Speciation and reduced hybrid female fertility in house mice

Taichi A. Suzuki[1,2,3] and Michael W. Nachman[1,2]

[1]*Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona 85721*

[2]*Current Address: Museum of Vertebrate Zoology and Department of Integrative Biology, University of California Berkeley, Berkeley, California 94720*

[3]*E-mail: taichis@berkeley.edu*

In mammals, intrinsic postzygotic isolation has been well studied in males but has been less studied in females, despite the fact that female gametogenesis and pregnancy provide arenas for hybrid sterility or inviability that are absent in males. Here, we asked whether inviability or sterility is observed in female hybrids of *Mus musculus domesticus* and *M. m. musculus*, taxa which hybridize in nature and for which male sterility has been well characterized. We looked for parent-of-origin growth phenotypes by measuring adult body weights in F1 hybrids. We evaluated hybrid female fertility by crossing F1 females to a tester male and comparing multiple reproductive parameters between intrasubspecific controls and intersubspecific hybrids. Hybrid females showed no evidence of parent-of-origin overgrowth or undergrowth, providing no evidence for reduced viability. However, hybrid females had smaller litter sizes, reduced embryo survival, fewer ovulations, and fewer small follicles relative to controls. Significant variation in reproductive parameters was seen among different hybrid genotypes, suggesting that hybrid incompatibilities are polymorphic within subspecies. Differences in reproductive phenotypes in reciprocal genotypes were observed and are consistent with cyto-nuclear incompatibilities or incompatibilities involving genomic imprinting. These findings highlight the potential importance of reduced hybrid female fertility in the early stages of speciation.

KEY WORDS: Haldane's rule, hybrid sterility, *Mus musculus*, reproductive isolation.

A detailed understanding of the process of speciation requires an understanding of the components of reproductive isolation and their relative importance. In the progeny of crosses between taxa with heteromorphic sex chromosomes, if one sex shows a reduction in fertility or viability, it is typically the heterogametic sex (Haldane 1922). The ubiquity of this pattern in diverse taxa has led to a tremendous interest in documenting the details of postzygotic reproductive isolation in males, the heterogametic sex in mammals and *Drosophila* (which are among the best-studied groups for postzygotic isolation). Much less attention has been paid to the nature of reproductive isolation in females. This is unfortunate because there are at least two major reasons why some components of postzygotic isolation might be specific to females.

First, reproduction in males and females is fundamentally different. Thus, even if the genetic basis of hybrid inviability may be similar for males and females (Orr 1993), the genetic basis of hybrid sterility may be different (Coyne 1985). In all animals, spermatogenesis and oogenesis are distinct developmental processes with genes uniquely expressed in each sex (Su et al. 2002). In viviparous animals, female fertility further depends on the successful provisioning of embryos until birth. The complicated anatomical and physiological requirements of pregnancy are unique to females and provide an arena in which hybrid incompatibilities may arise that is absent from males. Thus, there is no reason to expect that those alleles causing problems in male reproduction will be the same as those alleles causing problems in female reproduction.

Second, in viviparous animals, the intimate association between mother and offspring creates a context for genomic conflict over the on-going provisioning of embryos that is absent from oviparous animals (Zeh and Zeh 2008). In particular, conflict may
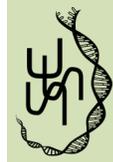
arise between mother and developing embryo, between different embryos, and between maternal and paternal genomes within an embryo (Haig 1993, 1997; Spencer et al. 1999). Such conflict may underlie the evolution of genomic imprinting, in which the expression of an allele depends on its parent of origin. In some cases of imprinting, paternal expression is associated with growth genes, whereas maternal expression is associated with repressors of growth (Ferguson-Smith 2011). If imprinting is disrupted in hybrids, hybrid inviability may result from undergrowth or overgrowth. For example, in crosses between female oldfield mice (*Peromyscus polionotus*) and male deer mice (*Peromyscus maniculatus*), imprinting is disrupted in F1 embryos leading to an overgrowth phenotype and embryonic death (Vrana et al. 1998, 2000). In this case, reproductive isolation is a consequence of F1 inviability, and it is not sex specific, but the arena for this effect is specific to females. In principle, such conflicts could also arise in later generations, although the effects are expected to be weaker (Zeh and Zeh 2008). Parent-of-origin growth effects have been documented in a number of mammalian groups (summarized in Table 1 of Brekke and Good 2014), although it is unclear whether they are generally due to imprinting defects. Some authors have argued that maternal–fetal interactions leading to hybrid inviability are likely to be the primary postzygotic barrier and even more important than hybrid sterility (Zeh and Zeh 2008), an idea that is consistent with the data from *Peromyscus* (Dawson 1965). For example, in the *Peromyscus* crosses described above, F1 inviability is seen in both directions, yet when F1 offspring survive, both males and females are fertile (e.g., Dawson 1965).

It is important to recognize that the genetic architecture of postzygotic isolation may be different in males and females. Many problems in hybrids are thought to arise from disrupted epistatic interactions between genes (Bateson 1909; Dobzhansky 1936; Muller 1942), referred to as Bateson-Dobzhansky-Muller (BDM) incompatibilities. Because males have one X chromosome whereas females have two, BDM incompatibilities involving recessive X-linked alleles will be exposed in F1 hybrid males yet masked in F1 hybrid females. In contrast, BDM incompatibilities involving dominant alleles, either on the X chromosome or on the autosomes, will be visible in hybrids of both sexes. The observation that most hybrid sterility alleles appear to act in a recessive fashion (e.g., Masly and Presgraves 2007) helps explain the prevalence of sterility in the heterogametic sex (Turelli and Orr 1995). Thus, there are more genetic paths to sterility in males compared to females (when males are heterogametic). Nonetheless, dominant interactions underlying male sterility are also known to occur (e.g., White et al. 2011) and thus dominant BDM interactions might also underlie female sterility.

House mice (*Mus musculus*) provide a useful system for studying hybrid female fertility and viability and for comparing

these to patterns seen in hybrid males. This species consists of three major lineages, *M. m. domesticus*, *M. m. musculus*, and *M. m. castaneus*, which diverged from a common ancestor approximately 0.35 million years ago (Boursot et al. 1993; Geraldes et al. 2011). *M. m. domesticus* and *M. m. musculus* form a narrow hybrid zone across Europe (Boursot et al. 1993; Sage et al. 1993; Baird and Macholán 2012). Males from the hybrid zone have reduced fertility (Albrechtova 2012; Turner et al. 2012), and male progeny of laboratory crosses between *M. m. domesticus* and *M. m. musculus* also show reduced fertility (Forejt and Ivanyi 1974; Britton-Davidian et al. 2005; Good et al. 2008a; Mihola et al. 2009; White et al. 2011).

Only one study has examined fertility in F1 hybrid females between *M. m. domesticus* and *M. m. musculus* (Britton-Davidian et al. 2005). They found reduced fertility in F1 hybrid females from a cross between the wild-derived outbred strains, *M. m. domesticus*[DDO] and *M. m. musculus*[MDH]. However, *M. m. domesticus*[DDO] is fixed for three Robertsonian translocations ($2n = 34$) whereas *M. m. musculus*[MDH] has a standard karyotype ($2n = 40$), so reduced fertility of hybrid females in this cross may be due to problems of chromosomal pairing and segregation during meiosis, rather than to BDM incompatibilities. A reanalysis of these data by Macholán et al. (2008) suggested that there might also be a slight reduction in embryonic viability in F1 females, although another recent study failed to find a reduction in fetal or placental growth in crosses between different inbred strains of *M. m. musculus* and *M. m. domesticus* (Kropackova et al. 2015). Using inbred strains with a standard karyotype (*M. m. domesticus*[WLA] and *M. m. musculus*[PWK]) Britton-Davidian et al. (2005) found no evidence for reduced fertility in hybrid females. However, the fertility of their control crosses may be underestimated due to inbreeding because they only used a single strain to represent each subspecies. Furthermore, Britton-Davidian et al. (2005) measured only a few reproductive parameters (proportion of successful pregnancies, litter size, sex ratio, and placental scars). Therefore, the extent and the importance of F1 hybrid female inviability or sterility between *M. m. domesticus* and *M. m. musculus* remain largely unknown.

Here, we assess the extent to which females contribute to reproductive isolation in crosses between *M. m. domesticus* and *M. m. musculus*. First, we look for evidence of overgrowth or undergrowth phenotypes that might reflect reduced viability in F1 females, as predicted by viviparity-driven conflict. Next, we characterize F1 fertility in detail. The hybrid female genotypes and the experimental design employed here parallel those used by Good et al. (2008a) to characterize hybrid male sterility, permitting a direct comparison between male and female components of reproduction. We found no evidence for hybrid inviability, but we found significant reductions in female fertility.

