Journal of Heredity 2007:98(2):115–122
doi:10.1093/jhered/esl064
Advance Access publication January 5, 2007

Searching the Genomes of Inbred Mouse Strains for Incompatibilities That Reproductively Isolate Their Wild Relatives

Bret A. Payseur and Michael Place

From the Laboratory of Genetics, University of Wisconsin, Madison, WI 53706.

Address correspondence to the author at the address above, or e-mail: payseur@wisc.edu.

Abstract

Identification of the genes that underlie reproductive isolation provides important insights into the process of speciation. According to the Dobzhansky–Muller model, these genes suffer disrupted interactions in hybrids due to independent divergence in separate populations. In hybrid populations, natural selection acts to remove the deleterious heterospecific combinations that cause these functional disruptions. When selection is strong, this process can maintain multilocus associations, primarily between conspecific alleles, providing a signature that can be used to locate incompatibilities. We applied this logic to populations of house mice that were formed by hybridization involving two species that show partial reproductive isolation, *Mus domesticus* and *Mus musculus*. Using molecular markers likely to be informative about species ancestry, we scanned the genomes of 1) classical inbred strains and 2) recombinant inbred lines for pairs of loci that showed extreme linkage disequilibria. By using the same set of markers, we identified a list of locus pairs that displayed similar patterns in both scans. These genomic regions may contain genes that contribute to reproductive isolation between *M. domesticus* and *M. musculus*. This hypothesis can now be tested using laboratory crosses and surveys of introgression in the wild.

The identification of the genes that underlie reproductive isolation between species is an exciting goal because it provides access to the ultimate genetic mechanisms of speciation. Empirical studies across a broad range of species indicate that intrinsic postzygotic isolation is caused by the independent evolution of interacting genes in separate populations that disrupts functional interactions between these genes in hybrids (Hollingshead 1930; Dobzhansky 1936; Wu and Beckenbach 1983; Christie and Macnair 1984; Orr 1987, 1997; Pantazidis and Zouros 1988; Coyne and Orr 1989, 1997, 2004; Perez and Wu 1995; True et al. 1996; Fishman and Willis 2001; Presgraves 2002, 2003; Price and Bouvier 2002; Tao et al. 2003). In addition to providing predictions about the evolution of reproductive barriers (Orr 1995), this "Dobzhansky–Muller model" (Bateson 1909; Dobzhansky 1936, 1937; Muller 1940, 1942) suggests a useful framework for finding the genes involved. Although the relevant loci are difficult to identify in the allopatric populations in which they evolved, the incompatible changes at these genes become visible in hybrids.

Two general approaches to locating incompatibilities in the genomes of hybrid individuals have been employed. In the first method (Dobzhansky 1936), species pairs are crossed to produce F_1's, and a population that includes recombinant genomes is generated by hybridizing these F_1's to the parental species (backcross) or to each other (intercross). Associations between molecular markers and sterility or inviability phenotypes in this population reveal the genomic locations of genes whose normal interactions have been compromised and therefore contribute to reproductive isolation. The mapping resolution of this strategy can be substantially increased by subsequent generations of crossing (True et al. 1996; Tao and Hartl 2003). Results from experimental crosses between species pairs have revolutionized our understanding of speciation genetics (Coyne and Orr 2004) and have revealed the identities of specific genes that cause hybrid sterility (Ting et al. 1998) and hybrid inviability (Wittbrodt et al. 1989; Barbash et al. 2003; Presgraves et al. 2003).

The second approach examines differential introgression of molecular markers through hybrid populations in nature (Hunt and Selander 1973; Barton and Bengtsson 1986; Dowling et al. 1989; Harrison 1990; Rieseberg et al. 1999). By comparing patterns of variation at many unlinked markers, the effects of genome-wide forces (such as migration) can

be measured and regions with reduced introgression can be identified. Locus-specific retardation in gene flow reflects natural selection against particular hybrid genotypes (Barton and Hewitt 1985), revealing the genomic location of genes that maintain reproductive barriers between nascent species (Rieseberg et al. 1999). Although restricted to groups that hybridize in nature, this approach has the advantage of focusing on patterns of gene flow in a natural setting. The strategy is most powerful in organisms with well-characterized genomes.

One such organism, the house mouse, holds great promise for identifying the incompatible changes that underlie reproductive isolation. The house mouse group comprises several closely related species whose natural histories have been documented as a result of human commensalism. As the premier model system in mammalian genetics, the house mouse offers several useful resources for speciation genetics, including a large catalog of mutants with relevant phenotypes, a highly developed system for the generation of knockouts, platforms for tissue-specific gene expression surveys, a complete genome sequence, and extensive information on DNA polymorphism between available inbred strains.

The two tactics for finding incompatibilities described above have been applied with some success to house mouse species, particularly *Mus domesticus* and *Mus musculus*. These two species diverged between 0.5 and 1 million years ago (Boursot et al. 1993) and display partial reproductive isolation. F_1 hybrid males are often sterile, and females are fertile in crosses between inbred lines derived from *M. domesticus* and *M. musculus* (Britton-Davidian et al. 2005), in accordance with Haldane's (1922) rule. Additionally, naturally occurring hybrids bear higher parasite loads than do pure-species individuals (Sage et al. 1986; Moulia et al. 1993), suggesting that hybrids may suffer reduced viability in nature.

