drives the evolution of BDMIs. Alternatively, BDMIs could involve a combination of unsorted ancestral variation or modifying loci that segregate neutrally within species (Rieseberg and Blackman 2010; Scopece *et al.* 2010; Cutter 2012; Matute *et al.* 2014) or are subject to balancing selection (Cutter 2012). Finally, polymorphic incompatibilities may reflect the breakdown of reproductive barriers due to gene flow between partially isolated populations. Hybrid incompatible alleles are generally assumed to be resistant to introgression (Barton and Hewitt 1985; Harrison 1990; Payseur 2010), but epistatic barriers may quickly erode in the face of gene flow (Bank *et al.* 2012; Lindtke and Buerkle 2015). Differentiating between these alternatives is crucial to understanding the evolution of reproductive isolation and the nature of species boundaries.

House mice provide a powerful system to understand the causes of polymorphic barriers during the early stages of speciation. There are three major lineages within *Mus musculus*—*M. m. musculus*, *M. m. domesticus*, and *M. m. castaneus*—that diverged ~0.35–0.50 MYA (Geraldes *et al.* 2011) and show partial reproductive isolation primarily due to hybrid male sterility (HMS). However, there appears to be considerable standing genetic variation for the strength of HMS (Britton-Davidian *et al.* 2005; Vyskočilová *et al.* 2005; Good *et al.* 2008b; Turner *et al.* 2012). For example, crosses between *M. m. musculus* females and *M. m. domesticus* males typically yield sterile F1 hybrid males due, in part, to negative interactions between *M. m. musculus* Chr X and the autosomal gene *Prdm9*, a DNA binding protein that directs the location of double-strand breaks during recombination (Mihola *et al.* 2009). PRDM9 binding sites evolve rapidly (Baker *et al.* 2015), leading to asymmetric binding and autosomal asynapsis that disrupts sex chromosome expression during spermatogenesis (Bhattacharyya *et al.* 2013; Campbell *et al.* 2013; Turner *et al.* 2014; Davies *et al.* 2016; Larson *et al.* 2017). *Prdm9* appears to be polymorphic for sterile and fertile alleles within both *M. m. domesticus* and *M. m. musculus* (Forejt and Ivanyi 1974; Vyskocilová *et al.* 2009; Flachs *et al.* 2012), and the strength of *Prdm9*-associated sterility is variable within *M. m. musculus* (Bhattacharyya *et al.* 2014; Flachs *et al.* 2014; Turner *et al.* 2014).

There is also variation in the severity of HMS in house mice that is independent of the *M. m. musculus* X (Good *et al.* 2008b). For example, crosses between *M. m. domesticus* females and *M. m. musculus* males produce sterile or fertile hybrid males dependent on the genotype of the *M. m. musculus* sire (Vyskočilová *et al.* 2005; Good *et al.* 2008b; Bhattacharyya *et al.* 2014; Flachs *et al.* 2014). The autosomal variants contributing to HMS in these crosses are unresolved and, aside from the rapid evolution of *Prdm9* (Davies *et al.* 2016), the causes of standing variation for HMS are not clear. One possible factor is that the small effective population sizes of house mice results in strong genetic drift and local inbreeding (Geraldes *et al.* 2011). Further, *M. m. domesticus* and *M. m. musculus* form a narrow hybrid zone in central Europe (Janoušek *et al.* 2012), which may weaken reproductive barriers through introgression (Turner and Harr 2014).

In this study, we used the genetic variation segregating between two wild-derived inbred strains of *M. m. musculus* to begin to characterize the genetic architecture of polymorphic barriers between *M. m. domesticus* females and *M. m. musculus* males. We found that polymorphic HMS encompasses at least five autosomal regions of the genome. We then used additional genetic crosses, whole genome sequencing of sperm pools, and population genomic analyses to explore the mechanistic and evolutionary drivers contributing to variation at one of these regions on the distal portion of Chr 4. These diverse genetic and genomic experiments further reveal the complex genetic basis of reproductive isolation in this system and demonstrate how these reproductive barriers have shaped introgression among mouse subspecies and the genomic composition of common laboratory strains of mice.

Materials and Methods

Mouse strains and experimental crosses

We focused on two wild-derived inbred strains of *M. m. musculus* (PWK/PhJ and CZECHII/EiJ, hereafter *musculus*^{PWK} and *musculus*^{CZII}) that differ in the degree of HMS when crossed to *M. m. domesticus* (Good *et al.* 2008b). The *musculus*^{PWK} strain was originally isolated near the hybrid zone in Prague, Czechia (50.0216°N, 14.4350°E) and yields weak HMS when crossed to female *M. m. domesticus*. The *musculus*^{CZII} was isolated further from the hybrid zone in Bratislava, Slovakia (48.1492°N, 17.1070°E) and produces mostly sterile males when crossed to female *M. m. domesticus*. We used two wild-derived strains of *M. m. domesticus* (WSB/PhJ and LEWES/PhJ, hereafter *domesticus*^{WSB} and *domesticus*^{LEW}) derived from natural populations in North America (MD, 39.3358°N, 77.3282°W and DE, 39.1453°N, 75.4188°N). Mice were originally purchased from Jackson Laboratory (Bar Harbor, ME). All animal use was approved by the University of Montana (protocol 002–13) and the University of Southern California (protocol 11,394) Institutes for Animal Care and Use Committees.

Hybrid incompatibilities underlying F1 hybrid phenotypes usually cannot be mapped because of a lack of genetic variation. However, the existence of polymorphic sterility factors within *musculus* provides an elegant way to resolve these incompatibilities directly in F1 hybrid males. We first quantified HMS in F1 crosses between female *domesticus*^{WSB} and either *musculus*^{PWK} or *musculus*^{CZII} males. To control for the effects of inbreeding depression in inbred strains, we compared the fertility of these F1 hybrids to *M. m. domesticus* interstrain F1 males (*domesticus*^{WSB} × *domesticus*^{LEW}),

evaluating each cross at different time points [60–65, 70–80, and 85–95 days postpartum (dpp)]. We then used an F1 hybrid test cross between *domesticus*^{WSB} females and interstrain F1 males from reciprocal crosses of *musculus*^{PWK} and *musculus*^{CZII} (Figure 1). This design maintains the F1 hybrid genotype, while segregating variation between two different *M. m. musculus* genomes. Finally, we backcrossed *M. m. musculus* interstrain F1 males to *musculus*^{PWK} females to determine if loci on Chr 4 showing strong transmission ratio distortion (TRD) in our hybrid crosses also showed TRD within *M. m. musculus*.

Male reproductive phenotypes

We quantified reproductive phenotypes of virgin males weaned in same-sex sibling groups at 21 dpp and housed singly at 45 dpp to mitigate dominance interactions (Snyder 1967). Males were killed using carbon dioxide followed by cervical dislocation at 58–70 dpp (F1 hybrid test cross) or up to 90 dpp (aged F1 males). Following Good *et al.* (2008b), we measured paired testes (an overall measure of fertility) and seminal vesicles (correlated with serological testosterone levels) relative to body weight. We isolated sperm from caudal epididymides diced in 1 ml of Dulbecco's PBS (Sigma, St. Louis, MO) and incubated at 37° for 10 m. The proportion of motile sperm and total sperm numbers were estimated from 5 μl suspensions (regular and heat-shocked, respectively) viewed in a Makler counting chamber on a light microscope over a fixed area and observation time. To evaluate sperm morphology, 25 μl sperm suspensions were fixed and stained, and ≥100 intact sperm were visually classified by a single individual (E.L.L.) while blind to genotype. Sperm head morphology were (1) normal with a long apical hook, (2) slightly abnormal (*i.e.*, shortened hook), (3) abnormal (*i.e.*, short hook and rounded shape), or (4) severely abnormal (*i.e.*, amorphous shape). We summarized these categories with a weighted index that ranged from high (3) to low (0) quality sperm (Oka *et al.* 2004; Good *et al.* 2008a). Sperm tail morphology were (1) normal, (2) bent at the base of the sperm head, (3) bent in the center of the tail forming a loop, or (4) twisted distally (White *et al.* 2011).

Genotyping and genome sequencing

We genotyped 468 individuals from our two genetic mapping experiments (*i.e.*, 156 F1 hybrid test cross males and 312 *M. m. musculus* backcross) and eight reference samples (two of each parent strain and *domesticus*^{WSB} × *musculus*^{CZII} F1 hybrids) using double-digest restriction site–associated DNA sequencing (ddRADseq; Peterson *et al.* 2012), with minor modifications. DNA was extracted from liver tissue using the NucleoSpin Tissue kit (Machery-Nagel, Düren, Germany) and incubated in 5 μl RNAase A (Fisher Scientific, Waltham, MA) at 37° for 15 m. We digested 1 μg of DNA with *Msp*I and *Sbf*I-High-Fidelity enzymes (New England Biolabs, Beverly, MA), ligated unique adaptors, and selected 200–500 bp fragments using a two-step size selection with AMPure XP beads (Agencourt Bioscience, Beverly, MA). Individual libraries were amplified 16 cycles in three 20 μl reactions using Phusion High-Fidelity DNA Polymerase (New England BioLabs), cleaned using AMPureXP, and quantified with a NanoPhotometer (IMPLEN, München, Germany). F1 hybrid libraries were paired-end sequenced on an Illumina HiSeq 2000 at the California Institute for Quantitative Biosciences, University of California, Berkley and on a MiSeq at the IBEST Genomic Resources Core, University of Idaho. The *M. m. musculus* backcross libraries were single-end sequenced on an Illumina HiSeq 4000 at the University of Oregon Genomics and Cell Characterization Core Facility.

Libraries were checked for intact barcodes, restriction enzyme cut-sites, and demultiplexed using *preprocess_radtag_lane.py* (Peterson *et al.* 2012). We used Trimmomatic v0.32 (Lohse *et al.* 2012) to remove adaptor sequences and low-quality bases and mapped reads to the Genome Reference Consortium mouse build 38 (GRCm38) using BWA-MEM v0.7.10 (Li 2013). We applied the Genome Analysis Tool Kit (GATK) v3.4 (McKenna *et al.* 2010) to call SNPs (HaplotypeCaller) that we then filtered (minDP 10, maxDP 150, minGQ 20) using VCFtools v0.1.14 (Danecek *et al.* 2011). We retained biallelic SNPs that were homozygous in the parent references, heterozygous in the F1 hybrid reference, genotyped in ≥95% of individuals, and >1000 bp from other SNPs. We retained individuals that were genotyped in ≥85% of markers and showed normal crossover rates.

We also performed two targeted genotyping assays. Males from the F1 hybrid test cross were genotyped for microsatellite length variants that encompass different *Prdm9* alleles. We used modified versions of D17Mit78 (forward: CACAGTGAGTCTGGGCTAGTC, reverse: GCATCTTATGGATTGAAATACGG) and D17Mit261 (forward: CCCTTGCTCTCCTTCATTCA, reverse: AATGCCAAATGGTCAGCC; Copeland *et al.* 1993) in 10 μl PCR reactions using MangoTaq (Bioline, Luckenwalde, Germany), run at 35 cycles of 94° for 30 sec, 58–48° for 30 sec (decreased by 1° per cycle for the first 10 cycles), and 72° for 1 m. We also expanded the genotyping of our *M. m. musculus* backcross using diagnostic microsatellites from the middle (D4Mit64: 140.08–141.03 bp) and distal end (D4Mit127: 148.60–151.60 bp) of Chr 4 (Copeland *et al.* 1993). These markers spanned SNPs genotyped using ddRADseq, and we genotyped 88 mice using both methods to allow cross validation. All fragments were analyzed on an ABI 3130xl at the University of Montana Genomics Core.