Table 1. Mean reproductive parameters for *M. m. musculus* and *M. m. domesticus*, and their F1 hybrids (*n* = 106).

F1 genotype (female × male)	Successful pregnancies[1] %	Body weight g (SE)	ROW[2] mg/g (SE)	Smaller follicles ×10² (SE)	Larger follicles ×10² (SE)	Corpora lutea (SE)	Placental scars (SE)	Embryo survival[3] % (SE)	Litter size (SE)	Individual pup weight g (SE)	Total litter weight g (SE)
Intrasubspecific (control) crosses											
M. m. musculus^PWK × *M. m. musculus*^CZECH	100 (11/11)	12.6 (0.3)	0.45 (0.03)	69.6 (4.1)	8.0 (1.0)	20.8 (2.3)	8.1 (0.3)	92.7 (4.4)	7.4 (0.3)	7.8 (0.1)	56.5 (2.9)
M. m. domesticus^LEWES × *M. m. domesticus*^WSB	90 (9/10)	13.8 (0.1)	0.35 (0.01)	84.8 (5.5)	6.5 (0.7)	21.5 (1.4)	6.2 (0.7)	90.1 (2.5)	5.6 (0.7)	10.7 (0.1)	67.2 (2.6)
Combined controls	95 (20/21)	13.2 (0.2)	0.41 (0.01)	77.6 (3.8)	7.2 (0.6)	21.2 (1.3)	7.2 (0.4)	91.5 (2.6)	6.5 (0.4)	9.1 (0.1)	61.8 (2.2)
Intersubspecific (hybrid) crosses											
M. m. musculus^PWK × *M. m. domesticus*^LEWES	91 (10/11)	13.2 (0.4)	0.35 (0.02)	65.1 (6.6)	7.2 (1.3)	13.7 (2.0)↓	6.0 (0.6)↓	95.8 (2.8)	5.7 (0.6)	8.9 (0.1)	49.9 (3.3)↓[4]
M. m. domesticus^LEWES × *M. m. musculus*^PWK	90 (9/10)	13.6 (0.5)	0.33 (0.02)↓[4]	55.2 (4.5)↓[4]	6.9 (0.6)	15.6 (1.6)↓	6.6 (0.8)	78.9 (5.4)↓	5.2 (0.7)	9.4 (0.1)	56.4 (3.8)
M. m. musculus^PWK × *M. m. domesticus*^WSB	90 (9/10)	12.1 (0.4)↓	0.41 (0.02)	52.4 (2.8)↓[4]	5.4 (0.5)	14.3 (1.5)↓	5.9 (0.2)↓[4]	70.0 (9.3)↓	4.2 (0.6)↓[4]	9.7 (0.2)	43.5 (4.1)↓[4]
M. m. domesticus^WSB × *M. m. musculus*^PWK	90 (9/10)	14.7 (0.5)↑	0.37 (0.01)	60.1 (4.2)↓	7.3 (0.6)	19.0 (1.6)	7.8 (1.0)	73.7 (5.1)↓[4]	5.7 (0.8)	9.9 (0.1)↑	60.3 (6.3)
M. m. musculus^CZECH × *M. m. domesticus*^LEWES	58 (7/12)	12.9 (0.2)	0.37 (0.04)	37.1 (3.9)↓[4]	7.8 (0.8)	14.2 (2.3)↓	4.4 (1.2)	98.4 (1.6)	4.3 (1.1)	9.4 (0.1)	53.7 (8.3)
M. m. domesticus^LEWES × *M. m. musculus*^CZECH	90 (9/10)	12.7 (0.2)	0.43 (0.02)	56.1 (5.4)↓[4]	9.3 (0.6)	17.2 (1.6)	6.6 (0.8)	79.2 (9.8)	5.6 (0.8)	9.3 (0.1)	60.1 (6.4)
M. m. musculus^CZECH × *M. m. domesticus*^WSB	78 (7/9)	12.4 (0.6)↓	0.40 (0.02)	39.8 (2.9)↓[4]	6.2 (0.8)	16.7 (2.6)	7.3 (0.3)	66.3 (13.1)↓	4.9 (1.0)	9.8 (0.1)	59.8 (3.3)
M. m. domesticus^WSB × *M. m. musculus*^CZECH	100 (13/13)	13.2 (0.4)	0.45 (0.01)↑	52.2 (4.2)↓[4]	8.9 (0.7)	21.8 (1.9)	8.3 (0.5)	77.7 (5.2)↓	6.4 (0.5)	10.6 (0.1)↑[4]	65.8 (5.2)
Combined hybrids	86 (73/85)	13.1 (0.2)	0.38 (0.01)	52.1 (1.8)↓[4]	7.5 (0.3)	16.6 (0.7)↓	6.6 (0.3)	79.5 (2.8)↓	5.3 (0.3)↓	9.7 (0.1)↑[4]	56.5 (1.9)

[1]Successful pregnancies; number of females that reproduced/total number of females.

[2]Relative ovary weight; milligrams of ovary/gram of body weight.

[3]Embryo survival; litter size at birth/placental scars.

[4]Comparisons that remained significant after Bonferroni correction α = 0.0056.

↑ or ↓ indicate significantly larger or smaller versus the combined intraspecific controls based on Wilcoxon rank sum test based on raw *P*-values (*P* < 0.05).

SE = standard error.

Materials and Methods

ANIMALS

Four wild-derived inbred strains were used in this study and were purchased from the Jackson Laboratory (Bar Harbor, ME). *M. m. domesticus*[LEWES/EiJ] and *M. m. domesticus*[WSB/EiJ] were used to represent *M. m. domesticus* (hereafter *M. m. domesticus*[LEWES] and *M. m. domesticus*[WSB]). *M. m. musculus*[PWK/PhJ] and *M. m. musculus*[CZECHII/EiJ] were used to represent *M. m. musculus* (hereafter *M. m. musculus*[PWK] and *M. m. musculus*[CZECH]). Detailed information on strain history is available from the Jackson Laboratory (www.jax.org).

EXPERIMENTAL DESIGN

Crosses were conducted in two steps. First, intrasubspecific (control) and intersubspecific (hybrid) F1 females were generated (Fig. 1). Because these strains are inbred, two strains of the same subspecies were crossed to control for the effects of inbreeding depression. This approach assumes that heterotic effects, if any, will be similar within and between subspecies. For the *M. m. musculus* control, *M. m. musculus*[PWK] females and *M. m. musculus*[CZECH] males were crossed. For the *M. m. domesticus* control, *M. m. domesticus*[LEWES] females and *M. m. domesticus*[WSB] males were crossed. For the F1 hybrid genotypes, all eight possible pairwise combinations were performed using two strains per subspecies as in Good et al. (2008a). A total of 21 females were used for controls and 85 females were used for intersubspecific hybrids. The average sample size for each genotype was 10.6; details are in Table S1.

Second, fertility of control and hybrid females was measured by crossing females to a tester male (*M. m. musculus*[PWK]). A single F1 female was paired with a single tester male in a fresh cage. Pairs were kept together for seven days and then separated. All F1 females were paired at 63 days of age. Gestation lasts 21 days. Pups were weaned at 21 days after birth. The mothers were then sacrificed (105–112 days old). Ovaries and the uterus from the mother were collected and used for measuring reproductive parameters. Females with and without litters were dissected during the same age window.

M. m. musculus[PWK] were chosen as the tester males because they are known to be good breeders and to mate readily with con-subspecifics, hetero-subspecifics, and with hybrids (Good et al. 2008a,b). All tester males were crossed to a female at least once to confirm fertility before being used in the experiment. Tester males varied in age from 56 to 240 days (paternal age mean $\pm$ SE = 158.2 $\pm$ 4.9 days). Although there are no changes in age-related testis morphology in mice until around 360 days (Tane-mura et al. 1993), we randomized the age of the tester male across female genotypes to avoid potential paternal age effects. There was no significant correlation between paternal age and female fertility ($r^2 = 0.009$, $P = 0.34$, see Fig. S1). All tester males were given a minimum four days refractory period to restore sperm between matings.

Additional crosses were performed to test for delayed puberty in female hybrids. Two hybrid genotypes were generated (*M. m. musculus*[PWK]/*M. m. domesticus*[LEWES] and *M. m. domesticus*[LEWES]/*M. m. musculus*[PWK]) for a total of 22 mice for this test. To test their fertility, we repeated the experiments described above using 120-day-old mice instead of using 63-day-old mice (Table S2).