An extended area of sympatry between *M. domesticus* and *M. musculus* that stretches across central Europe is one of the most intensively studied hybrid zones in the world (Payseur and Nachman 2005). Several decades of research using molecular markers that differentiate the two species have documented two patterns in this hybrid zone that are relevant to reproductive isolation: 1) introgression is fairly limited at most surveyed molecular markers (despite large ancestral species ranges), suggesting that selection acts against hybrids, and 2) there is clear heterogeneity in gene flow among different genomic regions, suggesting that the targets of selection can be located (Tucker et al. 1992; Boursot et al. 1993; Dod et al. 1993; Sage et al. 1993; Munclinger et al. 2002). Motivated by these observations, Payseur et al. (2004) documented differential patterns of introgression across the X chromosome, including a region with substantially reduced gene flow that likely contains genes that confer reproductive barriers between these species.

Attempts to locate incompatibilities between *M. domesticus* and *M. musculus* have also involved controlled crosses in the laboratory. Matings between wild-derived inbred lines of *M. musculus* and some classical inbred strains (which are primarily descended from *M. domesticus*) yield sterile hybrid males, whereas crosses between *M. musculus* and other classical strains produce fertile males (Forejt and Ivanyi 1974; Forejt 1996). Part of this difference is attributable to a gene on chromosome 17, *Hst1*, which has recently been localized to a 360-kbp region (Gregorova et al. 1996; Trachtulec et al. 2005). These findings were enabled by the unusual history of the classical strains, which are ultimately descended from crosses involving wild *M. domesticus* and *M. musculus* (Morse 1978; Silver 1995; Beck et al. 2000). The hybrid nature of classical strain genomes has been confirmed by molecular polymorphism data, including contrasting histories for the Y chromosome (which is primarily of *M. musculus* origin; Bishop et al. 1985) and the mitochondrial DNA (which is primarily of *M. domesticus* origin; Yonekawa et al. 1980; Ferris et al. 1982), and autosomal loci that apparently segregate interspecific variation (Wade et al. 2002; Wade and Daly 2005). Due to recent descriptions of genome-wide polymorphism (Mural et al. 2002; Wiltshire et al. 2003; Witmer et al. 2003; Frazer et al. 2004; Petkov et al. 2004; Pletcher et al. 2004; Yalcin et al. 2004; Zhang et al. 2005), the classical strains are now one of the most exhaustively surveyed hybrid populations at the molecular level.

Although hybrid zone surveys and laboratory crosses have yielded candidate regions for incompatibilities in house mice, the identification of the full set of partners whose disrupted interaction leads to reproductive isolation has been more challenging. These approaches are primarily designed to find "individual" loci that participate in incompatibilities—neither strategy explicitly tests for the existence of epistasis. Fortunately, the Dobzhansky–Muller model suggests a diagnostic tool for finding the interacting loci that cause incompatibilities.

In hybrid populations, natural selection acts to remove the deleterious multilocus combinations that cause hybrid sterility or inviability. When selection is strong, this process can maintain statistical associations ("linkage disequilibrium") among conspecific alleles at the participating loci, even in the face of recombination. Because this pattern will extend to linked loci, incompatibility partners can be located by scanning hybrid genomes for strong associations between conspecific alleles at marker loci. In the first application of this approach, Gardner et al. (2002) showed that genomic regions in linkage disequilibrium were often associated with phenotypes related to reproductive isolation in sunflower hybrids. Using single-nucleotide polymorphism (SNP) data from across the genomes of classical mouse strains, Payseur and Hoekstra (2005) found unlinked loci showing strong linkage disequilibrium and discovered that a disproportionate fraction of these outliers were driven by associations between conspecific alleles, as predicted under the Dobzhansky–Muller model.

The recent availability of dense genotypes for recombinant inbred lines (RILs) of mice now allows the application of this approach to additional populations that segregate variation from *M. domesticus* and *M. musculus*. RIL panels are generated by a defined crossing scheme and therefore provide a conservative, independent experiment in which to evaluate evidence for epistatic selection (Petkov et al. 2005).

Here, we report the results of separate genomic scans for incompatibilities between *M. domesticus* and *M. musculus* in the

classical inbred strains and a panel of RILs. Using a common set of markers allows us to directly compare patterns in these two hybrid populations and to nominate a list of candidate regions that show similar results in both scans.

Materials and Methods

SNP genotypes for all strains were obtained from the Wellcome-CTC Mouse Strain SNP Genotype Set (http://www.well.ox.ac.uk/mouse/INBREDS/). Strains and SNPs were selected for analyses using several criteria. First, to focus on markers most likely to be informative about species ancestry, SNPs that showed fixed differences between available wild-derived strains of *M. domesticus* (PERA/EiJ, PERC/EiJ, LEWES/EiJ, TIRANO/EiJ, WMPPasDn/J, WSB/EiJ, and ZALENDE/EiJ) and *M. musculus* (CZECHI/EiJ, MAI/Pas, MBT/Pas, PWK/Pas, PWKPh/J, PWK/Ros, PWK/Rbrc, and SKIVE/EiJ) and also had genotypes available for 22 of the classical strains (A/J, AKR/J, BTBR T+tf/J, BUB/BnJ, C3H/HeJ, C57BL/10J, DBA/1J, FVB/NJ, I/LnJ, KK/HIJ, LG/J, LP/J, MA/MyJ, NOD/LtJ, NON/LtJ, NZB/BINJ, PL/J, RIIIS/J, SEA/GnJ, SJL/J, ST/bJ, and 129 × 1/SvJ) were selected ($n = 973$). These 22 classical strains, which formed one hybrid population for analyses, were chosen to exclude wild-derived strains and very closely related strains. Next, the subset of these SNPs that also differed between C57BL/6J and DBA/2J, the parents of 89 "B × D" RILs (Taylor et al. 1999; Williams et al. 2001; Peirce et al. 2004; http://www.well.ox.ac.uk/mouse/INBREDS/), was chosen ($n = 303$) to enable direct comparison between the classical strains and this group of RILs. Finally, the remaining SNPs with *M. domesticus* allele frequencies of greater than 1/22 or less than 21/22 in the classical strains were retained for linkage disequilibrium analyses. The final group of 256 SNPs with informative genotypes in both the classical strains and the RILs was fairly evenly distributed across the autosomes (no X-linked SNPs were included; see Discussion).