Reference genomes have been published for *musculus*^{PWK} and *domesticus*^{WSB} (Keane *et al.* 2011). We generated whole genome shotgun sequences of a female *musculus*^{CZII} and a female *domesticus*^{LEW} from NEXTflex DNA sequencing genomic libraries (Bio Scientific, Austin, TX) that were paired-end sequenced on an Illumina HiSeq 2000 at the California Institute for Quantitative Biosciences University of California, Berkley. We generated additional whole genome sequence data from the same *musculus*^{CZII} female at GENEWIZ (South Plainfield, NJ) using Illumina TruSeq libraries, paired-end sequenced on an Illumina HiSeq 2500. These data were processed as described above.

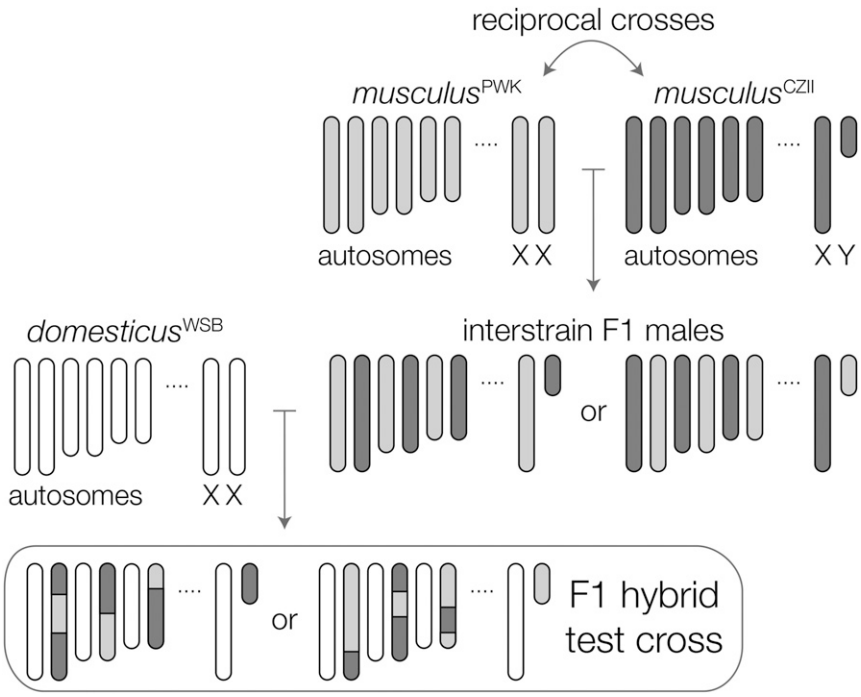

Figure 1 Crossing design to map polymorphic HMS loci in *musculus*. The fertility of F1 hybrids from crosses between *M. m. domesticus* females and *M. m. musculus* males depends on the strain of *M. m. musculus*; hybrids with *musculus*^{CZII} sires have more severe sterility. We test crossed interstrain *M. m. musculus* F1s to *M. m. domesticus* females to map F1 hybrid sterility alleles segregating within *M. m. musculus*.

QTL mapping

We performed QTL mapping in R/qtl v1.40-8 (Broman *et al.* 2003) with an assumed genotyping error rate of 0.001 and the Carter–Falconer mapping function (Carter and Falconer 1951). For all QTL analyses we used a grid size of 1 cM and 5% genome-wide significance thresholds estimated from 1000 permutations. We used standard interval QTL interval mapping (*scanone*) with a Haley–Knott regression for normally distributed traits, and nonparametric interval mapping for the proportion of motile sperm, and sperm head and tail morphologies. We used two-dimensional QTL mapping (*scantwo*) and multiple QTL model selection (*stepwiseqtl*) to identify additional QTL that may be involved in epistatic interactions. Multiple QTL models were compared using penalized LOD score with thresholds calculated from *scantwo* permutations. We tested for TRD in our cross using a χ^2 test of Mendelian proportions.

Genomic analyses

To investigate the evolutionary history of genomic regions associated with polymorphic sterility, we first analyzed the newly sequenced *musculus*^{CZII} and *domesticus*^{LEW} genomes and published genomic data from GRCm38 (*domesticus*^{C57}), *domesticus*^{WSB}, *musculus*^{PWK}, and *Mus spretus* SPRET/EiJ (Keane *et al.* 2011). For each genome, we called SNPs using the GATK (HaplotypeCaller) and filtered SNPs with VCFtools (minDP 10, maxDP 150, minGQ 30). We generated a BED file with all SNP positions and used the GATK to re-call genotypes in each genome for our target SNPs (HaplotypeCaller) and filter our VCF files (SelectVariants, minGQ 30, biallelic). We then tested for introgression using the four-taxon *D*-statistic (Green *et al.* 2010; Durand *et al.* 2011) as implemented in dfoil (Pease and Hahn 2015). Here, the *D*-statistic is the normalized difference in site pattern counts that support a closer relationship between *musculus*^{CZII} and a focal *M. m. domesticus* (ABBA, negative *D*-statistic) or *musculus*^{PWK} and a focal *M. m. domesticus* (BABA, positive *D*-statistic) with variants polarized using *M. spretus*. We calculated the *D*-statistic per chromosome and for nonoverlapping 100 kb and 1 Mb windows. We repeated these analyses using three different strains of *M. m. domesticus* (*domesticus*^{C57}, *domesticus*^{WSB}, and *domesticus*^{LEW}).

Second, we used published genotype data for classic laboratory strains, wild-derived strains (Yang *et al.* 2011), and wild populations of house mice (Harr *et al.* 2016) to evaluate genetic structure with principal components analysis using the R package *SNPRelate* (Zheng *et al.* 2012). We then used genotype data from 76 classic laboratory strains (Yang *et al.* 2011) to test for gametic disequilibrium (r^2) between candidate sterility regions and SNPs on other chromosomes using PLINK v2.0 (Chang *et al.* 2015). We restricted these analyses to SNPs between *musculus*^{PWK} and *musculus*^{CZII} $\geq$1 Mb apart with no missing data, minor allele frequencies $\geq$0.1, and that were also fixed between strains of *M. m. musculus* (CZECHII, STUS, and STUP) and *M. m. domesticus* (LEWES, ZALENDE, and PERA) with very low levels of introgression (Yang *et al.* 2011; Didion and Pardo-Manuel de Villena 2012).

Third, we evaluated phylogenetic discordance using whole exome data from 10 species of *Mus* (Sarver *et al.* 2017) and whole genomes from *domesticus*^{C57} and *domesticus*^{WSB} (Keane *et al.* 2011). We cleaned and mapped reads to species-specific exome-pseudoreferences generated by Sarver *et al.* (2017). We used RAxML v8.2.3 (Stamatakis 2014) to estimate maximum likelihood phylogenies (rapid bootstrapping and a GTR + Γ model of sequence evolution) for nonoverlapping 100 kb windows, and used these windows to produce a concatenated

Table 1 Reproductive phenotypes for *M. m. domesticus* and hybrid males

	M. m. domesticus	F1 hybrids		F1 hybrid test cross	
	domesticus^{WSB} × *domesticus*^{LEW}	*domesticus*^{WSB} × *musculus*^{PWK}	*domesticus*^{WSB} × *musculus*^{CZII}	*domesticus*^{WSB} × *musculus*$^{PWK \times CZII}$	*domesticus*^{WSB} × *musculus*$^{CZII \times PWK}$
Sample size	21	28	30	78	78
Body weight (g)	17.32 ± 3.1e−01	18.26 ± 4.6e−01	17.15 ± 2.8e−01	18.39 ± 2.8e−01	17.87 ± 2.1e−01
Relative paired testis weight (mg/g)	11.5 ± 2.9e−01	**7.5 ± 1.5e−01** ▼	**6.2 ± 1.3e−01** ▼	6.6 ± 1.3e−01 ▼	6.95 ± 1.6e−01 ▼
Relative paired seminal vesicle weight (mg/g)	5.25 ± 3.7e−01	6 ± 2.2e−01 ▲	5.8 ± 1.7e−01	5.9 ± 1.4e−01 ▲	5.6 ± 1.1e−01
Proportion motile sperm	0.77 ± 3.1e−02	0.83 ± 1.9e−02	0.77 ± 3.4e−02	0.84 ± 1.6e−02 ▲	0.83 ± 2.7e−02 ▲
Sperm count (1×10^6)	19.4 ± 1.5e+00	**11.8 ± 1.6e+00**	**5.8 ± 9.6e−01** ▼	4.9 ± 4.2e−01 ▼	5.7 ± 4.2e−01 ▼
Sperm head morphology index	2.99 ± 4.7e−03	**2.84 ± 2e−02** ▼	**1.14 ± 4.8e−02** ▼	2.43 ± 6.6e−02 ▼	2.18 ± 7.6e−02 ▼
Proportion normal sperm head attachment	1 ± 8.8e−04	**0.99 ± 2.7e−03** ▼	**0.96 ± 6.2e−03** ▼	0.99 ± 2.8e−03 ▼	0.98 ± 4.4e−03 ▼
Proportion straight proximal sperm tail	0.97 ± 6.9e−03	**0.96 ± 4.6e−03**	**0.99 ± 4.2e−03** ▲	0.97 ± 2.5e−03	0.96 ± 3e−03
Proportion straight distal sperm tail	1 ± 3.1e−03	1 ± 2.6e−03	1 ± 3.4e−03	0.99 ± 1.5e−03 ▼	0.99 ± 1.9e−03 ▼

Only F1 hybrids between 58 and 70 days postpartum (dpp) are included to allow direct comparison with the F1 hybrid test cross. See Figure S1 for reproductive traits across an F1 hybrid's reproductive lifespan. Traits are summarized as median values (±SE), and arrows represent significant increase (▲) or decrease (▼) in trait values relative to the *M. m. domesticus* control cross. Values in bold are significantly different between F1 hybrids (Wilcoxon test, false discovery rate–corrected $P \leq 0.05$). Trait proportions are polarized so that higher values indicate higher quality reproductive traits.

species tree for each chromosome. We then used ASTRAL v4.10.11 (Mirarab and Warnow 2015) to estimate the species tree while accounting for phylogenetic discordance among individually estimated gene trees. Trees were visualized with FigTree v1.4.3.