FEMALE VIABILITY AND FERTILITY PARAMETERS

To look for evidence of overgrowth or undergrowth phenotypes, body weight of control and hybrid females was measured before mating at 63 days. Following mating, several fertility parameters were measured. Number of pups (i.e., litter size) was measured at birth (<2 days) with minimum disruption. Individual pup weight and total litter weight were measured at 21 days. Relative ovary weight (i.e., ovary weight divided by body weight), number of placental scars, and number of follicles were examined in 105-to 112-day-old females. Animals included in the study of delayed puberty were sacrificed and reproductive parameters were measured at 162–169 days. The proportion of females that gave birth for each genotype was calculated as the number of females that reproduced divided by the total number of females that were paired. Embryo survival rate (i.e., the proportion of embryos that survived to birth) was estimated as the litter size divided by number of placental scars. Embryo survival rate was not calculated for females with no placental scars. Total litter weight was not calculated for females without litters.

OVARY HISTOLOGY AND DETECTION OF APOPTOSIS

Eggs develop from follicles in the ovary. Following ovulation, the ruptured follicle develops into a corpus luteum. Thus, counting the number of follicles at different stages of development provides an assessment of female fertility. The number of corpora lutea provides an estimate of the number of eggs that were ovulated. To make these measurements, both right and left ovaries were fixed in Bouin's fixative solution (Ricca Chemical Company, Arlington, TX) for 4–6 h. After fixation, ovaries were dehydrated, embedded in Paraplast (Fisher Scientific, Houston, TX), serially sectioned (5 μm), mounted on glass slides, and stained by standard Hematoxylin–Eosin staining. Every 40th section per ovary (approximately six to 10 sections per ovary) was used to estimate the total number of primordial, small primary, large primary, secondary, and antral follicles per ovary. In addition, two additional categories were defined: "smaller follicles" include both primordial and small primary follicles, whereas "larger follicles" include secondary and antral follicles (see Table S3 for follicle

A   Cross design

B   Intrasubspecific (Control) crosses

C   Intersubspecific (Hybrid) crosses
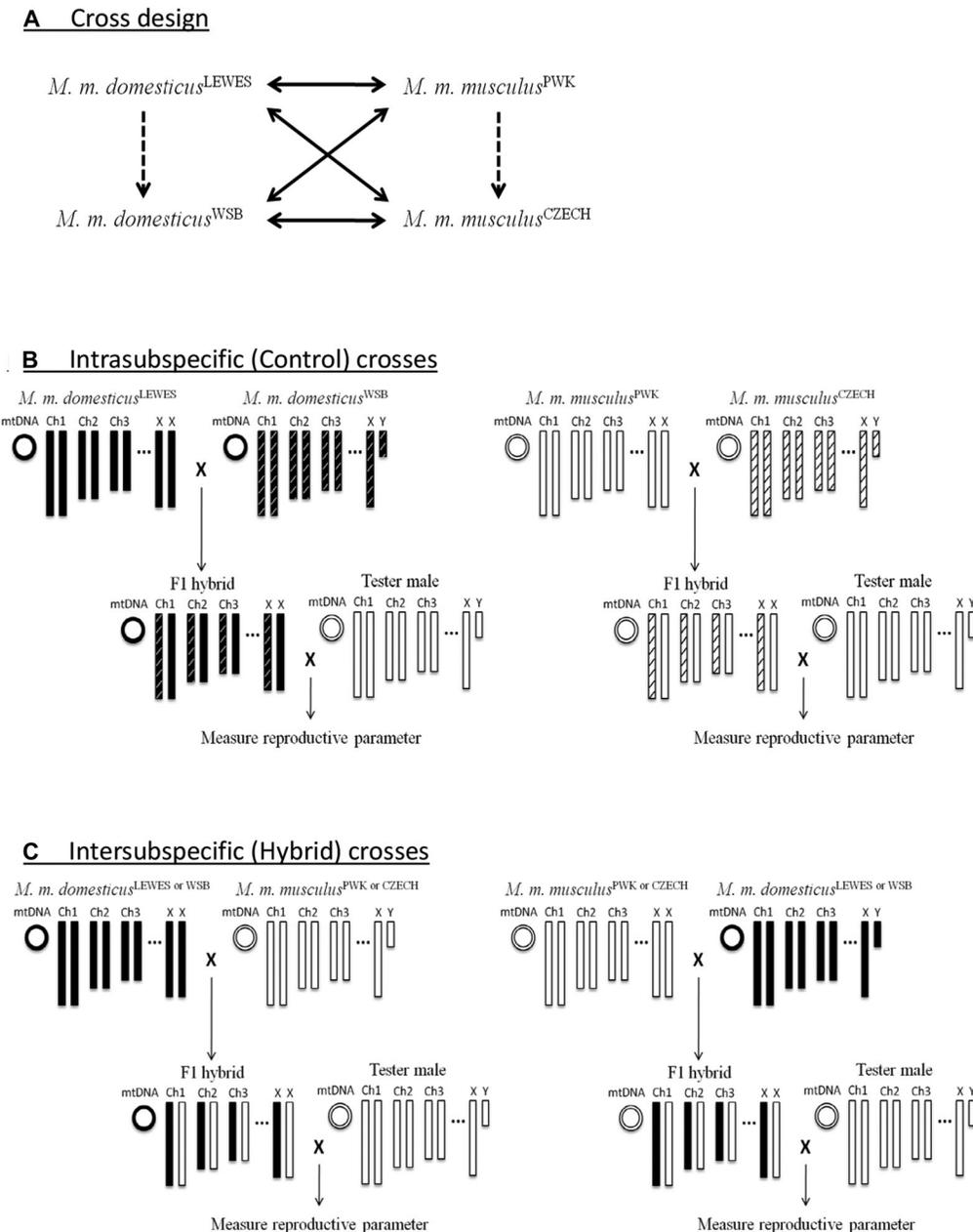
Figure 1. Experimental design and genetic composition of F1 females. (A) Schematic of intrasubspecific cross (dashed arrows) and intersubspecific cross (solid arrow). Arrows indicate direction of cross from female parent to male parent. Reciprocal crosses indicated by double head arrows. (B) Schematic of intrasubspecific (control) F1 females crossed with tester male (*M. m. musculus*[PWK]). Chromosomes and mitochondria (mtDNAs) are indicated in *M. m. domesticus*[LEWES] (black), *M. m. domesticus*[WSB] (black with white stripes), *M. m. musculus*[PWK] (white), and *M. m. musculus*[CZECH] (white with black stripes). (C) Schematic of intersubspecific (hybrid) F1 females crossed with tester male.

classification). Only follicles containing an oocyte with a clearly visible nucleus were scored to avoid double counting. All sections were scored without knowledge of the genotype. To estimate the total number of follicles per ovary, follicle number was multiplied by 40 because every 40th section was counted from each ovary (e.g., Hirshfield and Midgley 1978). Although scoring every 40th section per ovary provides only a rough estimate of the actual num-

ber of follicles per ovary, the relative numbers of follicles among different genotypes are unaffected by this calculation because it was applied uniformly to all genotypes in this study (e.g., Tilly 2003). In addition, the total number of fresh corpora lutea was scored. Newly formed corpora lutea were distinguished from older corpora lutea by larger overall size, smaller luteal cell size, and eosinophilic color (e.g., Felicio et al. 1983; Durlinger et al. 2002).

To evaluate whether rates of cell death associated with ooge-nesis differ between hybrid and control females, apoptotic cells were evaluated by TUNEL assay (In Situ Cell Death Detection kit, AP; Roche, Basel, Switzerland) following the manufacturer's protocol. Larger follicles were compared between controls (*M. m. musculus*[PWK]/*M. m. musculus*[CZECH]) and the hybrid genotype that had the smallest litter at 63 days (*M. m. musculus*[PWK]/*M. m. domesticus*[WSB]). Five 21-day-old females from each genotype were used for this assay.

DATA ANALYSIS

We tested for differences between genotypes using Wilcoxon rank sum tests. Raw *P*-values are reported as well as significance following a Bonferroni correction. A Fisher's exact test was used to compare the rates of successful pregnancies. Linear regression was performed to test for age effects of tester male on female litter size and to test for correlations between reduced litter size and other reproductive parameters (i.e., smaller follicles, corpora lutea, and embryo survival). All analyses were performed with JMP 11 (SAS Institute Inc., Cary, North Carolina).

RESULTS

HYBRID FEMALE VIABILITY

Viviparity-driven conflict predicts that disrupted interactions between mother and embryo may lead to parent-of-origin undergrowth or overgrowth phenotypes both in utero and into adulthood (Zeh and Zeh 2008), as seen in a number of mammals (Dawson 1965; Brekke and Good 2014). To test this idea, we measured adult body weight of F1 females at 63 days in control and inter-subspecific crosses. Across all crosses, the mean body weight of hybrids was 13.1 g, a value that is intermediate between the body weights of the two subspecies, *M. m. musculus* (12.6 g) and *M. m. domesticus* (13.8 g; Table 1). When the mother was *M. m. domesticus*, the mean body weight of F1 females was 13.6 g, and when the mother was *M. m musculus*, the body weight of F1 females was 12.7 g. Both values fall between the controls and thus provide no evidence for significant undergrowth or overgrowth. Moreover, these slight differences are consistent with maternal effects, where the larger mother produces slightly larger offspring.