Each inbred line was assumed to be completely homozygous at all markers, allowing direct inference of linkage disequilibrium. Linkage disequilibrium was estimated separately for the classical strain ($n = 22$) and the RIL ($n = 89$) populations using several standard metrics. Because results using different measures were similar, we focus on the squared correlation coefficient, R^2 (Hill and Robertson 1968), calculated as

$$R^2 = \frac{(p_{d1,d2} - p_{d1}p_{d2})^2}{p_{d1}(1-p_{d1})p_{d2}(1-p_{d2})},$$

where p_{d1} and p_{d2} are the frequencies of the *M. domesticus* allele at locus 1 and locus 2, and $p_{d1,d2}$ is the frequency of the gametic type carrying *M. domesticus* alleles at both loci.

To ensure independent assortment between markers in each generation, linkage disequilibrium was calculated for all SNP pairs located on different chromosomes (a total of 29 648 tests in each strain set). For each test, a P value was assigned by randomly permuting SNP genotypes at one locus across strains 1000 times, calculating the resulting distribution of R^2 values, and comparing the original R^2 value to this permuted distribution. Both Bonferroni and false-discovery rate (Storey and Tibshirani 2003) approaches indicated that only tests in which the observed R^2 was greater than all permuted values were statistically significant after accounting for the performance of multiple tests. In interpreting the results, we identified extreme tests for comparison in the classical strains and the RILs using two significance thresholds. First, we considered those tests for which the observed R^2 was greater than all permuted values ($P < 0.001$). Second, we considered those tests with $P < 0.05$. We intentionally applied this more liberal significance criterion, at the cost of including false positives, to search for overlap between the classical strains and the RILs.

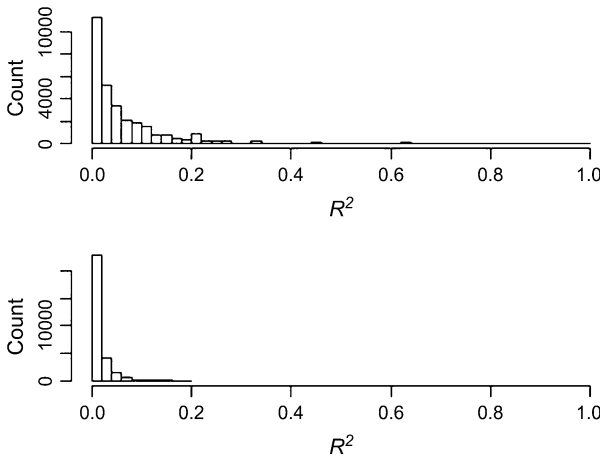

Figure 1. Genomic distributions of pairwise linkage disequilibria (R^2) among SNPs located on different chromosomes in 22 classical inbred strains **(A)** and in 89 B × D RILs **(B)**.

Results

The distributions of R^2 across 22 classical strains and 89 RILs for pairs of SNPs located on different chromosomes are shown in Figure 1. Average R^2 values were different than 0 in both strain sets (classical strains, mean = 0.06, $P < 10^{-15}$; RILs, mean = 0.01, $P < 10^{-15}$; one-sample t-test), with the classical strains exhibiting higher R^2 than the RILs in the same set of comparisons (mean difference in R^2 = 0.05; $P < 10^{-15}$; paired t-test). This disparity might have been caused by the nonequilibrium demographic history of the classical strains. Despite this evidence for a departure from linkage equilibrium in the two strain sets, the low mean values of R^2 indicate that strong associations were unlikely to arise in the absence of selection.

The overall relationship between R^2 values in the classical strains and in the RILs for the same set of tests is displayed in Figure 2; these values were not significantly correlated (Spearman's $\rho = 0.006$; $P = 0.34$). Because most locus pairs are unlikely to be affected by epistatic selection, this pattern is

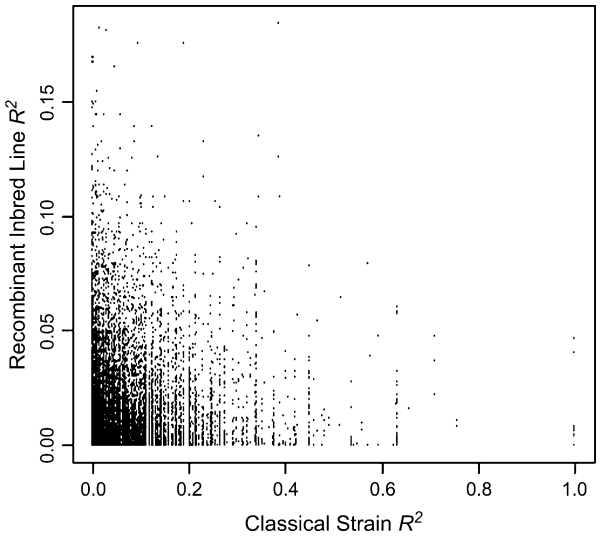

Figure 2. Scatterplot of R^2 values for the classical strains versus the RILs.

expected. However, the prediction that tests showing strong associations should overlap between the two strain sets was supported. Those locus pairs that showed extreme associations in the classical strains had significantly higher R^2 values ($P = 0.03$ when using a $P < 0.001$ significance criterion in classical strains; $P < 10^{-8}$ when using a $P < 0.05$ significance criterion; Wilcoxon signed rank test) and lower P values ($P = 0.03$ when using a $P < 0.001$ significance criterion in classical strains; $P < 10^{-15}$ when using a $P < 0.05$ significance criterion) in the RILs. This pattern suggests that signatures of epistatic selection were replicated across the two strain sets.