Genome-wide assessment of TRD

To test for TRD associated with sperm function, we used low-coverage whole genome sequencing of motile and immotile sperm populations collected from four F1 male *M. m. musculus* (*musculus*^{PWK} × *musculus*^{CZII}). Epididymal sperm were collected from euthanized adult males (106 dpp) in 2 ml Dulbecco's PBS (equilibrated at 37° and 5% carbon dioxide overnight). We applied 1 ml aliquots of sperm to a Percoll gradient (1 ml layers of 90 and 45% Percoll at 37°; GE Healthcare Life Sciences) and centrifuged (300 *g* for 13 min) to separate cellular debris (top), immotile sperm (middle), and motile sperm (bottom) (Ng *et al.* 1992; Phelps *et al.* 1999). Immotile and motile sperm fractions (400 μl each) were rinsed (1 ml 1.5 M NaCl, centrifuged at 10,000 *g* for 10 min) and stored at −80°. We purified DNA using the MasterPure Complete DNA purification kit (Epicentre Biotechnologies). Sperm fractions were rinsed in 600 μl of 70% EtOH (centrifuged at 14,000 g for 5 min) and incubated overnight at 55° in 600 μl lysis buffer, 25 μl of 1 M dithiothreitol, and 10 μl of 20 mg/ml proteinase K. We treated samples with RNase A (3 μl, for 30 min at 37°), precipitated the sperm in 200 μl of protein precipitation buffer (centrifuged at 14,000 *g* for 30 min), and incubated in 600 μl of isopropanol at −80° for 2–3 hr (centrifuged at 14,000 g for 20 min). The pellet was rinsed with 500 μl 75% ethanol (centrifuged at 14,000 g for 10 min) and dried overnight. We then constructed sequencing libraries using the NEBNext Ultra DNA Library Prep Kit for Illumina with Bio Scientific NEXTflex DNA Barcodes and generated 76 bp paired-end sequences on a HiSeq 2000 at the Epigenome Center, University of Southern California.

We conducted all analyses using reads mapped to strain-specific pseudoreferences for *musculus*^{CZII} or *musculus*^{PWK} (Sarver *et al.* 2017). Briefly, for each whole genome (described above), we called SNPs relative to GRCm38 (GATK HaplotypeCaller), hard-filtered our SNPs (maskExtension 5, QD < 2.0, FS > 60.0, MQ < 40.0, MQRankSum < −12.5, ReadPosRankSum < −8.0, QUAL < 30.0, minDP 10, maxDP 150), recalled SNPs that passed filtering at a base-pair resolution in each genome, and used this high-confidence SNP set to inject variants into GRCm38 using the GATK FastaAlternativeReferenceMaker. We trimmed and quality-filtered sperm fraction reads using expHTS (Streett *et al.* 2015), mapped reads to each pseudoreference using BWA-MEM, and called SNPs using the GATK (HaplotypeCaller). We assigned reads (MQ ≥ 56) that overlapped at least one diagnostic SNP as either *musculus*^{PWK} or *musculus*^{CZII} origin, and summarized read counts in 1 Mb sliding windows (step size 0.5 Mb). We tested for TRD in windows with ≥100 reads in all samples using a χ^2 test (false discovery rate–corrected $P < 0.01$;

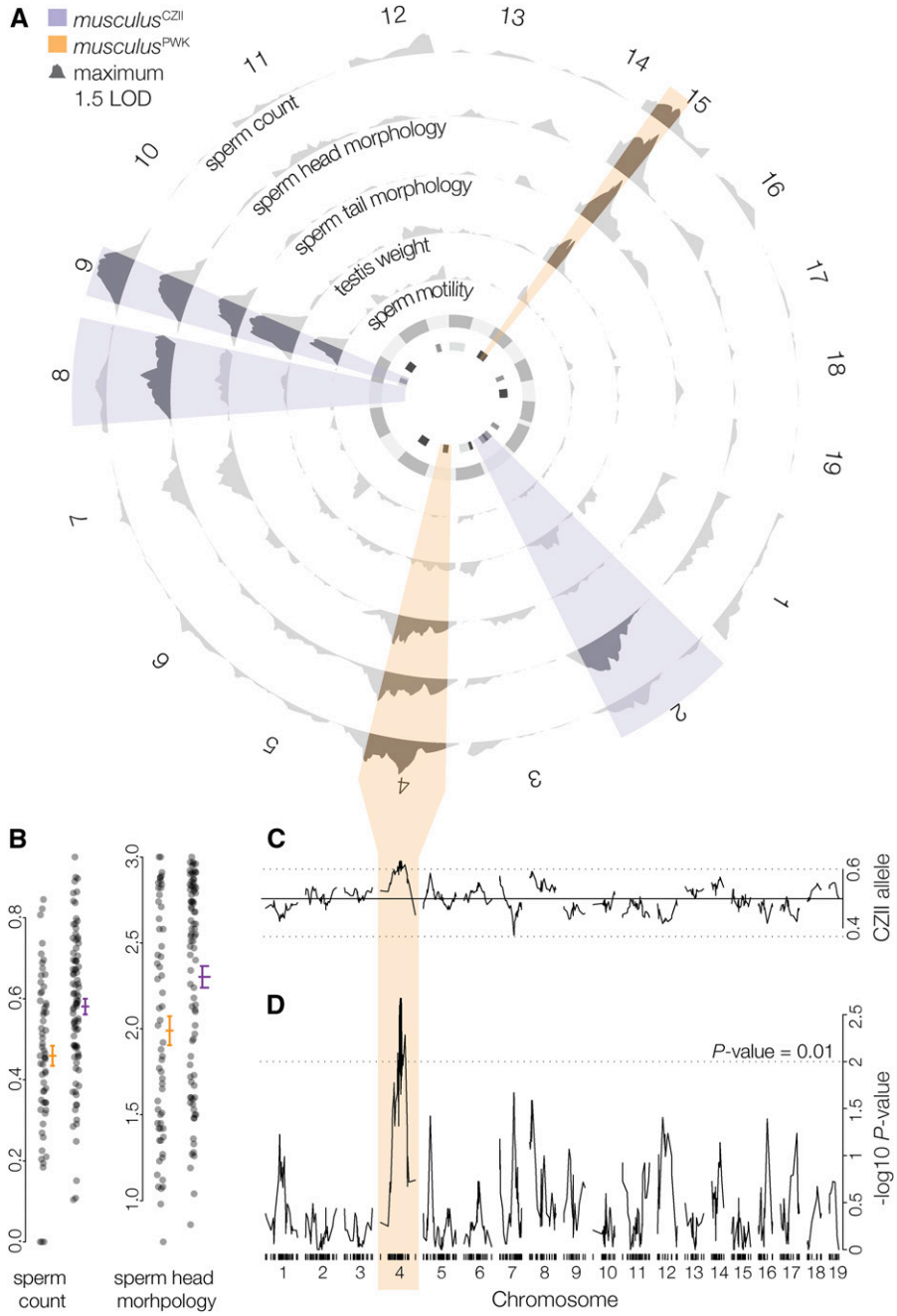

Figure 2 QTL for polymorphic HMS in *musculus*. (A) LOD curves (standard interval mapping) for HMS phenotypes. Highlighted intervals are the maximum LOD intervals (across all traits) on each chromosome for QTL associated with lower fertility in *musculus*^{CZII} (purple) and *musculus*^{PWK} (orange). The inner circle is QTL LOD support intervals for previously reported hybrid sterility loci mapped in *M. m. domesticus* and *M. m. musculus* F2 crosses (White *et al.* 2011; dark gray), wild mice from the hybrid zone (Turner *et al.* 2014; medium gray), and an F1 hybrid test cross (Bhattacharyya *et al.* 2014; light gray). (B) Normalized sperm count and sperm head morphology index plotted against the genotype of the marker with the largest Chr 4 LOD score. Lines indicate mean trait values (±SE). (C) Frequency of the *musculus*^{CZII} allele at each marker (Mendelian expectations 0.5:0.5, genome wide average: 0.496:0.504). (D) TRD plotted as the −log10P value from χ^2 test for Mendelian segregation per chromosome. Tick marks indicate SNP positions.

Benjamini and Hochberg 1995) to test the proportions of *musculus*^{PWK} *vs.* *musculus*^{CZII} reads in motile *vs.* immotile sperm fractions. We considered a window skewed if the proportion of *musculus*^{CZII}-derived reads significantly differed by at least 0.15 between sperm fractions.

To validate our Percoll method, we repeated our pipeline with experimentally combined normal and heat-shocked (immotile) sperm samples from two predominantly *M. m. domesticus* inbred strains (C57BL/6J and DBA/2J). We pooled normalized sperm extractions from each strain and then mixed sperm in equal proportions from the two strains, and repeated our experiment to vary which strain was heat shocked. We isolated DNA as described above, and PCR-amplified and Sanger-sequenced through a marker containing a diagnostic SNP.

Data availability

All data are available through NCBI under projects SRP093943 (F1 hybrid test cross RADseq), SRP094878 (*M. m. musculus* CZECHII/EiJ whole genome sequencing), SRP094877 (*M. m. domesticus* LEWES/EiJ whole genome sequencing), SRP082237 (sperm pools whole genome sequencing), and SPR102485 (backcross RADseq). Supplemental Material, File S1 contains phenotype data for F1 hybrids and F1 hybrid test cross. File S2 contains microsatellite genotypes for markers inside and outside the Chr 4 TRD

region. Supplemental material available at Figshare: https://doi.org/10.25386/genetics.6149399.

Results

HMS is polymorphic and polygenic in M. m. musculus

We found that F1 *M. m. domesticus* × *M. m. musculus* hybrids had variable fertility that was dependent on the strain of *M. m. musculus* sires (Table 1), extending previous results (Good *et al.* 2008b). Compared to fertile *M. m. domesticus* F1 males (*domesticus*^{WSB} × *domesticus*^{LEW}), hybrid males with *musculus*^{PWK} sires had smaller testes and more abnormal sperm morphologies. Hybrid males with *musculus*^{CZII} sires were even more severely sterile. These males had smaller testes, lower sperm counts, and a high proportion of abnormal sperm head and tail morphologies compared to hybrid males with *musculus*^{PWK} sires (Table 1). The fertility of *domesticus*^{WSB} × *musculus*^{PWK} hybrids was lower than previously reported from *domesticus*^{LEW} × *musculus*^{PWK} crosses (Good *et al.* 2008b; Campbell *et al.* 2013; Larson *et al.* 2017), suggesting that *domesticus*^{WSB} has sterility factors not present in other strains of *M. m. domesticus* (Odet *et al.* 2015). Differences among F1 crosses remained qualitatively consistent as males aged and sperm head morphology actually worsened with age (Figure S1). Therefore, HMS was not due to delayed reproductive maturity, as has been observed in other crosses (Campbell and Nachman 2014; Flachs *et al.* 2014). Overall, we found that *M. m. domesticus* × *M. m. musculus* HMS was dependent on the paternal strain, indicating autosomal and/or Y-linked sterility loci contribute to polymorphic sterility in *M. m. musculus*.