Individual crosses show some slight variations from this overall pattern. Most noteworthy is the cross between *M. m. musculus*[PWK] and *M. m domesticus*[WSB] (Table 1). When the mother was *musculus*, F1 females were slightly smaller than controls, whereas in the reciprocal cross, F1 females were significantly larger than the controls. This might reflect a mild, strain-specific parent-of-origin growth phenotype, although we note that these differences are smaller than previously reported for other taxa. Moreover, this parent-of-origin difference is not seen in F1 males in this same cross; in fact, no parent-of-origin

overgrowth or undergrowth phenotypes are seen in the F1 males reported by Good et al. (2008a).

In some species, parent-of-origin overgrowth or undergrowth phenotypes are manifest mainly in the placenta and influence the viability of embryos. To see if there were any differences in survival of embryos, we compared litter sizes of females mated to consubspecific males and litter sizes of females mated to hetero-subspecific males. Litter sizes in these two kinds of matings were nearly identical (Table 2), suggesting that there is no difference in F1 viability during early development.

HYBRID FEMALE FERTILITY

Reproductive measures were lower in hybrid females compared to control females for several different traits (Table 1, Fig. 2). Overall, hybrid females had significantly fewer primordial follicles, fewer small primary follicles, fewer large primary follicles, fewer corpora lutea, lower embryo survival, and smaller litter size compared to control females based on raw *P*-values (Fig. S2). Although differences between hybrid and control genotypes were not significant for some traits after Bonferroni correction, there was a consistent trend toward reduced reproductive performance for nearly all traits (the main exception was individual pup weight). Among the reproductive parameters that showed a significant reduction in combined hybrids, litter size was significantly positively correlated with embryo survival ($r^2 = 0.64$, $P < 0.0001$) and with the number of corpora lutea ($r^2 = 0.49$, $P < 0.0001$), suggesting that reduced ovulation and reduced embryo survival in utero may both contribute to reduced litter size in hybrids.

Hybrid females also showed considerable heterogeneity in reproductive traits when compared to controls (Table 1). For example, the litter size of *M. m. musculus*[PWK]/*M. m. domesticus*[WSB] hybrids (mean = 4.2) was significantly reduced relative to combined controls (mean = 6.5) whereas *M. m. domesticus*[WSB]/*M. m. musculus*[CZECH] hybrids (mean = 6.4) had comparable litter size to controls.

Twelve out of 85 hybrid females (14.1%) did not produce litters whereas only one out of 21 control females (4.8%) failed to produce a litter, a difference that was not significant (Table 1, Fisher's exact test, $P = 0.5$). Nonetheless, hybrid females without litters had significantly lower values for reproductive traits compared to hybrids with litters, with smaller relative ovary weights, fewer smaller follicles, fewer larger follicles, fewer corpora lutea, and fewer placental scars suggesting that these differences contributed to reduced fertility (Table 3). In contrast, reproductive values for the single control female that failed to produce a litter were normal. The ovary histology of this single female appeared normal whereas the ovary histology of the 12 hybrid females without litters appeared abnormal (Fig. 3).

Table 2. Litter sizes in intra- and intersubspecific crosses.

Maternal strain	Paternal strain(s)	Number of litters	Litter size (SE)[1]
M. m. musuclus[PWK]	*M. m. musuclus*[CZECH]	3	7.33 (0.33)
M. m. musuclus[PWK]	*M. m. domesticus*[LEWES and WSB]	15	7.46 (0.32)
M. m. domesticus[LEWES]	*M. m. domesticus*[WSB]	5	5.80 (0.49)
M. m. domesticus[LEWES]	*M. m. musuclus*[PWK and CZECH]	20	6.05 (0.29)

[1]Litter sizes were not significantly different when *M. m. musuclus*[PWK] females were crossed to *M. m. musuclus* or *M. m. domesticus* males ($P = 1.0$) nor when *M. m. domesticus*[LEWES] females were crossed to *M. m. domesticus* or *M. m. musuclus* males ($P = 0.7$) based on Wilcoxon rank sum test.
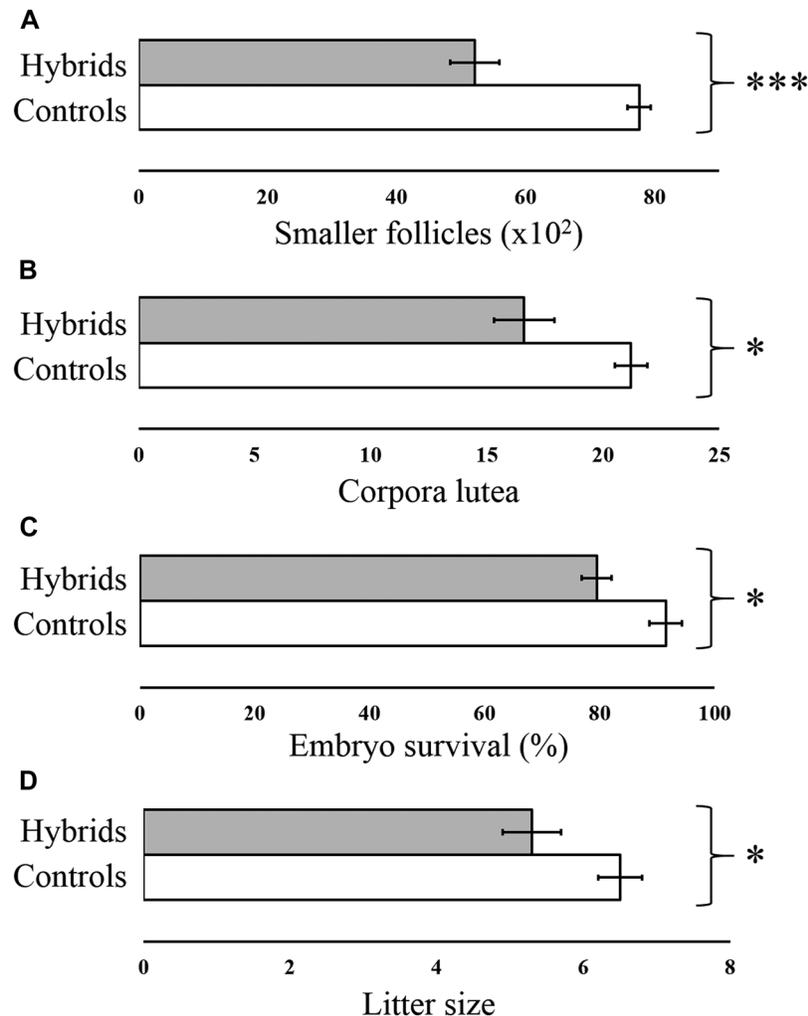
Figure 2. Mean reproductive parameters in hybrids (shaded) relative to controls (open). Wilcoxon rank sum test was performed for number of (A) smaller follicles, (B) number of corpora lutea, (C) embryo survival rate, (D) and litter size. Error bars indicate standard error. $^*P < 0.05$; $^{***}P < 0.0001$.

One potential explanation for reduced fertility in hybrid females is a developmental delay in the onset of reproductive maturity. To test for delayed puberty in hybrids, we compared fertility in 63-day-old females to fertility in 120-day-old females for a subset of hybrid genotypes. Both 120-day-old genotypes showed an increase in body weight, decrease in smaller and larger follicles, and also a trend toward lower rates of successful pregnancies, all of which are expected with aging (Table S4). Because there was no significant difference in litter size between 63-day-old and 120-day-old mice, there is no evidence of delayed puberty in the two hybrid genotypes suggesting that mice were fully mature at 63 days.

Table 3. Mean reproductive parameters for females without litters.

Cross	No. of females	Body weight g (SE)	ROW[1] mg/g (SE)	Smaller follicles $\times 10^2$ (SE)	Larger follicles $\times 10^2$ (SE)	Corpora lutea (SE)	Placental scars (SE)
Controls (produced a litter)	20	13.2 (0.2)	0.41 (0.01)	77.6 (3.8)	7.2 (0.6)	21.2 (1.3)	7.2 (0.4)
Hybrids (produced a litter)	73	13.1 (0.2)	0.38 (0.01)	52.1 (1.8)	7.5 (0.3)	16.6 (0.7)	6.6 (0.3)
Controls (produced no litter)	1	13.9 (-)	0.41 (-)	109.6 (-)	7.2 (-)	22.0 (-)	0.0 (-)
Hybrids (produced no litter)	12	12.3 (0.4)	0.29 (0.02)↓[2]	36.9 (4.2)↓[2]	5.9 (1.0)↓	6.1 (1.0)↓[2]	1.7 (0.8)↓[2]

[1]**Relative ovary weight: milligrams of ovary/gram of body weight.**

[2]**Remained significant after Bonferroni correction α = 0.0083.**

↑ or ↓ indicate significantly larger or smaller versus the combined hybrids with litters based on Wilcoxon rank sum test using raw *P*-values (*P* < 0.05).