We expect selection against incompatibilities that reproductively isolate *M. domesticus* and *M. musculus* and are segregating in these strain sets to maintain linkage disequilibrium between conspecific alleles in the face of recombination. If many incompatibilities are present, we might predict that those locus pairs showing extreme disequilibrium will be enriched for conspecific associations. Consistent with previous results (Payseur and Hoekstra 2005), we observed a bias toward conspecific associations (reductions of heterospecific genotype frequencies) among extreme tests in the classical strains using both $P < 0.001$ ($P = 0.07$; Fisher's exact test) and $P < 0.05$ ($P < 10^{-15}$; Fisher's exact test) significance thresholds. In contrast, there was no clear pattern in the direction of association for locus pairs showing extreme disequilibrium in the RILs ($P > 0.05$ in both tests). Despite the lack of evidence for a genome-wide pattern in the RILs, locus pairs showing extreme associations between conspecific alleles still represent reasonable candidates for incompatibilities. Because the classical strains and the RILs constitute independent evolutionary experiments, those tests that showed strong linkage disequilibrium between conspecific alleles in both strain sets provide our best incompatibility candidates.

Using a significance criterion of $P < 0.001$, 12 tests showed extreme associations biased toward conspecific combinations in the classical strains and 10 tests showed this pattern in the RILs. Among these tests, there were no cases in which both SNPs were the same across the two strain sets. However, one SNP on chromosome 11 (29.45 Mbp) appeared as an outlier in both scans. This SNP (rs13480935) showed associations with different regions in the classical strains (chromosome 6, 72.17 Mbp) and in the RILs (chromosome 15, 53.83 Mbp) and caused a nonsynonymous substitution (Gln → Arg) in a predicted gene (1700034F02Rik).

Ten tests involving the same SNP pairs were driven by associations between conspecific alleles and achieved significance at the $P < 0.05$ level in both the classical strains and the RILs (Table 1). Although these disequilibria were not significant after adjusting for multiple testing, their detection in both strain sets suggests that these locus pairs marked disrupted functional interactions between hybrid genotypes (incompatibilities).

Discussion

Natural selection against incompatibilities can maintain linkage disequilibrium in the face of recombination in hybrid populations (Gardner et al. 2002; Payseur and Hoekstra 2005). Using this rationale, we conducted genomic scans for incompatibilities in two hybrid populations of house mice: a group of classical strains and a set of RILs. Focusing on the same markers allowed us to compare results from both scans. Multiple factors, including the chance fixation of alleles within and between lines during the process of inbreeding, probably contributed to associations among SNPs on different chromosomes in both the classical strains and the RILs (Williams et al. 2001). However, these processes were not expected to generate associations at the same loci in both strain sets. Consequently, genomic regions from our list of marker pairs that showed extreme linkage disequilibrium driven by conspecific associations in the classical strains and the RILs might contain targets of epistatic selection against hybrid genotypes.

One SNP, located at 29.45 Mbp on chromosome 11, showed very strong associations among conspecific alleles in both scans. Interestingly, disrupted interactions between this region and a locus on the X chromosome have recently been shown to cause hybrid male sterility in crosses between *M. molossinus* (a lineage closely related to *M. musculus*) and C57BL/6J (Oka et al. 2006). Our analyses did not consider X-linked SNPs (see below). Although the chromosome 11 SNP we identified might have been in linkage disequilibrium with the mutations driving the observed associations, this nonsynonymous change was found in a gene expressed in adult testis, a pattern consistent with its involvement in hybrid male sterility. In contrast to the prediction for a simple 2-locus incompatibility, this SNP was associated with different loci in the classical strains and the RILs. This pattern might indicate that incompatibilities involving this locus are complex, with fitness effects of the 2-locus allelic combinations depending on the genetic background. Such a complex epistasis has been previously observed to contribute to

Table 1. SNP pairs that showed associations that were extreme ($P < 0.05$) and biased toward conspecific allelic combinations in both the classical inbred strains and the RILs