Next, we used an F1 hybrid test cross to generate 156 males (62 litters) that ranged from reproductively normal to mostly sterile. On average these males had smaller testes, lower sperm counts, and more abnormal sperm head and tail morphologies (Table 1). After filtering for coverage, we retained ddRADseq data for 150 males that had between 118,000 and 921,000 uniquely mapped paired reads (median 297,634, total mapped reads of 48.5 million paired reads). We constructed a genetic map using 582 high-quality SNPs between *musculus*^{CZII} and *musculus*^{PWK}. Using standard interval mapping we detected two regions of the *M. m. musculus* genome on Chr 9 and Chr 15 that contributed to multiple sterility phenotypes (Figure 2A), suggesting a shared genetic and/or developmental basis. Chr 9 QTL reduced the fertility of hybrids carrying a *musculus*^{CZII} allele and Chr 15 QTL reduced the fertility of hybrids carrying a *musculus*^{PWK} allele (Table 2). We identified two additional QTL on Chr 2 and Chr 8 that contributed to abnormal sperm head morphologies associated with the *musculus*^{CZII} allele, and QTL on Chr 4 that contributed to lower sperm counts and abnormal sperm head and tail morphologies associated with the *musculus*^{PWK} allele (Figure 2B). Using two-dimensional QTL mapping, we identified pairs of QTL that additively contributed to sperm count (Chr 4 and Chr 9) and abnormal sperm head morphologies

Table 2 Polymorphic hybrid sterility QTL detected using standard interval QTL mapping

	Chr	Position (cM)	LOD score	*P*-value	Position (Mb)	1.5 LOD interval (Mb)	%Var[a]	Effect[b]	Phenotype means ± SE	
									musculus^{PWK}	*musculus*^{CZII}
Relative paired testis weight (mg/g)	9	36.49	4.79	0.001	92.81	49.08–113.49	11.22	−0.88	7.3 ± 0.13	6.34 ± 0.15
	15	22.74	3.12	0.023	78.25	29.87–84.06	6.9	0.69	6.49 ± 0.14	7.28 ± 0.15
Proportion motile sperm[c]	9	35.49	3.11	0.012	88.11	73.06–122.96	6.42	−0.1	0.84 ± 0.02	0.74 ± 0.02
Normalized sperm count[d]	4	41.82	3.25	0.031	125.29	90.72–155.46	9.83	0.12	0.46 ± 0.02	0.58 ± 0.02
	9	36.49	3.03	0.051	92.81	59.96–111.01	8.77	−0.12	0.58 ± 0.02	0.47 ± 0.02
Sperm head morphology index	2	75.42	3.06	0.023	174.03	148.24–198.11	5.03	−0.29	2.37 ± 0.07	1.97 ± 0.07
	4	67.84	2.87	0.041	154.51	4.17–180.22	3.62	0.25	1.95 ± 0.08	2.33 ± 0.06
	8	45.05	2.86	0.042	113.45	58.94–169.42	7.03	−0.34	2.38 ± 0.08	2.02 ± 0.07
	9	41.49	3.21	0.015	105.4	64.49–122.96	5.98	−0.32	2.38 ± 0.07	1.96 ± 0.07
	15	22.74	4.94	<0.001	78.25	44.19–84.06	8.48	0.38	1.95 ± 0.07	2.43 ± 0.07
Proportion normal sperm head attachment[c]	4	65.23	2.76	0.061	152.74	4.17–180.22	5.75	0.02	0.96 ± 0.004	0.98 ± 0.004
	9	42.49	2.69	0.069	106.53	64.50–122.96	3.77	−0.01	0.98 ± 0.004	0.96 ± 0.004
	15	22.15	4.85	<0.001	77.78	71.46–86.98	7.13	0.02	0.96 ± 0.004	0.98 ± 0.004

[a] The percent of the phenotypic variance explained by the QTL, calculated using Haley–Knott regression.
[b] The difference between the phenotype averages of the *musculus*^{PWK} and *musculus*^{CZII} alleles. A negative effect indicates the *musculus*^{CZII} allele lowers the reproductive phenotype value. Effects were estimated using Haley–Knott regression.
[c] Phenotypes were analyzed using nonparametric interval mapping and %var and effect were estimated using standard interval mapping.
[d] Square-root transformed sperm count (1×10^6).

Table 3 Polymorphic F1 hybrid sterility QTL detected using multiple QTL mapping

	Chr	Position (cM)	LOD score	*P*-value	Position (Mb)	1.5 LOD interval (Mb)	%Var QTL[a]	%Var Full[b]	Effect ± SE[c]
Relative paired testis weight (mg/g)	9	34.65	4.71	<0.001	85.99	59.96–113.49	13.46	NA	−0.96 ± 0.20
Normalized sperm count[d]	4	41.82	3.7	<0.001	125.29	90.72–155.46	9.83	18.30	0.12 ± 0.03
	9	34.65	3.32	<0.001	85.99	59.96–104.82	8.77		−0.12 ± 0.03
Sperm head morphology index	4	67.84	6.82	<0.001	154.51	149.56–155.46	12.93	45.35	0.35 ± 0.08
	8	45.05	5.08	<0.001	113.45	95.68–129.09	9.36		−0.4 ± 0.08
	9	34.65	9.33	<0.001	85.99	85.99–85.99	18.41		−0.5 ± 0.08
	15	22.74	5.72	<0.001	78.25	76.5–78.25	10.65		0.42 ± 0.08
	4:9	NA	3.95	<0.001	NA	NA	7.14	NA	0.70 ± 0.16

QTL identified using nonparametric interval mapping were not assessed using multiple QTL mapping.
[a] The percent of the phenotypic variance explained by each QTL.
[b] The percent of the phenotypic variance explained by all terms (*e.g.*, all QTL) in the model.
[c] The difference between the phenotype averages of the *musculus*^{PWK} and *musculus*^{CZII} alleles. A negative effect indicates the *musculus*^{CZII} allele lowers the reproductive phenotype value.
[d] Square-root transformed sperm count (1×10^6).

(Chrs 1, 2, 4, 7, 8, 9, and 15). We found no evidence of epistatic interactions (Table S1), although sample sizes were likely too small to detect such effects. Multiple QTL models supported several loci contributing to sperm count and abnormal sperm head morphology, consistent with our single QTL results (Table 3). Neither Chr Y origin nor genotyped *Prdm9* alleles were associated with hybrid sterility phenotypes (Table S2).

Hybrid sterility QTL colocalized with TRD on Chr 4

Sterility phenotypes associated with the *musculus*^{PWK} allele on Chr 4 colocalized with a large region (46.91:153.39 Mb) that had a deficit of *musculus*^{PWK} alleles at 50 consecutive markers (expected allelic ratio: 50:50, median observed 39.5:60.5, χ^2 test, $P \leq 0.05$; Figure 2, C and D). The most extreme TRD was observed at 116.01:151.14 Mb (median observed 38.7:61.3, χ^2 test, $P \leq 0.001$) and was also observed when crosses were parsed by sire (*musculus*$^{PWK \times CZII}$ sire, $N = 75$, 33.3:66.7; *musculus*$^{CZII \times PWK}$ sire $N = 75$, 44.3:55.7). Sex ratios were normal in these crosses (females: males 51:49, χ^2 test, $P = 0.701$). The Chr 4 region showing TRD overlapped with QTL for lower sperm count (±1.5 LOD interval 90.72–155.46) and more abnormal sperm head and tail morphology (±1.5 LOD interval 4.17–180.22) in males with the *musculus*^{PWK} allele. This could be due to chance given that the ±1.5 LOD intervals for all sterility QTL encompassed 19.1% (275.59 cM) of the total genome, although sterility QTL associated with *musculus*^{PWK} alleles encompassed only 5.5% of the genome (79.16 cM total).

Chr 4 sterility and TRD loci showed unusual patterns of introgression

The distal region of Chr 4 showing TRD contained an unusually high density of SNPs between *musculus*^{CZII} and *musculus*^{PWK} (hypergeometric test, $P < 0.001$). Previous work has shown appreciable subspecific introgression into *musculus*^{PWK}, including a large tract of *M. m. domesticus* introgression on the distal portion of Chr 4 (Yang *et al.* 2011). Consistent with this, we found considerable genome-wide introgression between *musculus*^{PWK} and *M. m. domesticus* (median *D*-statistic of 0.253, Figure 3A). Across most chromosomes, *D*-statistic estimates were similar regardless of which *M. m. domesticus* strain was used. On Chr 4 we detected introgression between *M. m. domesticus* (*domesticus*^{WSB}, *domesticus*^{LEW}) and *musculus*^{PWK}, but not between *domesticus*^{C57} and *musculus*^{PWK}. Discordance in the *D*-statistic among *M. m. domesticus* strains was localized to a 20 Mb region on the distal end of Chr 4 (130–150 Mb), coincident with the *musculus*^{PWK} sterility QTL and the region with the strongest TRD (Figure 3B). For simplicity, we will refer to this narrower region as the Chr 4 TRD locus.

We then contrasted patterns of divergence outside and inside of the Chr 4 TRD locus using other *M. m. musculus* inbred strains and wild house mice (Yang *et al.* 2011; Harr *et al.* 2016). Wild mice strongly clustered by subspecies both outside and inside of the TRD locus (Figure 4A). In contrast, some *M. m. musculus* wild-derived strains (PWK and PWD) clustered with *M. m. domesticus* at the TRD region while classic laboratory strains (primarily *M. m. domesticus* in origin) showed a mosaic of subspecific origins within the Chr 4 TRD locus (Figure 4B). We then estimated gametic disequilibrium between Chr 4 (130–150 Mb) and 1313 autosomal and Y-linked SNPs to test for other genomic regions that may be associated with the Chr 4 TRD locus. Overall, gametic disequilibrium was low (r^2: median 0.014, maximum 0.645), and similar to prior estimates (Payseur and Hoekstra 2005). Seven SNPs showed elevated r^2 (Table S3), but we did not find any SNPs that had high r^2 values across the Chr 4 TRD haplotype. There was no association between Chr 4 and Chr Y, which has either a *M. m. musculus* or *M. m. domesticus* origin in the classic strains (Bishop *et al.* 1985).

To evaluate the deeper evolutionary origin of the Chr 4 TRD locus, we estimated the phylogenies across 10 species of *Mus* for Chr 3 (a similar sized chromosome with limited introgression relative to other chromosomes and no TRD) and Chr 4 regions outside and inside the TRD locus. Concatenated trees for Chr 3 and non-TRD Chr 4 (Figure S2) were consistent with previous species tree estimates (Sarver *et al.* 2017).

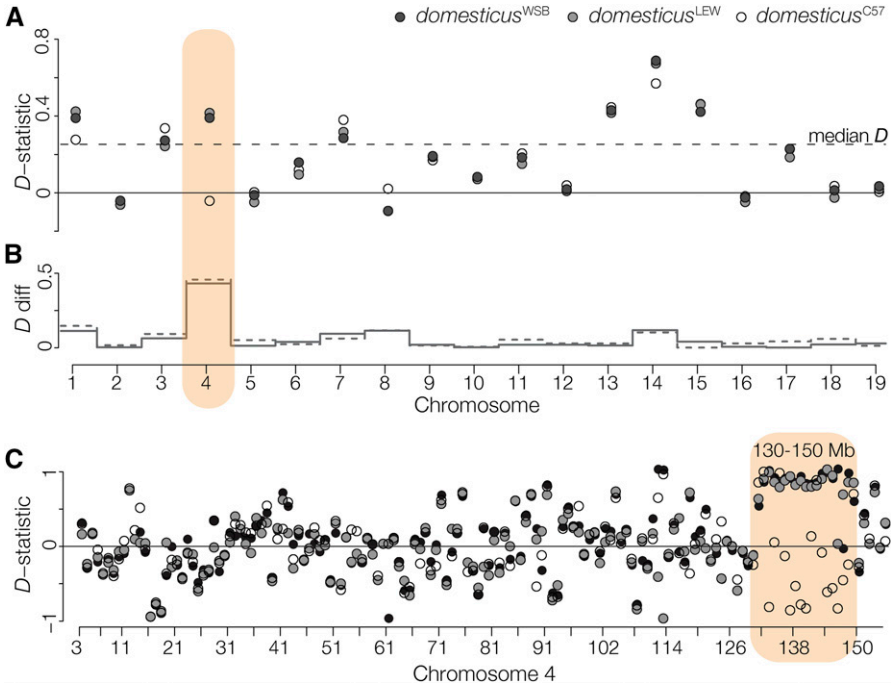

Figure 3 Introgression between *M. m. musculus* and *M. m. domesticus*. (A) *D*-statistics, calculated for each chromosome, testing for introgression between *musculus*^{CZII} or *musculus*^{PWK} and *M. m. domesticus*. The median *D*-statistic across all *M. m. domesticus* comparisons is represented by the dashed line. Patterns on Chr 4 vary depending on the strain of *M. m. domesticus*. (B) The absolute difference in the *D*-statistic using *domesticus*^{WSB} (solid gray line) or *domesticus*^{LEW} (dashed gray line) compared to *domesticus*^{C57}. (C) *D*-statistic calculated over 1 Mb nonoverlapping windows localizes discordant introgression to a 20 Mb window on the distal end of Chr 4.