SE = standard error.
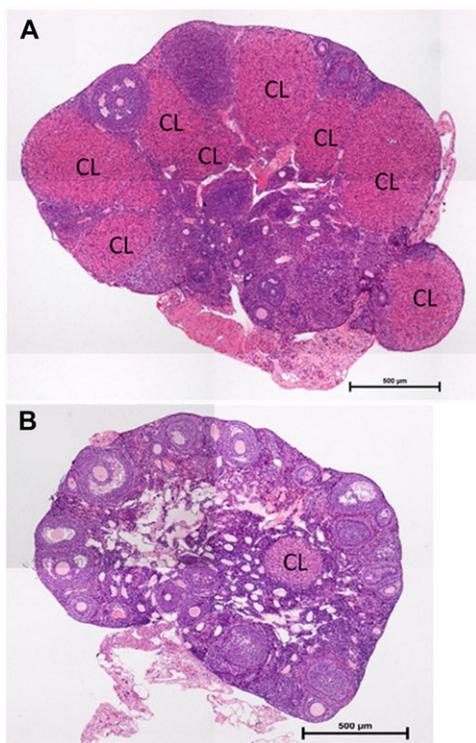
Figure 3. Photomicrographs of ovarian histology showing fewer fresh corpora lutea (CL) in hybrids without litters. **(A)** Control female (*M. m. domesticus*^{LEWES} x *M. m. domesticus*^{WSB}) without a litter. **(B)** Hybrid female (*M. m. domesticus*^{LEWES} x *M. m. musculus*^{PWK}) without a litter.

HYBRID FEMALE FERTILITY DIFFERS AMONG GENOTYPES

Asymmetry of hybrid female reproduction

Heterogeneity among crosses in hybrid female reproductive performance can be used to answer questions concerning the genetic basis of reproductive incompatibilities. For example, if reproductive parameters are different between reciprocal genotypes, the pattern might be explained by cyto-nuclear incompatibility or parent-of-origin expression (i.e., imprinting). We observed variation in hybrid female fertility in some reciprocal genotypes (Tables 1 and S5). For example, hybrid females from *M. m. musculus*^{PWK} × *M. m. domesticus*^{WSB} and *M. m. domesticus*^{WSB} × *M. m. musculus*^{PWK} differed significantly in body weight ($P = 0.003$), placental scars ($P = 0.005$), and total litter weight ($P = 0.03$, Table S5). Moreover, the litter size of *M. m. musculus*^{PWK}/*M. m. domesticus*^{WSB} hybrids (mean $\pm$ SE $= 4.2 \pm 0.6$) was significantly reduced relative to controls (mean $\pm$ SE $= 6.5 \pm 0.4$, $P = 0.0007$), whereas progeny from the reciprocal cross had litter sizes comparable to controls (mean $\pm$ SE $= 5.7 \pm 0.8$, $P = 0.14$). All other reciprocal comparisons also yielded significant differences for at least two reproductive traits (Table S5). For example, individual pup weight was significantly different in the comparison between *M. m. musculus*^{PWK}/*M. m. domesticus*^{LEWES} hybrids and *M. m. domesticus*^{LEWES}/*M. m. musculus*^{PWK} hybrids ($P = 0.006$), and also between *M. m. musculus*^{CZECH}/*M. m. domesticus*^{WSB} hybrids and *M. m. domesticus*^{WSB}/*M. m. musculus*^{CZECH} hybrids ($P < 0.0001$, Table S5).

Polymorphism of reproductive incompatibilities

Comparisons between progeny of crosses in which one parental strain is the same but the other parental strain is different within subspecies can provide insight into polymorphism for hybrid sterility alleles. For example, litter size and embryo survival were reduced in *M. m. musculus*^{PWK}/*M. m. domesticus*^{WSB} hybrids compared to *M. m. musculus*^{PWK}/*M. m. domesticus*^{LEWES} hybrids (Tables 1 and S6). Differences in fertility of these hybrids indicate that *M. m. domesticus* is polymorphic for alleles contributing to reduced litter size or embryo loss when *M. m. musculus*^{PWK} is the mother. Interestingly, in the reciprocal crosses (i.e., when *M. m. domesticus* is the mother), there is no evidence of polymorphism within *M. m. domesticus* for litter size nor embryo survival suggesting that the polymorphic alleles depend on the parent of origin (Table S7). Similarly, we observed significant polymorphisms within *M. m. musculus* for alleles contributing to body weight, relative ovary weight, smaller follicles, larger follicles,

placental scars, and individual pup weight but not for litter size (Tables S6 and S7). It should be noted that some of this variation may simply reflect strain-specific reproductive differences rather than polymorphism for BDM incompatibilities. However, some polymorphic alleles must be contributing to BDM incompatibilities because most of the reproductive parameters are lower in hybrids compared to controls (Table 1).

Discussion

Crosses between many different species of mammals reveal that hybrid male sterility is much more common than hybrid female sterility; as a group, mammals clearly obey Haldane's rule (Gray 1972). Because of this widespread pattern, there have been fewer studies of postzygotic reproductive isolation in female hybrids. Nonetheless, hybrid females might be important in reproductive isolation, either due to reduced viability as a consequence of disrupted maternal–fetal interactions or due to reduced fertility as a consequence of disrupted epistasis involving genes specific to female reproduction. We found no evidence for reduced viability in hybrid females, but the results presented here demonstrate a small but significant reduction in hybrid female fertility. Below we discuss the viviparity-conflict hypothesis in light of our results, the potential physiological and genetic mechanisms that might lead to reproductive problems in hybrid females, differences between male and female hybrid sterility, and the contribution of hybrid females to reproductive isolation in the early stages of speciation.

HYBRID FEMALE VIABILITY AND VIVIPARITY-DRIVEN CONFLICT

Parent-of-origin growth effects in hybrids are known from at least four orders of mammals (summarized in Brekke and Good 2014) and thus might be fairly common. In some cases, the observed effects are very large. For example, in crosses between female oldfield mice and male deer mice, overgrowth phenotypes in embryos are so severe as to frequently result in maternal death, while the reciprocal cross yields viable and fertile offspring that are 25% smaller than the parents at six months of age (Dawson 1965). Based on observations such as this, Zeh and Zeh (2008) have argued that hybrid inviability should contribute more to postzygotic isolation than hybrid sterility in viviparous taxa. This effect is expected to be particularly pronounced in polyandrous species, where conflict can arise among embryos with different fathers, a situation that is common in house mice (Dean et al. 2006). Despite these predictions, we found no evidence for consistent overgrowth or undergrowth phenotypes in hybrids, suggesting that hybrid sterility in both sexes is more important than hybrid inviability in causing reproductive isolation. Our results are consistent with another recent study that showed no differences in fetal or placental growth between intrasubspecific and intersub-

specific crosses using different strains of *M. m. musculus* and *M. m. domesticus* (Kropackova et al. 2015). These findings are noteworthy because this same species (*M. musculus*) is involved in parent-of-origin growth effects in hybrids when crossed to *Mus spretus*, a slightly more distantly related species with which *M. musculus* is broadly sympatric and sometimes hybridizes (Zechner et al. 1996). Similar effects are also seen in crosses between *M. musculus* and *Mus specilegus* or *Mus macedonicus* (Zechner et al. 1996). In all of these crosses, the parent-of-origin growth effects are manifest in the development of the placenta and result in lowered viability of embryos. Here, we only measured the adult weight of F1 hybrids, so we could not evaluate whether subtle effects might exist in the development of the placenta. However, if differences in placental development exist, it is unlikely that they have a large effect on hybrid embryonic viability because litter sizes for females mated to hetero-subspecific males were indistinguishable from litter sizes of females mated to consubspecific males (Table 2).

THE PHYSIOLOGY OF REDUCED FERTILITY IN HYBRID FEMALES

In contrast to the normal viability of F1 hybrid females, the fertility of F1 hybrid females was slightly reduced. Although 95% of control females successfully produced litters, only 86% of hybrid females produced a litter. Moreover, the average litter size for hybrid females was 5.3 whereas the average litter size for the controls was 6.5, a reduction of 18%. The smaller litter sizes observed in hybrid female mice could, in principal, be explained by several physiological mechanisms, including reduced number of smaller follicles, lower rates of ovulation, greater embryo loss in utero, and/or delays in the developmental onset of puberty.