Chromosome	Position[a]	Genic location	Named genes within 1 Mbp[b]	Chromosome	Position	Genic location	Named genes within 1 Mbp	Classical strain R^2	Classical strain P	RIL R^2	RIL P
2	51959411	Coding (Met-Ile, Neb)	Rnd3, Tas2r134, Nmi, Tnfaip6, Rif1, Neb, Arl5, Cacnb4, Stam2, Fmnl2	4	54744598	Intergenic	Fcmd, Tal2, Tmem38b, Zfp462, Rad23b, Klf4	0.26	0.047	0.11	0.001
2	70040299	Intronic, A430065P19Rik	Abcb11, Dhrs9, Lrp2, Bbs5, Kbtbd10, Ppig, Phospho2, Klhl23, Ssb, Mettl5, Sp5, Gad1, Gorasp2, Tlk1, Cybrd1, Dyncl12	17	79269800	Intergenic	Crim1, Fez2, Vit, Strn, Eif2ak2, Cebpz, Prkcn, Qpct, Cdc42cp3, Cyp1b1, Arl6ip2, Hnrpll, Galm, Sfrs7, Gemin6, Dhx57, Morn2, Gm941	0.63	0.009	0.06	0.045
2	103085173	Intronic, Ehf	Trim44, Fjx1, Slc1a2, Cd44, Pdhx, Apip, Ehf, Elf5, Abtb2, Nat10, Gpiap1, Lmo2, Fbxo3, Cd59b, Cd59a	8	75340444	Intronic, Eps15l1	Abhd8, Mrpl34, Tmem16h, Gtpbp3, Plvap, Bst2, Txnl6, Slc27a1, Pgls, Glt25d1, Unc13a, Jak3, Insl3, B3gnt3, Fcho1, Zfp709, Zfp617, Cyp4f18, Olfr372, Olfr373, Olfr374, Tpm4, Rab8a, Hsh2d, Cib3, Ap1m1, Klf2, Eps15l1, Calr3, Cherp, Crsp7, Tmem38a, Sin3b, F2rl3, Large	0.23	0.046	0.13	0.002
4	55597684	Intergenic	Zfp462, Rad23b, Klf4	6	111318505	Intronic, Grm7	Grm7, Lmcd1	0.38	0.030	0.05	0.048
5	87403662	Intronic, Tmprss11d	Cenpc1, Ube1l2, Gnrhr, Tmprss11c, Tmprss11d, Tmprss11a, Tmprss11f, Tmprss11b, Tmprss11e, Ugt2b34, Ugt2b1, Ugt2b35, Ugt2b36, Ugt2b5, Ugt2b37, Ugt2a3	17	77627900	Intergenic	Crim1, Fez2, Vit, Strn	0.34	0.033	0.08	0.015
6	98378988	Intergenic	Frmd4b, Mitf, Foxp1	12	34836766[c]	Intergenic	Atxn7l1, Twistnb, Ferd31, Twist1, Hdac9, Snx13	0.32	0.026	0.08	0.012
6	134587480	Coding (Val-Leu, Mansc1)	Kap, Etv6, Bcl2l14, Lrp6, Mansc1, Loh12cr1, Dusp16, Crebl2, Gpr19, Cdkn1b, Apold1, Ddx47, Gprc5a, Gprc5d, Hebp1, Gsg1, Pbp2, Emp1	16	41259652	Intergenic	Lsamp, Gap43	0.42	0.008	0.06	0.027
7	83593190	Intronic (Il16)	Eftud1, Rkhd3, Tmc3, Stard5, Il16, Mesdc1, Mesdc2, Arnt2, Fah, Za20d3	13	20913932	Intronic (Aoah)	Elmo1, Aoah, Olfr1370, Olfr42, Olfr1368, Trim27, Gpx5, Olfr1367, Zfp96, Zfp306, Zfp187, Zfp192, Olfr1366, Olfr1365, Olfr1364, Olfr1362, Olfr11, Olfr1361, Olfr1360, Olfr1359, Hist1h2bl, Hist1h2ai, Hist1h3h, Hist1h2aj, Hist1h2bm, Hist1h4j, Hist1h4k, Hist1h2ak, Hist1h2bn, Hist1h1b, Hist1h3i, Hist1h2an, Hist1h2bp, Hist1h4m, Hist1h2ao	0.35	0.021	0.13	<0.001
10	103537270	Intergenic	Lrriq1, Slc6a15	17	81580956[c]	Intergenic	Map4k3, Thumpd2, Slc8a1	0.27	0.047	0.06	0.02
11	38173654	Intergenic	None	16	41259652	Intergenic	Lsamp, Gap43	0.32	0.021	0.06	0.031

[a] Base-pair position from The Jackson Laboratory SNP database (http://phenome.jax.org/pub-cgi/phenome/mpdcgi?rtn=snps/list_pre).

[b] Named genes within a 2-Mbp window centered on the SNP position.

[c] Updated position could not be found in The Jackson Laboratory SNP database. Original Wellcome Trust position is provided instead.

reproductive isolation (Wu and Beckenbach 1983; Orr and Irving 2001; Storchova et al. 2004). Further evidence for the existence of complex interactions was provided by the association of a SNP on chromosome 16 with multiple regions in both the classical strains and the RILs (Table 1). Additionally, three SNPs across a distance of about 4 Mbp on chromosome 17 showed extreme associations in both strain sets, suggesting that this genomic region might be involved in higher order incompatibilities.

Two genomic regions thought to be involved in reproductive isolation between *M. domesticus* and *M. musculus*, the central region of the X chromosome and the proximal tip of chromosome 17 (where *Hst1* is located), were absent from our list of incompatibility candidates. Available X-linked SNPs that were diagnostic of species were eliminated prior to our genomic scans, primarily because the *M. domesticus* allele was often fixed or segregating at a very high frequency in the classical strains. This pattern is consistent with a role for the *M. musculus* X chromosome in reproductive isolation (Gregorova and Forejt 2000; Storchova et al. 2004; Harr 2006). The SNP closest to *Hst1* in our survey was located approximately 3 Mbp away (proximal; Trachtulec et al. 2005) and showed conspecific associations in the classical strains and the RILs (data not shown) but not with the same SNPs. This result might indicate that incompatibilities involving *Hst1* are complex. Alternatively, marker density might have been too sparse in one or both strain sets for *Hst1* and/or its partner loci to be in linkage disequilibrium with the surveyed markers. Furthermore, our focus on one set of RILs restricted the SNPs we used to those that segregated between C57BL/6J and DBA/2J. If *Hst1* and/or its partner loci did not vary among these strains, we would have no power to find these incompatibilities in the RILs, even if there was strong selection against them in the classical strains. In fact, selection against these incompatibilities in the early recombinant generations of the classical strains might have left C57BL/6J and DBA/2J identical at these loci.