In contrast, trees from the Chr 4 TRD locus showed reciprocal swapping of *M. m. musculus* and *M. m. domesticus* strains, with *musculus*^{PWK} closest to *domesticus*^{WSB} and *musculus*^{CZII} closest to *domesticus*^{C57}. These conflicting patterns were most apparent using a gene-tree approach to characterize patterns of fine-scale topological discordance across Chr 4 (Figure 4C). We found no other *Mus* lineages with variant topologies for the Chr 4 TRD region.

TRD was restricted to hybrid crosses and was not associated with sperm motility

Male meiotic drivers often operate through various mechanisms of sperm impairment (Lindholm *et al.* 2016). In the classic house mouse *t* complex drive system, heterozygous males show a higher frequency in motile sperm of the sperm killing Chr 17 *t* haplotype (Lyon 2003). We used whole genome sequencing of sperm pools to test if the higher frequency of the *musculus*^{CZII} Chr 4 TRD haplotype in the offspring of the F1 hybrid test cross reflected motility differences in the sperm of the *M. m. musculus* (*musculus*^{PWK} × *musculus* ^{CZII}) sires. We generated between 40 and 64 million uniquely mapped reads (MQ ≥ 56) from the motile and immotile sperm fractions of four F1 *M. m. musculus* males. An average of 7 million reads per sample spanned at least one diagnostic SNP. We parsed reads into 5463 overlapping 1 Mb windows and analyzed an average of 4777 windows with ≥100 mapped reads and ≥1 diagnostic SNP (QUAL ≥ 24). No window showed significant skew between motile and immotile sperm pools in any male. In an additional experiment, we confirmed that the Percoll method was effective in separating motile from immotile sperm (Figure S3).

Our sequencing experiment demonstrated that Chr 4 TRD is likely not related to sperm motility and therefore must reflect genotypic differences in sperm competitive interactions (including female choice), fertilization ability, or postzygotic development. To localize the timing of distortion, we crossed the same *M. m. musculus* sire genotype (*musculus*^{PWK} × *musculus* ^{CZII}) to *musculus*^{PWK} females to generate 602 backcross offspring (319 female and 283 male) from 133 litters. We generated ddRADseq libraries for 312 backcross mice. After removing individuals with low coverage, we retained 303 mice that had between 156,791 and 3,315,067 uniquely mapped reads (median of 664,627 per mouse, total of 232,515,877 single reads) and we constructed a genetic map using 358 high-quality SNPs. We found no evidence of TRD on Chr 4 (expected allelic ratio, 50:50; median observed, 50:50) and this pattern held when parsed by sire, sex, or sire and sex. To confirm these results, we genotyped an additional 193 backcross mice using microsatellite markers spanning Chr 4 and still found no evidence of TRD (N = 496 mice; Table 4). Thus, Chr 4 TRD between *musculus*^{CZII} and *musculus*^{PWK} alleles was only observed in crosses involving *domesticus*^{WSB} females (Figure 5A).

Discussion

Polymorphic HMS has a polygenic basis in house mice

Individuals often vary in the degree that they are reproductively isolated from other lineages, but the genetic basis and evolutionary origin of such variation remains poorly understood. In house mice, there is considerable variability in the strength of F1 HMS in crosses using different inbred strains or wild isolates of *M. m. musculus* and *M. m. domesticus* (Britton-Davidian *et al.* 2005; Vyskočilová *et al.* 2005; Good *et al.* 2008b; Bhattacharyya *et al.* 2014). One simple interpretation

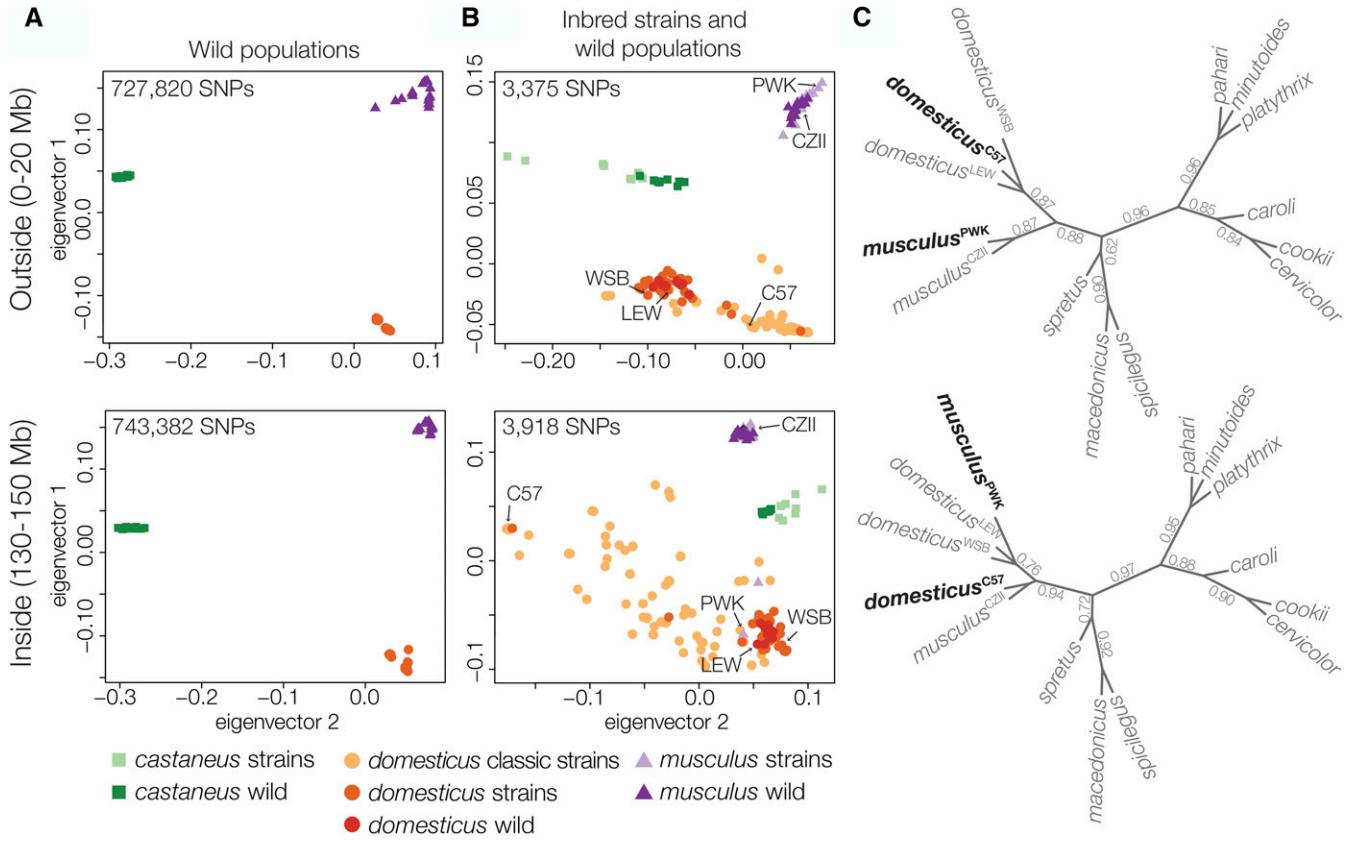

Figure 4 Patterns of discordance outside and inside the Chr 4 TRD locus. (A) Principal components analysis of SNPs from whole genome sequencing of eight wild populations of *M. m. musculus*, *M. m. domesticus*, and *M. m. castaneus* (Harr *et al.* 2016) and (B) the Mouse Diversity Array for wild mice (dark colors) and classic laboratory and wild-derived strains of mice (light colors) (Yang *et al.* 2011). There was strong clustering of SNPs in wild populations, both outside and inside the TRD region, but the classic strains showed mixed SNP clustering. (C) Unrooted species trees estimated across 100 kbp windows outside and inside the Chr 4 TRD locus. Branches are annotated with their local quartet scores.

of these results is that there are one or a few common incompatibilities that are polymorphic within *M. m. musculus* and/or *M. m. domesticus* populations. Consistent with this, the only HMS gene yet identified in mammals, *Prdm9*, appears to be polymorphic for sterile and fertile alleles within both *M. m. musculus* and *M. m. domesticus* (Forejt and Ivanyi 1974; Vyskocilová *et al.* 2009; Flachs *et al.* 2012). While the evolutionary origin and extent of *Prdm9*-linked HMS variation remains unclear in natural populations, our results reveal that there is likely to be considerable polymorphism at other HMS loci.

We identified five autosomal regions that contributed to variation in HMS in crosses between *M. m. domesticus* females and *M. m. musculus* males, despite sampling just two wild-derived inbred strains of *M. m. musculus* (Figure 2). F1 hybrid males from crosses between female *M. m. domesticus* and *musculus*^{PWK} yield only weak sterility phenotypes, while crosses involving *musculus*^{CZII} are more severely sterile in both directions of the cross (Table 1; Good *et al.* 2008b). Surprisingly, considerable variation exists beneath this seemly simple F1 architecture. Sterility loci were associated with both strains; sterility alleles on Chrs 2, 8, and 9 derived from *musculus*^{CZII}, while *musculus*^{PWK} sterility variants were mapped to Chr 4 and Chr 15.