Hybrids showed a significant reduction in the number of smaller follicles (Figs. 2, S2). The greatest reduction was observed in hybrid females that failed to produce a litter (Table 2). However, the association between the number of smaller follicles and litter size among females was weak ($r^2 = 0.06$, $P = 0.006$). Because primordial follicles eventually give rise to larger follicles from which eggs are released during ovulation and because primordial follicles also produce hormones critical to follicle development and fertility (Kevenaar et al. 2006), one might expect a direct correlation between the number of early follicles and litter size. However, the actual relationship may be more complicated. For example, the strength of the correlation between the number of follicles and the reproductive output is strain specific in laboratory mice (Ratts et al. 1995; Rucker et al. 2000) and species specific in viviparous lizards (Mendez De La Cruz et al. 1993). Thus, the consequence of fewer small follicles for fitness likely depends on many factors.

The number of corpora lutea provides a more direct assessment of fertility. Fewer ovulations will reduce the number of potential oocytes to be fertilized and may result in reduced litter size. Several observations suggest that lowered rates of ovulation contribute to reduced fertility in hybrid females: there was a significant reduction of corpora lutea in hybrids relative to controls (Table 1, Fig. 2), a significant reduction of corpora lutea in hybrid females without litters relative to hybrid females with litters (Table 3, Fig. 3), and a significant correlation between number of corpora lutea and litter size ($r^2 = 0.49$, $P < 0.0001$). We hypothesized that increased cell death after the development of secondary follicles in hybrids might explain reduced rates of ovulation. However, apoptotic cells, as measured by TUNEL assays, were not significantly different between hybrids and controls in 21-day-old mice (Fig. S3). We also found no evidence that delayed puberty contributes to reduced ovulation in hybrid females (Table S4), although delayed puberty has been reported in some hybrid males (Flachs et al. 2014). Further investigation is required to reveal the physiological cause of reduced ovulation and the contribution to reduced litter size.

Reduced embryo survival in utero is another potential explanation for reduced litter size in hybrids (Fig. 2). Because the number of placental scars formed after implantation was not significantly different between hybrids and controls (Table 1), postimplantation embryo mortality may contribute to reduced litter size. The death of early embryos could be due to incompatibilities in the mother or in the embryos or due to incompatibilities between the mother and the embryos (discussed below).

In summary, a subtle yet significant reduction in litter size was observed in F1 hybrid female mice, and this may reduce the fitness of hybrids in natural populations. The underlying physiological mechanisms causing reduced litter size are undoubtedly complex but are likely to include a combination of reduced ovulation during each estrus cycle and increased embryo loss in utero. A reduction in the number of smaller follicles may alter the production of hormone synthesis (Kevenaar et al. 2006) and thus could also contribute to both ovulation and implantation problems.

THE GENETICS OF REDUCED FERTILITY IN HYBRID FEMALES

In *Drosophila*, recessive incompatibilities outnumber dominant ones (Presgraves 2003; Masly and Presgraves 2007) and the same trend is apparent in house mice (White et al. 2011). Because the genomes of F1 hybrid females are completely heterozygous, neither recessive X-linked nor recessive autosomal hybrid incompatibilities are expected to affect the phenotypes measured here. Thus, alternative genetic models must be considered to explain the reduced fertility seen in F1 hybrid females. These include incompatibilities with dominant–dominant interactions, cyto-nuclear incompatibilities, and incompatibilities involving imprinted genes. Some phenotypes, such as the reduction in embryo survival in the backcross offspring of F1 females could be due to incompatibilities underlying fertility in the mother, incompatibilities underlying viability in the offspring due to recessive alleles (which are not exposed in the F1s), or maternal–fetal interactions that involve recessive alleles in the offspring.

Under the dominance theory, a minority of sterility factors is expected to be partially dominant (Turelli and Orr 1995). If most of the reduction in female fertility is due to interactions between nuclear dominant–dominant incompatibilities, reduced fertility should be seen in reciprocal genotypes. This was mostly the case for the number of smaller follicles and to a lesser extent for corpora lutea but not for the other traits (Table 1). The occurrence of dominant–dominant incompatibilities for hybrid lethality seems to be rare in *Drosophila* (Presgraves 2003 but see Barbash et al. 2000). The occurrence of dominant–dominant incompatibilities for hybrid sterility, however, is not well characterized. Thus, disruptive epistatic interactions between dominant–dominant loci may explain some (e.g., follicles, corpora lutea, etc.) but not all of the observed patterns.

Cytoplasmic effects on hybrid female sterility have been proposed in several studies (Orr 1987; Orr and Coyne 1989; Davis et al. 1994; Slotman et al. 2005; Lee et al. 2008; Meiklejohn et al. 2013). Because the nuclear genome of F1 hybrid females is fully heterozygous, reciprocal genotypes differ only by their cytoplasm (i.e., mtDNA). Moreover, mtDNA in mammals evolves quickly and encodes proteins that interact closely with many nuclear-encoded proteins, a situation that may be conducive to BDM incompatibilities. Orr (1987) suggested an effect of cytoplasm on female sterility in crosses between *D. pseudoobscura* and *D. persimilis* by showing that crosses in one direction produce F1 hybrid females with significantly reduced fertility, whereas F1 hybrids from the reciprocal cross are fully fertile. Similar patterns were seen in this study; asymmetry in hybrid female fertility was seen across reciprocal crosses (Table S5). This suggests that cyto-nuclear interactions and/or imprinting (see below) may contribute to the reduction in hybrid female fertility. Recently, a mitochondrial–nuclear incompatibility was shown to delay development time and to reduce female fecundity by 50% in *Drosophila* (Meiklejohn et al. 2013). Moreover, cyto-nuclear incompatibilities are known to cause hybrid problems in mammals (reviewed in St. John et al. 2004), insects (reviewed in Rousset and Raymond 1991), copepods (Burton 1990; Edmands and Burton 1999; Rawson and Burton 2002), and plants (reviewed in Saumitou-Laprade et al. 1994). The role of mitochondrial–nuclear interactions in mice could be tested by measuring mitochondrial function in these hybrids, as has been done in copepods (Ellison and Burton 2006).

Although we found no evidence for overgrowth or undergrowth phenotypes in F1s, consistent with Kropackova et al.

(2015), a role for imprinting in some of the phenotypes cannot be completely ruled out. The fact that asymmetry is seen in some crosses suggests that either cyto-nuclear interactions or imprinting may be involved. Widespread disruption of imprinting seems highly unlikely because the growth of F1 females appears normal. There is one other way by which imprinting may contribute to speciation: because imprinted genes show expression of only one allele, they are effectively hemizygous and thus may expose recessive alleles that can participate in hybrid incompatibilities (much like the X chromosome in males).

Finally, the smaller litter size of F1 females might partly reflect decreased viability of their embryos due to genetic interactions that are not present in F1s. All F1 females were crossed to a tester *M. m. musculus*[PWK] male, so the backcross progeny could have incompatibilities between recessive *M. m. musculus* alleles and dominant *M. m. domesticus* alleles. Such interactions could contribute to the observed postimplantation embryo mortality and reduced litter size.

POLYMORPHISM OF INCOMPATIBILITIES IN HYBRID MALES AND FEMALES

Intrasubspecific variation in hybrid female fertility in crosses involving both *M. m. musculus* and *M. m. domesticus* was observed (Tables S6 and S7), suggesting that hybrid incompatibilities are polymorphic within both subspecies. Polymorphism in hybrid fertility is consistent with early stages in the evolution of reproductive isolation (Cutter 2012) and has been observed previously in crosses between these taxa. For example, incompatibilities in hybrid male sterility in house mice are known to be polymorphic within both subspecies, and allelic variation has been shown in classic inbred strains, wild-derived *M. m. musculus*, and multiple populations of *M. m. musculus* (Forejt and Ivanyi 1974; Good et al. 2008a; Vyskočilová et al. 2009). Using the same genotypes as in this study, Good et al. (2008a) found polymorphism for hybrid male sterility within *M. m. musculus* but not within *M. m. domesticus*. Interestingly, we found polymorphism in hybrid female fertility in both *M. m. musculus* and *M. m. domesticus*. We also found that *M. m. domesticus*[WSB] seems to have more reproductive incompatibilities with *M. m. musculus* compared to *M. m. domesticus*[LEWES] (Table S8). Consistent with Haldane's rule, male hybrids showed a greater reduction in fertility compared to controls than is seen in females (i.e., 39% reduction in relative testes weight vs. 2% reduction in relative ovary weight, and 74% reduction in sperm count vs. 22% reduction in counts of corpora lutea, which approximate the number of eggs ovulated; Good et al. 2008a; Table S8). This indicates that some hybrid sterility factors differ between males and females and is consistent with mapping experiments showing hybrid male and female sterility involve different loci in *Drosophila* (Orr and Coyne 1989). No sex bias

was observed among these progeny (207 females and 204 males) providing no evidence for sex-specific embryonic mortality.