Our list of incompatibility candidates (Table 1) does not satisfy requirements for statistical significance in light of the large number of tests performed in both scans. By randomizing the statistical significance ($P < 0.05$ vs. $P >= 0.05$) and nature of the associations (conspecific vs. heterospecific) across genomic locations, we estimate that the number of tests expected to show these diagnostic patterns in "both" the classical strains and the RILs by chance alone is similar to the number reported in Table 1 (data not shown). As a result, these candidates should be viewed as preliminary until they can be validated using a combination of approaches. First, if these genomic regions contain incompatibilities, we predict that similar patterns should be observed in other admixed populations, including natural hybrid zones. In addition to measuring linkage disequilibrium, geographic clines in multilocus genotypic frequencies could be used to test whether alleles at these loci "cointrogress." Second, recombinant generations of crosses between *M. domesticus* and *M. musculus* should show linkage disequilibrium or segregation distortion involving these genomic regions. The association of hybrid sterility or inviability phenotypes with these regions would provide more direct evidence of their importance in reproductive isolation. Finally, if patterns of variation at these locus pairs reflect selection against incompatibilities, we would expect genes in these regions (Table 1) to functionally interact. The house mouse is one of just a few systems for which this multifaceted approach—genomic scans for epistatic selection, surveys of differential introgression in hybrid zones, genetic mapping in crosses between species pairs, and functional tests for epistasis—is currently feasible.

It should also be possible to improve the approach used here by developing better multilocus signatures of selection against incompatibilities. Although measures of linkage disequilibrium are adept at detecting interlocus associations, they are not specifically designed to find a reduction in the frequency of one gametic type, the pattern predicted under the Dobzhansky–Muller model (Muller 1942; Orr 1995; Payseur and Hoekstra 2005). Moreover, multivariate measures of association (Nyholt 2004) may facilitate the detection of more complex incompatibilities by measuring higher order correlations rather than pairwise associations. Finally, alternative analyses of gametic frequencies that are not focused on linkage disequilibrium might prove to be more powerful for unlinked loci because linkage disequilibrium decays quickly with free recombination. Further development and application of these strategies seems warranted by the multilocus nature of the Dobzhansky–Muller model.

Acknowledgments

We thank Loren Rieseberg for organizing an excellent symposium and for inviting us to contribute to this special issue. We also thank the members of the Wellcome-CTC consortium for making the inbred strain SNP genotypes publicly available. Michael White and an anonymous reviewer provided useful comments on the manuscript. This paper is based on a presentation given at the 2006 Annual Meeting of the American Genetic Association, "Genetics of Speciation," University of British Columbia, Vancouver, Canada, July 21–24, 2006.

References

Barbash DA, Siino DF, Tarone AM, Roote J. 2003. A rapidly evolving MYB-related protein causes species isolation in Drosophila. Proc Natl Acad Sci USA. 100:5302–5307.

Barton N, Bengtsson BO. 1986. The barrier to genetic exchange between hybridising populations. Heredity. 57:357–376.

Barton NH, Hewitt GM. 1985. Analysis of hybrid zones. Annu Rev Ecol Syst. 16:113–148.

Bateson W. 1909. Heredity and variation in modern lights. In: Seward AC, editor. Darwin and modern science. Cambridge (UK): Cambridge University Press. p. 85–101.

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

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

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

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

Christie P, Macnair MR. 1984. Complementary lethal factors in 2 North-American populations of the yellow monkey flower. J Hered. 75: 510–511.

Coyne JA, Orr HA. 1989. Patterns of speciation in Drosophila. Evolution. 43:362–381.

Coyne JA, Orr HA. 1997. "Patterns of speciation in Drosophila" revisited. Evolution. 51:295–303.

Coyne JA, Orr HA. 2004. Speciation. Sunderland (MA): Sinauer Associates.

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

Dobzhansky T. 1937. Genetics and the origin of species. New York: Columbia University Press.

Dod B, Jermiin LS, Boursot P, Chapman VH, Tonnes-Nielsen J, Bonhomme F. 1993. Counterselection on sex chromosomes in the Mus musculus European hybrid zone. J Evol Biol. 6:529–546.

Dowling TE, Smith GR, Brown WM. 1989. Reproductive isolation and introgression between Notropis-cornutus and Notropis-chrysocephalus (Family Cyprinidae)—comparison of morphology, allozymes, and mitochondrial-DNA. Evolution. 43:620–634.

Ferris SD, Sage RD, Wilson AC. 1982. Evidence from mtDNA sequences that common laboratory strains of inbred mice are descended from a single female. Nature. 295:163–165.

Fishman L, Willis JH. 2001. Evidence for Dobzhansky-Muller incompatibilites contributing to the sterility of hybrids between Mimulus guttatus and M. nasutus. Evolution. 55:1932–1942.

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

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

Frazer KA, Wade CM, Hinds DA, Patil N, Cox DR, Daly MJ. 2004. Segmental phylogenetic relationships of inbred mouse strains revealed by fine-scale analysis of sequence variation across 4.6 mb of mouse genome. Genome Res. 14:1493–1500.

Gardner K, Buerkle A, Whitton J, Rieseberg LH. 2002. Inferring epistasis in wild sunflower hybrid zones. In: Wolf JB, Brodie ED 3rd, Wade MJ, editors. Epistasis and the evolutionary process. Oxford (UK): Oxford University Press. p. 264–279.