F1 HMS variability has been observed in other *M. m. musculus* strains (Piálek *et al.* 2008; Bhattacharyya *et al.* 2014) and in wild *M. m. musculus* isolated from eastern Czechia (Good *et al.* 2008b). Thus, the polymorphic HMS that we document here may be relatively widespread within *M. m. musculus* (Vyskočilová *et al.* 2005; Good *et al.* 2008b; Bhattacharyya *et al.* 2014). Consistent with this, there was some overlap between the HMS loci we identified and HMS QTL from other studies (Figure 2A). Bhattacharyya *et al.* (2014) used a similar experimental design between *domesticus*^{C57} females and *M. m. musculus* interstrain males (PWD and STUS) to map polymorphic hybrid sterility to Chr 9 (sperm count). Sterility loci were identified on Chr 4 (epididymis weight) in recombinant inbred lines derived from all three *M. m. musculus* subspecies (Shorter *et al.* 2017), Chr 4 (testis weight) and Chr 15 (abnormal sperm morphology) were identified in F2 crosses between *domesticus*^{WSB} and *M. m. musculus* PWD (White *et al.* 2011; Turner *et al.* 2014), and Chr 2 and Chr 9 were associated with low testis weights in wild-caught hybrid mice (Turner and Harr 2014). However, these studies also found sterility QTL on Chrs 1, 2, 3, 5, 6, 10, 12, 13, 14, 17, and 18, which implies that nearly every autosome is linked to some form of HMS. The emerging picture is

Table 4 Summary of genotype frequencies in *M. m. musculus* backcross

	Crosses	AA	AB	Total	%AB	*P*-value
Region 1	*musculus*^{CZII} × ^{PWK} sire	100	93	195	47.7	0.614
140,089,156–	females	54	50	106	47.2	0.695
141,037,913 bp	males	46	43	89	48.3	0.75
	musculus^{PWK} × ^{CZII} sire	144	157	301	52.2	0.454
	females	77	83	160	51.9	0.635
	males	67	74	141	52.5	0.556
	Total	244	250	496	50.4	0.787
Region 2	*musculus*^{CZII} × ^{PWK} sire	97	97	195	49.7	1
148,602,050–	females	53	53	106	50	1
151,609,913 bp	males	44	44	89	49.4	1
	musculus^{PWK} × ^{CZII} sire	138	162	301	53.8	0.166
	females	73	86	160	53.8	0.303
	males	65	76	141	53.9	0.354
	total	235	259	496	52.2	0.28

To test for TRD within *M. m. musculus*, female *musculus*^{PWK} were crossed to reciprocal interstrain F1s between *musculus*^{CZII} and *musculus*^{PWK}. Offspring were genotyped for two regions inside the introgressed TRD on Chr 4. AA and AB represent offspring genotype.

of an increasingly complex genetic basis to HMS that depends strongly on genotype. Indeed, multiple polymorphic hybrid sterility factors would account for the variable fertility of multigeneration hybrids from the center of the hybrid zone (Turner *et al.* 2012). Inbred line crosses remain one of the most powerful tools for the genetic dissection of hybrid incompatibilities. However, the common assumption that most incompatibilities reflect fixed differences between lineages appears increasingly tenuous, especially during the early stages of speciation. This realization has the potential to broadly impact important issues in speciation genetics. In addition to the need to incorporate population-level sampling into the design of mapping studies on the genetics of speciation, many theoretical predications on the accumulation of reproductive isolation are based on epistatic models that treat interacting hybrid incompatibilities as fixed within species (*e.g.*, Orr and Turelli 2001; Wang *et al.* 2013; Lindtke and Buerkle 2015).

The causes of polymorphic reproductive isolation

Several incompatibilities are polymorphic in house mice, but the origins of these variants are unclear. One possible source is introgression at previously fixed incompatibilities. Alleles contributing to hybrid incompatibilities should have restricted introgression relative to the rest of the genome. Indeed, the identification of loci showing restricted gene flow across hybrid zones is a powerful approach to identifying alleles that contribute to reproductive barriers (Barton and Hewitt 1985; Harrison 1990; Payseur 2010). However, gene flow and recombination within a hybrid zone can quickly break down epistatic interactions among BDMIs (Virdee and Hewitt 1994; Shuker *et al.* 2005; Bank *et al.* 2012; Lindtke and Buerkle 2015), which could in turn result in polymorphic incompatibilities. The house mouse hybrid zone is wide relative to the dispersal distances of mice. As a result, pure *M. m. domesticus* and *M. m. musculus* rarely come into contact and few F1 mice are found in the hybrid zone. The zone is primarily composed of complex, multigeneration hybrids that show extensive variation in the severity of HMS (Janoušek *et al.* 2012; Turner *et al.* 2012; Turner and Harr 2014), which likely reflects the partial breakdown of epistatic reproductive barriers.

Several of the common wild-derived strains show appreciable introgression between subspecies of *M. m. musculus*, including *musculus*^{PWK} (Yang *et al.* 2011; Sarver *et al.* 2017). Four of our polymorphic HMS regions did not colocalize with strong signatures of introgression (results not shown), although gene flow cannot be ruled out at our current mapping resolution. At least one HMS region (Chr 4) did coincide with introgression into *musculus*^{PWK} (Figure 3), but not necessarily in the direction predicted if HMS polymorphism reflects the partial erosion of reproductive barriers. Hybrid sterility QTL on Chr 4 contributed to low sperm counts and abnormal sperm morphology in males carrying the *musculus*^{PWK} allele (Figure 2A). Coincident with the Chr 4 HMS QTL, an ~20 Mb *M. m. domesticus* haplotype (represented here by *domesticus*^{WSB} and *domesticus*^{LEW}) was introgressed into *musculus*^{PWK}, while a *M. m. musculus* haplotype (represented by *musculus*^{CZII}) appears introgressed into *domesticus*^{C57} (Figure 3) and some other classic strains (Figure 4, Yang *et al.* 2011). At least two *M. m. musculus* strains (PWK and PWD) derived from different localities carry introgressed *M. m. domesticus* haplotypes. In other words, a *M. m. domesticus*-derived hybrid sterility locus has introgressed into at least two independent *M. m. musculus* strains. Transmission of the same introgressed *musculus*^{PWK} allele was also underrepresented in our hybrid test cross (Figure 2C). Thus, the distal end of Chr 4 shows a propensity to reciprocally introgress between *M. m. musculus* and *M. m. domesticus* genomes despite asymmetric TRD and detrimental effects on hybrid fertility.

How can recurrent reciprocal introgression be reconciled with the evolution of HMS and TRD in the same genomic region? Non-Mendelian segregation is common in divergent crosses and can reflect differences in gamete production, fertilization, and zygote survival (Lindholm *et al.* 2016). For example, sexual selection can lead to TRD when gametes carrying different alleles have contrasting fertilization abilities due to male gamete competition or cryptic female choice (*e.g.*, Fishman *et al.* 2008). We did not observe TRD in our independent *M. m. musculus* backcross or distal Chr 4 introgression in wild mice, arguing against simple competitive advantage of the *musculus*^{CZII} haplotype. In divergent crosses, TRD is often caused by biased transmission of selfish genetic elements (*i.e.*, meiotic drive or segregation distortion; McDermott and Noor 2010; Lindholm *et al.* 2016) as found, for example, at the *R2d2* locus in house mice (Didion *et al.* 2015, 2016). Drive elements generate intragenomic conflict, which should drive strong counter selection for unlinked drive suppressors. Drive systems coevolve independently in isolated populations, which can lead to sterility when drivers and suppressers are uncoupled in hybrid genomes (Frank 1991; Hurst and Pomiankowski 1991). Male meiotic drivers often act by impairing the development or fertilization capacity of nondriving sperm (Lindholm *et al.* 2016). We tested this

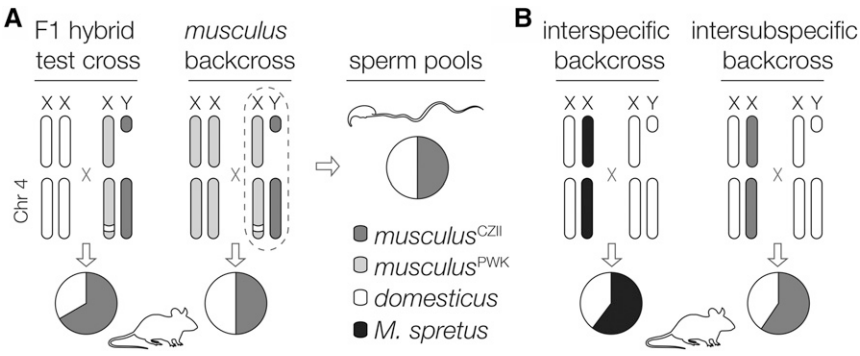

Figure 5 Summary of Chr 4 TRD. (A) The *musculus*^{PWK} haplotype on the distal portion of Chr 4, which is derived from *M. m. domesticus*, was undertransmitted in the offspring of crosses involving *M. m. domesticus* females and *M. m. musculus* males, but not in *M. m. musculus* backcrosses or the sperm of *M. m. musculus* males. (B) Two other crosses have reported reduced transmission of the distal portion of Chr 4 derived from *M. m. domesticus*, but through females. The first was an interspecific cross between F1 females (*M. m. domesticus* C57BL/6J × *M. spretus*) and *M. m. domesticus* C57BL/6J males (TRD ~73 Mb to end of Chr 4; Ceci *et al.* 1989), and the second was an interspecific cross between interstrain F1 females (*M. m. domesticus* C57BL/KsJ × *musculus*^{CZII}) and *M. m. domesticus* C57BL/KsJ males (TRD ~107–138 Mb; Fiedorek and Kay 1994).

scenario directly and found no TRD between motile and immotile sperm of *musculus*$^{PWK \times CZII}$ males. More broadly, TRD on the distal region of Chr 4 has also been reported in two other divergent crosses: TRD favoring the distal Chr 4 *M. spretus* allele in crosses between *domesticus*^{C57} × *M. spretus* F1 females and *domesticus*^{C57} males (Ceci *et al.* 1989), and TRD again favoring the *musculus*^{CZII} distal Chr 4 allele in crosses between *domesticus*$^{C57BL/KsJ}$ males and *domesticus*$^{C57BL/KsJ}$ × *musculus*^{CZII} F1 females (Fiedorek and Kay 1994). Importantly, both crosses reveal reduced transmission of the distal portion of Chr 4 derived from *M. m. domesticus* through female gametogenesis (Figure 5B). If these patterns reflect a common mechanism, then Chr 4 TRD must act independent of male-specific mechanisms.

Collectively, these results suggest that Chr 4 TRD and introgression are both a consequence of incompatibilities that reduce hybrid embryo viability (postzygotic inviability). In principle, TRD could occur because of a negative interaction between egg (or female reproductive tract) and sperm resulting in reduced fertilization (postmating prezygotic barriers) (Nadeau 2017), although incompatible egg–sperm interactions are often asymmetric (*i.e.*, depend on the parent of origin of gametes; Larson *et al.* 2012). Chr 4 TRD occurs in crosses involving both male and female *M. m. domesticus*, with consistent bias against the Chr 4 *M. m. domesticus* allele when backcrossed to *M. m. domesticus* (Figure 5). The simplest explanation for this pattern is a two locus BDMI involving a recessive Chr 4 incompatibility derived in the *M. m. domesticus* lineage. It remains unclear why TRD driven by hybrid inviability in crosses involving *M. m. domesticus* colocalizes with HMS QTL that manifests in the F1 offspring. It is possible that early-acting hybrid inviability leads to the pleiotropic impairment of other reproductive traits. Alternatively, this region may harbor multiple incompatibilities, which appears to be the case for TRD of polymorphic hybrid incompatibilities in monkeyflowers (Kerwin and Sweigart 2017).

Under an inviability model, introgression at the Chr 4 TRD locus in various classic and wild-derived inbred strains (*i.e.*, *musculus*$^{PWK, PWD}$) would reflect different outcomes of selection against particular incompatible allelic pairings. Such epistatic selection should generate linkage disequilibrium between distal Chr 4 and other genomic regions within hybrid genomes. Although our initial scan of genotypes from 76 classic laboratory strains failed to detect these associations (Table S3), multiple genome-wide studies have revealed that selection against other deleterious allelic combinations has shaped the mosaic composition of introgressed laboratory strains (Payseur and Hoekstra 2005; Petkov *et al.* 2005) and the *M. m. domesticus–M. m. musculus* hybrid zone (Turner *et al.* 2012; Turner and Harr 2014). There has been considerable effort to resolve the extent to which various classic and common wild-derived laboratory strains are introgressed, with an emphasis on overall strain genetic purity (Yang *et al.* 2011; Didion and Pardo-Manuel de Villena 2012). While overall admixture proportions are of some relevance, our results suggest that specific genome-wide patterns of introgression may be strongly shaped by selection with the unexpected result that selection against epistatic BDMIs may facilitate introgression at underlying loci. These results underscore the intricacies of nascent species boundaries during the early stages of speciation when reproductive isolation remains incomplete and genetically variable.