The question of whether alleles involved in incompatibilities become fixed by genetic drift or natural selection is important for understanding the evolution of reproductive isolation (Coyne and Orr 2004). Although transient polymorphism will exist regardless of whether mutations are fixed by drift or selection, it is expected to be very rare under strong selection. In contrast, polymorphism is expected to be more common if incompatible alleles are fixed by drift (Shuker et al. 2005). The observation of polymorphic hybrid sterility in both *M. m. musculus* and *M. m. domesticus* (Tables S6 and S7) raises the possibility that sojourn times for alleles underlying incompatibilities are longer than might be expected under simple models of selection on unconditionally beneficial mutations.

EVOLUTION OF REPRODUCTIVE ISOLATION

Has reduced fertility of hybrid females played a role in the evolution of reproductive isolation between *M. m. musculus* and *M. m. domesticus*? Although F1 hybrids are extremely rare or absent in the current hybrid zone (e.g., Wang et al. 2011), characterizing F1 hybrid fertility in the laboratory helps to understand the dynamics that took place during the initial period of secondary contact. It is well established that hybrid male sterility is a strong cause of isolation between these two subspecies, and possibly stronger than other aspects of fitness in hybrids such as higher parasite load in hybrids (Sage et al. 1986, but see Baird et al. 2012) or behavioral barriers to gene flow (Laukaitis et al. 1997; Smadja and Ganem 2002; Smadja et al. 2004; Bímová et al. 2011). The reduction in hybrid female fertility observed here is weaker than the hybrid male sterility observed by Good et al. (2008a), consistent with Haldane's rule (Table S8). However, even small reductions in hybrid female fertility might be important in reproductive isolation. For example, during the initial stages of secondary contact, gene flow between subspecies could be mediated by backcrosses of F1 females to either parent (because F1 males are often sterile). The results presented here indicate that such backcrosses will result in 18% smaller litter size on average and thus may provide an additional impediment to gene flow. Reductions in hybrid female fertility may also be more pronounced in later generations when recessive incompatibilities are exposed. In fact, this has been observed in crosses between the inbred strain C57BL/6 and *M. m. molossinus*[MOM]; in this case, F1 hybrids are fully fertile, but F2 hybrid females show low fertility due to defects in developing blastocysts (Niwa-Kawakita 1994). Such hybrid breakdown of female fertility is often seen in insects (e.g., Davis et al. 1994: *D. simulans* clade). Because most mice in the hybrid zone today are the product of late generation backcrosses or intercrosses, detailed studies of F2 or later generations in both males and females

would help us better understand the relative importance of female sterility in reproductive isolation in nature.

ACKNOWLEDGMENTS

LITERATURE CITED

Albrechtova, J., T. Albrecht, S. J. E. Baird, M. Macholán, G. Rudolfsen, P. Munclinger, P. K. Tucker, and J. Pialek. 2012. Sperm-related phenotypes implicated in both maintenance and breakdown of a natural species barrier in the house mouse. Proc. R. Soc. B. 279:4803–4810.

Baird, S. J. E., and M. Macholán. 2012. What can the *Mus musculus musculus/M. m. domesticus* hybrid zone tell us about speciation? Pp. 334–372 *in* M. Macholán, S. J. E. Baird, P. Munclinger, and J. Piálek, eds. Evolution of the house mouse. Cambridge Univ. Press, Cambridge, U . K.

Baird, S. J. E., A. Ribas, M. Macholán, T. Albrecht, J. Piálek, and J. Goüy de Bellocq. 2012. Where are the wormy mice? A re-examination of hybrid parasitism in the European house mouse hybrid zone. Evolution 66:2757–2772.

Barbash, D. A., J. Roote, and M. Ashburner. 2000. The *Drosophila melanogaster Hybrid male rescue* gene causes inviability in male and female species hybrids. Genetics 154:1747–1771.

Bateson, W. 1909. Heredity and variation in modern lights. Pp. 85–101 *in* A. C. Seward, ed. Darwin and modern science. Cambridge Univ. Press, Cambridge, U . K.

Bímová, B. V., M. Macholán, S. J. E. Baird, P. Munclinger, P. Dufková, C. M. Laukaitis, R. C. Karn, K. Luzynski, P. K. Tucker, and J. Piálek. 2011. Reinforcement selection acting on the European house mouse hybrid zone. Mol. Ecol. 20:2403–2424.

Boursot, P., J.-C. Auffray, J. Britton-Davidian, and F. Bonhomme. 1993. The evolution of house mice. Annu. Rev. Ecol. Syst. 24:119–152.

Brekke, T. D., and J. M. Good. 2014. Parent-of-origin growth effects and the evolution of hybrid inviability in dwarf hamsters. Evolution 68:3134–3148.

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

Burton, R. S. 1990. Hybrid breakdown in developmental time in the copepod *Tigriopus californicus*. Evolution 44:1814–1822.

Coyne, J. A. 1985. The genetic basis of Haldane's rule. Nature 314:736–738.

Coyne, J. A., and H. A. Orr. 2004. Speciation. Sinauer Associates, Sunderland, MA.

Cutter, A. D. 2012. The polymorphic prelude to Bateson-Dobzhansky-Muller incompatibilities. Trends Ecol. Evol. 27:209–218.

Davis, A. W., E. G. Noonburg, and C.-I. Wu. 1994. Evidence for complex genic interactions between conspecific chromosomes underlying hybrid female sterility in the *Drosophila simulans* clade. Genetics 137:191–199.

Dawson, W. D. 1965. Fertility and size inheritance in a *Peromyscus* species cross. Evolution 19:44–55.

Dean, M. D., K. G. Ardlie, and M. W. Nachman. 2006. The frequency of multiple paternity suggests that sperm competition is common in house mice (*Mus domesticus*). Mol. Ecol. 15:4141–4151.

Dobzhansky, T. 1936. Studies on hybrid sterility. II. Localization of sterility factors of sterility factors in *Drosophila pseudoobscura* hybrids. Genetics 21:113–135.

Durlinger, A. L. L., M. J. G. Gruijters, P. Kramer, B. Karels, H. A. Ingraham, M. W. Nachtigal, J. Th. J. Uilenbroek, J. A. Grootegoed, and A. P. N. Themmen. 2002. Anti-Müllerian hormone inhibits initiation of primordial follicle growth in the mouse ovary. Endocrinology 143:1076–1084.

Edmands, S., and R. S. Burton. 1999. Cytochrome *C* oxidase activity in interpopulation hybrids of a marine copepod: a test for nuclear-nuclear or nuclear-cytoplasmic coadaptation. Evolution 53:1972–1978.

Ellison, C. K., and R. S. Burton. 2006. Disruption of mitochondrial function in interpopulation hybrids of *Tigriopus californicus*. Evolution 60:1382–1391.

Felicio, L. S., J. F. Nelson, R. G. Gosden, and C. E. Finch. 1983. Restoration of ovulatory cycles by young ovarian grafts in aging mice: potentiation by long-term ovariectomy decreases with age. Proc. Natl. Acad. Sci. USA 80:6076–6080.

Ferguson-Smith, A. C. 2011. Genomic imprinting: the emergence of an epigenetic paradigm. Nat. Rev. Genet. 12:565–575.

Flachs, P., T. Bhattacharyya, O. Mihola, J. Pialek, J. Forejt, and Z. Trachtulec. 2014. Prdm9 incompatibility controls oligospermia and delayed fertility but no selfish transmission in mouse intersubspecific hybrids. PLoS ONE 9:e95806.

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.

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.

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

Good, J. M., M. D. Dean, and M. W. Nachman. 2008b. A complex genetic basis to X-linked hybrid male sterility between two species of house mice. Genetics 179:2213–2228.

Gray, A. P. 1972. Mammalian hybrids. Commonwealth Agricultural Bureaux, Slough, U . K.

Haig, D. 1993. Genetic conflicts in human pregnancy. Q. Rev. Biol. 68:495–532.

———. 1997. Parental antagonism, relatedness asymmetries, and genomic imprinting. Proc. R. Soc. Lond. B Biol. Sci. 264:1657–1662.

Haldane, J. B. S. 1922. Sex ratio and unisexual sterility in animal hybrids. J. Genet. 12:101–109.

Hirshfield, A. N., and A. R. Midgley, Jr. 1978. Morphometric analysis of follicular development in the rat. Biol. Reprod. 19:597–605.