Gregorova S, Forejt J. 2000. PWD/Ph and PWK/Ph inbred mouse strains of Mus m. musculus subspecies—a valuable resource of phenotypic variations and genomic polymorphisms. Folia Biol. 46:31–41.

Gregorova S, Mnukova-Fajdelova M, Trachtulec Z, Capkova J, Loudova M, Hoglund M, Hamvas R, Lehrach H, Vincek V, Klein J, et al. 1996. Sub-milli-Morgan map of the proximal part of mouse chromosome 17 including the hybrid sterility 1 gene. Mamm Genome. 7:107–113.

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

Harr B. 2006. Genomic islands of differentiation between house mouse subspecies. Genome Res. 16:730–737.

Harrison RG. 1990. Hybrid zones: windows on evolutionary processes. Oxf Surv Evol Biol. 7:69–128.

Hill WG, Robertson A. 1968. Linkage disequilibrium in finite populations. Theor Appl Genet. 38:226–231.

Hollingshead L. 1930. A lethal factor in *Crepis* effective only in interspecific hybrids. Genetics. 15:114–140.

Hunt WG, Selander RK. 1973. Biochemical genetics of hybridisation in European house mice. Heredity. 31:11–33.

Morse HC. 1978. Introduction. In: Morse HC, editor. Origins of inbred mice. New York: Academic Press. p. 1–31.

Moulia C, Le Brun N, Dallas J, Orth A, Renaud F. 1993. Experimental evidence of genetic determinism in high susceptibility to intestinal pinworm infection in mice: a hybrid zone model. Parasitology. 106:387–393.

Muller HJ. 1940. Bearing of the Drosophila work on systematics. Huxley JS, editor. The new systematics. Oxford (UK): Clarendon Press. p. 185–268.

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

Munclinger P, Bozikova E, Sugerkova M, Pialek J, Macholan M. 2002. Genetic variation in house mice (Mus, Muridae, Rodentia) from the Czech and Slovak Republics. Folia Zool. 51:81–92.

Mural RJ, Adams MD, Myers EW, Smith HO, Miklos GLG, Wides R, Halpern A, Li PW, Sutton GG, Nadeau J, et al. 2002. A comparison of whole-genome shotgun-derived mouse chromosome 16 and the human genome. Science. 296:1661–1671.

Nyholt DR. 2004. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. Am J Hum Genet. 74:765–769.

Oka A, Aoto T, Totsuka Y, Takahashi R, Ueda M, Mita A, Sakurai-Yamatani N, Yamamoto H, Kuriki S, Takagi N, et al. 2006. Disruption of genetic interaction between two autosomal regions and the X chromosome causes reproductive isolation between mouse strains derived from different subspecies. Genetics. Epub Ahead of Print.

Orr HA. 1987. Genetics of male and female sterility in hybrids of Drosophila pseudoobscura and D. persimilis. Genetics. 116:555–563.

Orr HA. 1995. The population genetics of speciation: the evolution of hybrid incompatibilities. Genetics. 139:1805–1813.

Orr HA. 1997. Haldane's rule. Annu Rev Ecol Syst. 28:195–218.

Orr HA, Irving S. 2001. Complex epistasis and the genetic basis of hybrid sterility in the Drosophila pseudoobscura Bogota-USA hybridization. Genetics. 158:1089–1100.

Pantazidis AC, Zouros E. 1988. Location of an autosomal factor causing sterility in Drosophila mojavensis males carrying the Drosophila arizonensis Y chromosome. Heredity. 60:299–304.

Payseur BA, Hoekstra HE. 2005. Signatures of reproductive isolation in patterns of single nucleotide diversity across inbred strains of mice. Genetics. 171:1905–1916.

Payseur BA, Krenz JG, Nachman MW. 2004. Differential patterns of introgression across the X chromosome in a hybrid zone between two species of house mice. Evolution. 58:2064–2078.

Payseur BA, Nachman MW. 2005. The genomics of speciation: investigating the molecular correlates of X chromosome introgression across the hybrid zone between Mus domesticus and Mus musculus. Biol J Linn Soc. 84: 523–534.

Peirce JL, Lu L, Gu J, Silver LM, Williams RW. 2004. A new set of BXD recombinant inbred lines from advanced intercross populations in mice. BMC Genet. 5:7.

Perez DE, Wu CI. 1995. Further characterization of the Odysseus locus of hybrid sterility in Drosophila: one gene is not enough. Genetics. 140: 201–206.

Petkov PM, Ding Y, Cassell MA, Zhang W, Wagner G, Sargent EE, Asquith S, Crew V, Johnson KA, Robinson P, et al. 2004. An efficient SNP system for mouse genome scanning and elucidating strain relationships. Genome Res. 14:1806–1811.

Petkov PM, Graber JH, Churchill GA, DiPetrillo K, King BL, Paigen K. 2005. Evidence of a large-scale functional organization of mammalian chromosomes. PLoS Genet. 1:312–322.

Pletcher MT, McClurg P, Batalov S, Su AI, Barnes SW, Lagler E, Korstanje R, Wang X, Nusskern D, Bogue MA, et al. 2004. Use of a dense single

nucleotide polymorphism map for in silico mapping in the mouse. PLoS Biol. 2:e393.

Presgraves DC. 2002. Patterns of postzygotic isolation in Lepidoptera. Evolution. 56:1168–1183.