Acknowledgments

We thank Kathleen Tsung, Joseph Dysthe, Brent Young, Nick Schultz, Selene Tyndale, and Charlie Nicolet for assistance with data collection and analysis. We also thank members of the Good laboratory; Ryan Bracewell, Lila Fishman, David Aylor, Bret Payseur, Michael Nachman, Peter Ellis, and two anonymous reviewers for helpful feedback; the Vincent J. Coates Genomics Sequencing Laboratory at the University of California Berkeley, supported by the National Institutes of Health S10 instrumentation grants S10RR029668 and S10RR027303; and the University of Montana Genomics Core, supported by a grant from the M.J. Murdock Charitable Trust. This research was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development of the National

Institutes of Health (R01-HD073439 and R01-HD094787 to J.M.G.), the National Institute of General Medical Sciences (R01-GM098536 to M.D.D.), and the National Science Foundation (1146525 to M.D.D.).

Author contributions: J.M.G. conceived the study. J.M.G., E.L.L., and M.D.D. designed the experiments. E.L.L., D.V., C.C., V.S., E.N., and L.P.P. performed experiments and collected data. E.L.L., D.V., B.A.J.S., S.K., L.P.P., M.D.K., and M.D.D. analyzed data. E.L.L. and J.M.G. wrote the manuscript with feedback from the authors.

Literature Cited

Baker, C. L., S. Kajita, M. Walker, R. L. Saxl, N. Raghupathy *et al.*, 2015 PRDM9 drives evolutionary erosion of hotspots in *Mus musculus* through haplotype-specific initiation of meiotic recombination. PLoS Genet. 11: e1004916. https://doi.org/10.1371/journal.pgen.1004916

Bank, C., R. Bürger, and J. Hermisson, 2012 The limits to parapatric speciation: dobzhansky–Muller incompatibilities in a continent–island model. Genetics 191: 845–863. https://doi.org/10.1534/genetics.111.137513

Barton, N. H., and G. M. Hewitt, 1985 Analysis of hybrid zones. Annu. Rev. Ecol. Syst. 16: 113–148. https://doi.org/10.1146/annurev.es.16.110185.000553

Bateson, W., 1909 Heredity and variation in modern lights, pp. 85–101 in *Darwin and Modern Science*, edited by A. C. Seward. Cambridge University Press, Cambridge, UK.

Benjamini, Y., and Y. Hochberg, 1995 Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. 57: 289–300.

Bhattacharyya, T., S. Gregorova, O. Mihola, M. Anger, J. Sebestova *et al.*, 2013 Mechanistic basis of infertility of mouse intersubspecific hybrids. Proc. Natl. Acad. Sci. USA 110: E468–E477. https://doi.org/10.1073/pnas.1219126110

Bhattacharyya, T., R. Reifova, S. Gregorova, P. Simecek, V. Gergelits *et al.*, 2014 X chromosome control of meiotic chromosome synapsis in mouse inter-subspecific hybrids. PLoS Genet. 10: e1004088. https://doi.org/10.1371/journal.pgen.1004088

Bishop, C. E., P. Boursot, B. Baron, F. Bonhomme, and D. Hatat, 1985 Most classical *Mus musculus domesticus* laboratory mouse strains carry a *Mus musculus musculus* Y chromosome. Nature 315: 70–72. https://doi.org/10.1038/315070a0

Britton-Davidian, J., F. Fel-Clair, J. Lopez, P. Alibert, and P. Boursot, 2005 Postzygotic isolation between the two European subspecies of the house mouse: estimates from fertility patterns in wild and laboratory-bred hybrids. Biol. J. Linn. Soc. Lond. 84: 379–393. https://doi.org/10.1111/j.1095-8312.2005.00441.x

Broman, K. W., H. Wu, S. Sen, and G. A. Churchill, 2003 R/qtl: QTL mapping in experimental crosses. Bioinformatics 19: 889–890. https://doi.org/10.1093/bioinformatics/btg112

Campbell, P., and M. W. Nachman, 2014 X-Y interactions underlie sperm head abnormality in hybrid male house mice. Genetics 196: 1231–1240. https://doi.org/10.1534/genetics.114.161703

Campbell, P., J. M. Good, and M. W. Nachman, 2013 Meiotic sex chromosome inactivation is disrupted in sterile hybrid male house mice. Genetics 193: 819–828. https://doi.org/10.1534/genetics.112.148635

Carter, T. C., and D. S. Falconer, 1951 Stocks for detecting linkage in the mouse, and the theory of their design. J. Genet. 50: 307–323. https://doi.org/10.1007/BF02996226

Case, A. L., F. R. Finseth, C. M. Barr, and L. Fishman, 2016 Selfish evolution of cytonuclear hybrid incompatibility in *Mimulus*. Proc. Biol. Sci. 283: 20161493–20161499. https://doi.org/10.1098/rspb.2016.1493

Ceci, J. D., L. D. Siracusa, N. A. Jenkins, and N. G. Copeland, 1989 A molecular genetic linkage map of mouse chromosome 4 including the localization of several proto-oncogenes. Genomics 5: 699–709. https://doi.org/10.1016/0888-7543(89)90111-0

Chang, C. C., C. C. Chow, L. C. Tellier, S. Vattikuti, S. M. Purcell *et al.*, 2015 Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaSci 4: 7–16. https://doi.org/10.1186/s13742-015-0047-8

Christie, P., and M. R. Macnair, 1987 The distribution of postmating reproductive isolating genes in populations of the yellow monkey flower, *Mimulus guttatus*. Evolution 41: 571–578. https://doi.org/10.1111/j.1558-5646.1987.tb05827.x

Copeland, N. G., N. A. Jenkins, D. J. Gilbert, J. T. Eppig, L. J. Maltais *et al.*, 1993 A genetic linkage map of the mouse: current applications and future prospects. Science 262: 57–66. https://doi.org/10.1126/science.8211130

Cutter, A. D., 2012 The polymorphic prelude to Bateson-Dobzhansky-Muller incompatibilities. Trends Ecol. Evol. 27: 209–218. https://doi.org/10.1016/j.tree.2011.11.004

Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks *et al.*, 2011 The variant call format and VCFtools. Bioinformatics 27: 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Davies, B., E. Hatton, N. Altemose, J. G. Hussin, F. Pratto *et al.*, 2016 Re-engineering the zinc fingers of PRDM9 reverses hybrid sterility in mice. Nature 530: 171–176. https://doi.org/10.1038/nature16931

Didion, J. P., and F. Pardo-Manuel de Villena, 2012 Deconstructing *Mus gemischus*: advances in understanding ancestry, structure, and variation in the genome of the laboratory mouse. Mamm. Genome 24: 1–20. https://doi.org/10.1007/s00335-012-9441-z

Didion, J. P., A. P. Morgan, A. M. F. Clayshulte, R. C. Mcmullan, L. Yadgary *et al.*, 2015 A multi-megabase copy number gain causes maternal transmission ratio distortion on mouse chromosome 2. PLoS Genet. 11: e1004850. https://doi.org/10.1371/journal.pgen.1004850

Didion, J. P., A. P. Morgan, L. Yadgary, T. A. Bell, R. C. Mcmullan *et al.*, 2016 *R2d2* drives selfish sweeps in the house mouse. Mol. Biol. Evol. 33: 1381–1395. https://doi.org/10.1093/molbev/msw036

Dobzhansky, T., 1937 *Genetics and the Origin of Species*. Columbia University Press, New York.

Durand, E. Y., N. Patterson, D. Reich, and M. Slatkin, 2011 Testing for ancient admixture between closely related populations. Mol. Biol. Evol. 28: 2239–2252. https://doi.org/10.1093/molbev/msr048

Fiedorek, F. T., and E. S. Kay, 1994 Mapping of PCR-based markers for mouse chromosome 4 on a backcross penetrant for the misty (m) mutation. Mamm. Genome 5: 479–485. https://doi.org/10.1007/BF00369316

Fishman, L., J. Aagaard, and J. C. Tuthill, 2008 Toward the evolutionary genomics of gametophytic divergence: patterns of transmission ratio distortion in monkeyflower (*Mimulus*) hybrids reveal a complex genetic basis for conspecific pollen precedence. Evolution 62: 2958–2970. https://doi.org/10.1111/j.1558-5646.2008.00475.x

Flachs, P., O. Mihola, P. Simecek, S. Gregorova, J. C. Schimenti *et al.*, 2012 Interallelic and intergenic incompatibilities of the *Prdm9* (Hst1) gene in mouse hybrid sterility. PLoS Genet. 8: e1003044. https://doi.org/10.1371/journal.pgen.1003044

Flachs, P., T. Bhattacharyya, O. Mihola, J. Piálek, J. Forejt *et al.*, 2014 *Prdm9* incompatibility controls oligospermia and delayed fertility but no selfish transmission in mouse intersubspecific hybrids. PLoS One 9: e95806. https://doi.org/10.1371/journal.pone.0095806

Forejt, J., and P. Ivanyi, 1974 Genetic studies on male sterility of hybrids between laboratory and wild mice (*Mus musculus* L.). Genet. Res. 24: 189–206. https://doi.org/10.1017/S0016672300015214

Frank, S. A., 1991 Divergence of meiotic drive-suppression systems as an explanation for sex-biased hybrid sterility and inviability. Evolution 45: 262–267. https://doi.org/10.1111/j.1558-5646

Geraldes, A., P. Basset, K. L. Smith, and M. W. Nachman, 2011 Higher differentiation among subspecies of the house mouse (*Mus musculus*) in genomic regions with low recombination. Mol. Ecol. 20: 4722–4736. https://doi.org/10.1111/j.1365-294X.2011.05285.x

Good, J. M., M. D. Dean, and M. W. Nachman, 2008a A complex genetic basis to X-linked hybrid male sterility between two species of house mice. Genetics 179: 2213–2228. https://doi.org/10.1534/genetics.107.085340

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

Gordon, M., 1927 The genetics of a viviparous top-minnow *Platypoecilus*; the inheritance of two kinds of melanophores. Genetics 12: 253–283.

Green, R. E., J. Krause, A. W. Briggs, T. Maricic, U. Stenzel *et al.*, 2010 A draft sequence of the Neandertal genome. Science 328: 710–722. https://doi.org/10.1126/science.1188021

Harr, B., E. Karakoc, R. Neme, M. Teschke, C. Pfeifle *et al.*, 2016 Genomic resources for wild populations of the house mouse, *Mus musculus* and its close relative *Mus spretus*. Sci. Data 3: 160075. https://doi.org/10.1038/sdata.2016.75

Harrison, R. G., 1990 Hybrid zones: windows on evolutionary process. Oxf. Surv. Evol. Biol. 7: 69–128.

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

Janoušek, V., L. Wang, K. Luzynski, P. Dufková, M. M. Vyskočilová *et al.*, 2012 Genome-wide architecture of reproductive isolation in a naturally occurring hybrid zone between *Mus musculus musculus* and *M. m. domesticus*. Mol. Ecol. 21: 3032–3047. https://doi.org/10.1111/j.1365-294X.2012.05583.x

Johnson, N. A., 2010 Hybrid incompatibility genes: remnants of a genomic battlefield? Trends Genet. 26: 317–325. https://doi.org/10.1016/j.tig.2010.04.005

Keane, T. M., L. Goodstadt, P. Danecek, M. A. White, K. Wong *et al.*, 2011 Mouse genomic variation and its effect on phenotypes and gene regulation. Nature 477: 289–294. https://doi.org/10.1038/nature10413

Kerwin, R. E., and A. L. Sweigart, 2017 Mechanisms of transmission ratio distortion at hybrid sterility loci within and between *Mimulus* species. G3 7: 3719–3730.