Kevenaar, M. E., M. F. Meerasahib, P. Kramer, B. M. N. vande Lang-Born, F. H. deJong, N. P. Groome, A. P. N. Themmen, and J. A. Visser. 2006. Serum anti-mullerian hormone levels reflect the size of the primordial follicle pool in mice. Endocrinology 147:3228–3234.

Kropáčková, L., J. Piálek, V. Gergelits, J. Forejt, and R. Reifová. 2015. Maternal-fetal genomic conflict and speciation: no evidence for hybrid placental dysplasia in crosses between two house mouse subspecies. J. Evol. Biol. 28:688–698.

Laukaitis, C. M., E. S. Crister, and R. C. Karn. 1997. Salivary androgen-binding protein (ABP) mediates sexual isolation in *Mus musculus*. Evolution 51:2000–2005.

Lee, H.-Y., J.-Y. Chou, L. Cheong, N.-H. Chang, S.-Y. Yang, and J.-Y. Leu. 2008. Incompatibility of nuclear and mitochondrial genomes causes hybrid sterility between two yeast species. Cell 135:1065–1073.

Macholán, M., S. J. E. Baird, P. Munclinger, P. Dufková, B. Bímová, and J. Piálek. 2008. Genetic conflict outweighs heterogametic incompatibility in the mouse hybrid zone? BMC Evol. Biol. 8:271–284.

Masly, J. P., and D. C. Presgraves. 2007. High-resolution genome-wide dissection of the two rules of speciation in *Drosophila*. PLoS Biol. 5:e243.

Meiklejohn, C. D., M. A. Holmbeck, and M. A. Siddiq. 2013. An incompatibility between a mitochondrial tRNA and its nuclear-encoded tRNA synthetase compromises development and fitness in *Drosophila*. PLoS Genet. 9:e1003238.

Mendez De La Cruz, F. R., L. J. Guillette, Jr., and M. V. Cruz. 1993. Differential atresia of ovarian follicles and its effect on the clutch size of two populations of the viviparous lizard *Sceloporus mucronatus*. Funct. Ecol. 7:535–540.

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.

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

Niwa-Kawakita, M. 1994. Reproductive depression of female mice in intersubspecific F2 hybrids of the species *Mus musculus*. Pp. 121–128 *in* K. Moriwaki, T. Shiroishi, and H. Yonekawa, eds. Genetics in wild mice: its application to biomedical research. Japan Scientific Press/Karger, Tokyo.

Orr, H. A. 1987. Genetics of male and female sterility in hybrids of *Drosophila pseudoobscura* and *Drosophila persimilis*. Genetics 116:555–563.

———. 1993. Haldane's rule has multiple genetic causes. Nature 361:532–533.

Orr, H. A., and J. A. Coyne. 1989. The genetics of postzygotic isolation in the *Drosophila virilis* group. Genetics 121:527–537.

Presgraves, D. C. 2003. A fine-scale genetic analysis of hybrid incompatibilities in *Drosophila*. Genetics 163:955–972.

Ratts, V. S., J. A. Flaws, R. Kolp, C. M. Sorenson, and J. L. Tilly. 1995. Ablation of *bcl-2* gene expression decreases the number of oocytes and primordial follicles established in the post-natal female mouse gonad. Endocrinology 136:3665–3668.

Rawson, P. D., and R. S. Burton. 2002. Functional coadaptation between cytochrome *c* and cytochrome *c* oxidase within allopatric populations of a marine copepod. Proc. Natl. Acad. Sci. USA 99:12955–12958.

Rousset, F., and M. Raymond. 1991. Cytoplasmic incompatibility in insects: why sterilize females? Trends Ecol. Evol. 6:54–57.

Rucker, E. B., P. Dierisseau, K. U. Wagner, L. Garrett, A. Wynshaw-Boris, J. A. Flaws, and L. Hennighausen. 2000. Bcl-x and Bax regulate mouse primordial germ cell survival and apoptosis during embryogenesis. Mol. Endocrinol. 14:1038–1052.

Sage, R. D., D. Heyneman, K.-C. Lim, and A. C. Wilson. 1986. Wormy mice in a hybrid zone. Nature 324:60–63.

Sage, R. D., W. R. Atchley, and E. Capanna. 1993. House mice as models in systematic biology. Syst. Biol. 42:523–561.

Saumitou-Laprade, P., J. Cuguen, and P. Vernet. 1994. Cytoplasmic male sterility in plants: molecular evidence and the nucleocytoplasmic conflict. Trends Ecol. Evol. 9:431–435.

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. R. Soc. Lond. B. 272:2491–2497.

Slotman, M., A. Della Torre, and J. R. Powell. 2005. Female sterility in hybrids between *Anopheles gambiae* and *An. arabiensis* and the causes of Haldane's rule. Evolution 59:1016–1026.

Smadja, C., and G. Ganem. 2002. Subspecies recognition in the house mouse: a study of two populations from the border of a hybrid zone. Behav. Ecol. 13:312–320.

Smadja, C., J. Catalan, and G. Ganem. 2004. Strong premating divergence in a unimodal hybrid zone between two subspecies of the house mouse. J. Evol. Biol. 17:165–176.

Spencer, H. G., A. G. Clark, and M. W. Feldman. 1999. Genetic conflicts and the evolutionary origin of genomic imprinting. Trends Ecol. Evol. 14:197–201.

St. John, J. C., R. E. Lloyd, E. J. Bowles, E. C. Thomas, and S. E. Shourbagy. 2004. The consequences of nuclear transfer for mammalian foetal development and offspring survival. A mitochondrial DNA perspective. Reproduction 127:631–641.

Su, A. I., M. P. Cooke, K. A. Ching, Y. Hakak, J. R. Walker, T. Wiltshire, A. P. Orth, R. G. Vega, L. M. Sapinoso, A. Moqrich, et al. 2002. Large-scale analysis of the human and mouse transcriptomes. Proc. Natl. Acad. Sci. USA 99:4465–4470.

Tanemura, K., M. Kurohmaru, K. Kuramoto, and Y. Hayashi. 1993. Age-related morphological changes in the testis of the BDF1 mouse. J. Vet. Med. Sci. 55:703–710.

Tilly, J. L. 2003. Ovarian follicle counts—not as simple as 1, 2, 3. Reprod. Biol. Endocrinol. 1:11.

Turelli, M., and H. A. Orr. 1995. The dominance theory of Haldane's rule. Genetics 140:389–402.

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.

Vrana, P. B., X.-J. Guan, R. S. Ingram, and S. M. Tilghman. 1998. Genomic imprinting is disrupted in interspecific *Peromuscus* hybrids. Nat. Genet. 20:362–365.

Vrana, P. B., J. A. Fossella, P. Matteson, T. del Rio, M. J. O'Neill, and S. M. Tilghman. 2000. Genetic and epigenetic incompatibilities underlie hybrid dysgenesis in *Peromyscus*. Nat. Genet. 25:120–124.

Vyskočilová, 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.

Wang, L., K. Luzynski, J. E. Pool, V. Janousek, P. Dufkova, M. M. Vyskocilova, K. C. Teeter, M. W. Nachman, P. Munclinger, M. Macholan, et al. 2011. Measures of linkage disequilibrium among neighbouring SNPs indicate asymmetries across the house mouse hybrid zone. Mol. Ecol. 20:2985–3000.

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.

Zechner, U., M. Reule, A. Orth, F. Bonhomme, B. Strack, J. L. Guenet, H. Hameister, and R. Fundele. 1996. An X-chromosome linked locus contributes to abnormal placental development in mouse interspecies hybrids. Nat. Genet. 12:398–403.

Zeh, J. A., and D. W. Zeh. 2008. Viviparity-driven conflict: more to speciation than meets the fly. Ann. N. Y. Acad. Sci. 113:126–148.

Associate Editor: M. Hahn
Handling Editor: M. Servedio

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Figure S1. No relationship between litter size and age of tester males.
Figure S2. Hybrid female fertility along reproductive time scale.
Figure S3. No significant differences between apoptotic cells between controls and hybrids.
Table S1. Phenotype scores for each of the 106 mice at 63 days old.
Table S2. Phenotype scores for each of the 22 mice at 120 days old.
Table S3. Classification of follicles used in the study.
Table S4. Test for delayed puberty using two hybrid genotypes.
Table S5. Comparisons of reproductive parameters between hybrid female progeny of reciprocal crosses ($n = 106$).
Table S6. Test for paternal polymorphism within subspecies for reproductive parameters ($n = 106$).
Table S7. Test for maternal polymorphisms within subspecies for reproductive parameters ($n = 106$).
Table S8. Summary of F1 hybrid sterility between males and females.