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

Presgraves DC, Balagopalan L, Abmayr SM, Orr HA. 2003. Adaptive evolution drives divergence of a hybrid inviability gene between two species of Drosophila. Nature. 423:715–719.

Price TD, Bouvier MM. 2002. The evolution of F1 postzygotic incompatibilities in birds. Evolution. 56:2083–2089.

Rieseberg LH, Whitton J, Gardner K. 1999. Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. Genetics. 152:713–727.

Sage RD, Atchley WR, Capanna E. 1993. House mice as models in systematic biology. Syst Biol. 42:523–561.

Sage RD, Heyneman D, Lim KC, Wilson AC. 1986. Wormy mice in a hybrid zone. Nature. 324:60–63.

Silver LM. 1995. Mouse genetics. New York: Oxford University Press.

Storchova R, Gregorova S, Buckiova D, Kyselova V, Divina P, Forejt J. 2004. Genetic analysis of X-linked hybrid sterility in the house mouse. Mamm Genome. 15:515–524.

Storey JD, Tibshirani R. 2003. Statistical significance for genome-wide studies. Proc Natl Acad Sci USA. 100:9440–9445.

Tao Y, Hartl DL. 2003. Genetic dissection of hybrid incompatibilities between Drosophila simulans and D. mauritiana. III. Heterogeneous accumulation of hybrid incompatibilities, degree of dominance, and implications for Haldane's rule. Evolution. 57:2580–2598.

Tao Y, Zeng ZB, Li J, Hartl DL, Laurie CC. 2003. Genetic dissection of hybrid incompatibilities between Drosophila simulans and D. mauritiana. II. Mapping hybrid male sterility loci on the third chromosome. Genetics. 164:1399–1418.

Taylor BA, Wnek C, Kotlus BS, Roemer N, MacTaggart T, Phillips SJ. 1999. Genotyping new BXD recombinant inbred mouse strains and comparison of BXD and consensus maps. Mamm Genome. 10:335–348.

Ting CT, Tsaur SC, Wu ML, Wu CI. 1998. A rapidly evolving homeobox at the site of a hybrid sterility gene. Science. 282:1501–1504.

Trachtulec Z, Mihola O, Vlcek C, Himmelbauer H, Paces V, Forejt J. 2005. Positional cloning of the Hybrid sterility 1 gene: fine genetic mapping and evaluation of two candidate genes. Biol J Linn Soc. 84:637–641.

True JR, Weir BS, Laurie CC. 1996. A genome-wide survey of hybrid incompatibility factors by the introgression of marked segments of Drosophila mauritiana chromosomes into Drosophila simulans. Genetics. 142: 819–837.

Tucker PK, Sage RD, Warner J, Wilson AC, Eicher EM. 1992. Abrupt cline for sex chromosomes in a hybrid zone between two species of mice. Evolution. 46:1146–1163.

Wade CM, Daly MJ. 2005. Genetic variation in laboratory mice. Nat Genet. 37:1175–1180.

Wade CM, Kulbokas EJ 3rd, Kirby AW, Zody MC, Mullikin JC, Lander ES, Lindblad-Toh K, Daly MJ. 2002. The mosaic structure of variation in the laboratory mouse genome. Nature. 420:574–578.

Williams RW, Gu J, Qi S, Lu L. 2001. The genetic structure of recombinant inbred mice: high-resolution consensus maps for complex trait analysis. Genome Biol. 2:RESEARCH0046.

Wiltshire T, Pletcher MT, Batalov S, Barnes SW, Tarantino LM, Cooke MP, Wu H, Smylie K, Santrosyan A, Copeland NG, et al. 2003. Genome-wide single-nucleotide polymorphism analysis defines haplotype patterns in mouse. Proc Natl Acad Sci USA. 100:3380–3385.

Witmer PD, Doheny KF, Adams MK, Boehm CD, Dizon JS, Goldstein JL, Templeton TM, Wheaton AM, Dong PN, Pugh EW, et al. 2003. The development of a highly informative mouse Simple Sequence Length Polymorphism (SSLP) marker set and construction of a mouse family tree using parsimony analysis. Genome Res. 13:485–491.

Wittbrodt J, Adam D, Malitschek B, Maueler W, Raulf F, Telling A, Robertson SM, Schartl M. 1989. Novel putative receptor tyrosine kinase encoded by the melanoma-inducing Tu locus in Xiphophorus. Nature. 341:415–421.

Wu CI, Beckenbach AT. 1983. Evidence for extensive genetic differentiation between the sex-ratio and the standard arrangement of Drosophila pseudoobscura and D. persimilis and identification of hybrid sterility factors. Genetics. 105:71–86.

Yalcin B, Fullerton J, Miller S, Keays DA, Brady S, Bhomra A, Jefferson A, Volpi E, Copley RR, Flint J, et al. 2004. Unexpected complexity in the haplotypes of commonly used inbred strains of laboratory mice. Proc Natl Acad Sci USA. 101:9734–9739.

Yonekawa H, Moriwaki K, Gotoh O, Watanabe J, Hayashi JI, Miyashita N, Petras ML, Tagashira Y. 1980. Relationship between laboratory mice and subspecies Mus musculus domesticus based on restriction endonuclease cleavage patterns of mitochondrial DNA. Jpn J Genet. 55:289–296.

Zhang J, Hunter KW, Gandolph M, Rowe WL, Finney RP, Kelley JM, Edmonson M, Buetow KH. 2005. A high-resolution multistrain haplotype analysis of laboratory mouse genome reveals three distinctive genetic variation patterns. Genome Res. 15:241–249.

Corresponding Editor: Loren Rieseberg