Larson, E. L., G. L. Hume, J. A. Andrés, and R. G. Harrison, 2012 Post-mating prezygotic barriers to gene exchange between hybridizing field crickets. J. Evol. Biol. 25: 174–186. https://doi.org/10.1111/j.1420-9101.2011.02415.x

Larson, E. L., S. Keeble, D. Vanderpool, M. D. Dean, and J. M. Good, 2017 The composite regulatory basis of the large X-effect in mouse speciation. Mol. Biol. Evol. 34: 282–295.

Li, H., 2013 Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997v2.

Lindholm, A. K., K. A. Dyer, R. C. Firman, L. Fishman, W. Forstmeier *et al.*, 2016 The ecology and evolutionary dynamics of meiotic drive. Trends Ecol. Evol. 31: 315–326. https://doi.org/10.1016/j.tree.2016.02.001

Lindtke, D., and C. A. Buerkle, 2015 The genetic architecture of hybrid incompatibilities and their effect on barriers to introgression in secondary contact. Evolution 69: 1987–2004. https://doi.org/10.1111/evo.12725

Lohse, M., A. M. Bolger, A. Nagel, A. R. Fernie, J. E. Lunn *et al.*, 2012 RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. Nucleic Acids Res. 40: W622–W627. https://doi.org/10.1093/nar/gks540

Lyon, M. F., 2003 Transmission ratio distortion in mice. Annu. Rev. Genet. 37: 393–408. https://doi.org/10.1146/annurev.genet.37.110801.143030

Maheshwari, S., and D. A. Barbash, 2011 The genetics of hybrid incompatibilities. Annu. Rev. Genet. 45: 331–355. https://doi.org/10.1146/annurev-genet-110410-132514

Matute, D. R., J. Gavin-Smyth, and G. Liu, 2014 Variable post-zygotic isolation in *Drosophila melanogaster/D. simulans* hybrids. J. Evol. Biol. 27: 1691–1705. https://doi.org/10.1111/jeb.12422

McDermott, S. R., and M. A. F. Noor, 2010 The role of meiotic drive in hybrid male sterility. Philos. Trans. R. Soc. Lond. B Biol. Sci. 365: 1265–1272. https://doi.org/10.1098/rstb.2009.0264

McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis *et al.*, 2010 The genome analysis toolkit: a mapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20: 1297–1303. https://doi.org/10.1101/gr.107524.110

Mihola, O., Z. Trachtulec, C. Vlcek, J. C. Schimenti, and J. Forejt, 2009 A mouse speciation gene encodes a meiotic histone H3 methyltransferase. Science 323: 373–375. https://doi.org/10.1126/science.1163601

Mirarab, S., and T. Warnow, 2015 ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. Bioinformatics 31: i44–i52. https://doi.org/10.1093/bioinformatics/btv234

Muller, H. J., 1942 Isolating mechanisms, evolution and temperature. Biol. Symp. 6: 71–125.

Nadeau, J. H., 2017 Do gametes woo? Evidence for their nonrandom union at fertilization. Genetics 207: 369–387.

Ng, F. L., D. Y. Liu, and H. G. Baker, 1992 Comparison of Percoll, mini-Percoll and swim-up methods for sperm preparation from abnormal semen samples. Hum. Reprod. 7: 261–266. https://doi.org/10.1093/oxfordjournals.humrep.a137628

Odet, F., W. Pan, T. A. Bell, S. G. Goodson, A. M. Stevans *et al.*, 2015 The founder strains of the collaborative cross express a complex combination of advantageous and deleterious traits for male reproduction. G3 (Bethesda) 5: 2671–2683. https://doi.org/10.1534/g3.115.020172

Oka, A., A. Mita, N. Sakurai-Yamatani, H. Yamamoto, N. Takagi *et al.*, 2004 Hybrid breakdown caused by substitution of the X chromosome between two mouse subspecies. Genetics 166: 913–924. https://doi.org/10.1534/genetics.166.2.913

Orr, A. H., and M. Turelli, 2001 The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities. Evolution 55: 1085–1094. https://doi.org/10.1111/j.0014-3820.2001.tb00628.x

Patterson, J. T., and W. S. Stone, 1952 *Evolution in the Genus Drosophila*. The Macmillan Company, New York.

Payseur, B. A., 2010 Using differential introgression in hybrid zones to identify genomic regions involved in speciation. Mol. Ecol. Resour. 10: 806–820. https://doi.org/10.1111/j.1755-0998.2010.02883.x

Payseur, B. A., and H. E. Hoekstra, 2005 Signatures of reproductive isolation in patterns of single nucleotide diversity across inbred strains of mice. Genetics 171: 1905–1916. https://doi.org/10.1534/genetics.105.046193

Pease, J. B., and M. W. Hahn, 2015 Detection and polarization of introgression in a five-taxon phylogeny. Syst. Biol. 64: 651–662. https://doi.org/10.1093/sysbio/syv023

Peterson, B. K., J. N. Weber, E. H. Kay, H. S. Fisher, and H. E. Hoekstra, 2012 Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS One 7: e37135. https://doi.org/10.1371/journal.pone.0037135

Petkov, P. M., J. H. Graber, G. A. Churchill, K. DiPetrillo, B. L. King *et al.*, 2005 Evidence of a large-scale functional organization of mammalian chromosomes. PLoS Genet. 1: e33. https://doi.org/10.1371/journal.pgen.0010033

Phelps, M. J., J. Liu, J. D. Benson, C. E. Willoughby, J. A. Gilmore *et al.*, 1999 Effects of Percoll separation, cryoprotective agents, and temperature on plasma membrane permeability characteristics of murine spermatozoa and their relevance to cryopreservation. Biol. Reprod. 61: 1031–1041. https://doi.org/10.1095/biolreprod61.4.1031

Piálek, J., M. Vyskocilová, B. Bímová, D. Havelková, J. Piálková *et al.*, 2008 Development of unique house mouse resources suitable for evolutionary studies of speciation. J. Hered. 99: 34–44. https://doi.org/10.1093/jhered/esm083

Presgraves, D. C., 2010 The molecular evolutionary basis of species formation. Nat. Rev. Genet. 11: 175–180. https://doi.org/10.1038/nrg2718

Reed, L. K., and T. A. Markow, 2004 Early events in speciation: polymorphism for hybrid male sterility in *Drosophila*. Proc. Natl. Acad. Sci. USA 101: 9009–9012. https://doi.org/10.1073/pnas.0403106101

Rieseberg, L. H., and B. K. Blackman, 2010 Speciation genes in plants. Ann. Bot. (Lond.) 106: 439–455. https://doi.org/10.1093/aob/mcq126

Sarver, B. A. J., S. Keeble, T. Cosart, P. K. Tucker, M. D. Dean *et al.*, 2017 Phylogenomic insights into mouse evolution using a pseudoreference approach. Genome Biol. Evol. 9: 726–739. https://doi.org/10.1093/gbe/evx034

Scopece, G., C. Lexer, A. Widmer, and S. Cozzolino, 2010 Polymorphism of postmating reproductive isolation within plant species. Taxon 59: 1367–1374.

Shorter, J. R., F. Odet, D. L. Aylor, W. Pan, C.-Y. Kao *et al.*, 2017 Male infertility is responsible for nearly half of the extinction observed in the mouse collaborative cross. Genetics 206: 557–572. https://doi.org/10.1534/genetics.116.199596

Shuker, D. M., K. Underwood, T. M. King, and R. K. Butlin, 2005 Patterns of male sterility in a grasshopper hybrid zone imply accumulation of hybrid incompatibilities without selection. Proc. Biol. Sci. 272: 2491–2497. https://doi.org/10.1098/rspb.2005.3242

Snyder, R. L., 1967 Fertility and reproductive performance of grouped male mice, pp. 458–472 in *Comparative Aspects of Reproductive Failure*, edited by K. Benirschke. Springer-Verlag, New York https://doi.org/10.1007/978-3-642-48949-5_26

Stamatakis, A., 2014 RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30: 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Streett, D. A., K. R. Petersen, A. T. Gerritsen, S. S. Hunter, and M. L. Settles, 2015 expHTS: analysis of high throughput sequence data in an experimental framework. BCB 15: 523–524. https://doi.org/10.1145/2808719.2811442

Sweigart, A. L., and L. E. Flagel, 2015 Evidence of natural selection acting on a polymorphic hybrid incompatibility locus in *Mimulus*. Genetics 199: 543–554. https://doi.org/10.1534/genetics.114.171819

Turner, L. M., and B. Harr, 2014 Genome-wide mapping in a house mouse hybrid zone reveals hybrid sterility loci and Dobzhansky-Muller interactions. eLife 3: e02504.

Turner, L. M., D. J. Schwahn, and B. Harr, 2012 Reduced male fertility is common but highly variable in form and severity in a natural house mouse hybrid zone. Evolution 66: 443–458. https://doi.org/10.1111/j.1558-5646.2011.01445.x

Turner, L. M., M. A. White, D. Tautz, and B. A. Payseur, 2014 Genomic networks of hybrid sterility. PLoS Genet. 10: e1004162. https://doi.org/10.1371/journal.pgen.1004162

Virdee, S. R., and G. M. Hewitt, 1994 Clines for hybrid dysfunction in a grasshopper hybrid zone. Evolution 48: 392–407. https://doi.org/10.1111/j.1558-5646.1994.tb01319.x

Vyskočilová, M., Z. Trachtulec, J. Forejt, and J. Pialek, 2005 Does geography matter in hybrid sterility in house mice? Biol. J. Linn. Soc. Lond. 84: 663–674. https://doi.org/10.1111/j.1095-8312.2005.00463.x

Vyskocilová, M., G. Pražanová, and J. Piálek, 2009 Polymorphism in hybrid male sterility in wild-derived *Mus musculus musculus* strains on proximal chromosome 17. Mamm. Genome 20: 83–91. https://doi.org/10.1007/s00335-008-9164-3

Wang, R. J., A. Cécile, and B. A. Payseur, 2013 The evolution of hybrid incompatibilities along a phylogeny. Evolution 67: 2905–2922.

White, M. A., B. Steffy, T. Wiltshire, and B. A. Payseur, 2011 Genetic dissection of a key reproductive barrier between nascent species of house mice. Genetics 189: 289–304. https://doi.org/10.1534/genetics.111.129171

Wright, K. M., D. Lloyd, D. B. Lowry, M. R. Macnair, and J. H. Willis, 2013 Indirect evolution of hybrid lethality due to linkage with selected locus in *Mimulus guttatus*. PLoS Biol. 11: e1001497. https://doi.org/10.1371/journal.pbio.1001497

Yang, H., J. R. Wang, J. P. Didion, R. J. Buus, T. A. Bell *et al.*, 2011 Subspecific origin and haplotype diversity in the laboratory mouse. Nat. Genet. 43: 648–655. https://doi.org/10.1038/ng.847

Zheng, X., D. Levine, J. Shen, S. M. Gogarten, C. Laurie *et al.*, 2012 A high-performance computing toolset for relatedness and principal component analysis of SNP data. Bioinformatics 28: 3326–3328. https://doi.org/10.1093/bioinformatics/bts606

Communicating editor: D. Barbash