

Genome-wide patterns of gene flow across a house mouse hybrid zone

Katherine C. Teeter,^{1,2,5} Bret A. Payseur,³ Leslie W. Harris,^{1,2} Margaret A. Bakewell,² Lisa M. Thibodeau,^{1,2} Janelle E. O'Brien,^{1,2} James G. Krenz,⁴ Maria A. Sans-Fuentes,⁴ Michael W. Nachman,⁴ and Priscilla K. Tucker^{1,2,6}

¹Museum of Zoology, University of Michigan, Ann Arbor, Michigan 48109-1079, USA; ²Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan 48109-1048, USA; ³Laboratory of Genetics, University of Wisconsin, Madison, Wisconsin 53706, USA; ⁴Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona 85721, USA

Hybrid zones between closely related species or subspecies provide useful settings for studying the genetic architecture of speciation. Using markers distributed throughout the mouse genome, we use a hybrid zone between two recently diverged species of house mice (*Mus musculus* and *Mus domesticus*) as a natural mapping experiment to identify genomic regions that may be involved in reproductive isolation. Using cline analysis we document a nearly 50-fold variation in level of introgression among markers. Some markers have extremely narrow cline widths; these genomic regions may contribute to reproductive isolation. Biological processes associated with these narrow clines include physiological and immune responses to the environment as well as physiological and behavioral aspects of reproduction. Other autosomal markers exhibit asymmetrically broad clines, usually with high frequencies of *M. domesticus* alleles on the *M. musculus* side of the hybrid zone. These markers identify genome regions likely housing genes with alleles that are spreading from one species to the other. Biological processes associated with these wide clines include cell signaling, olfaction, and pheromone response. These processes play important roles in survival and reproduction, and associated genes are likely targets of selection. Patterns of linkage disequilibrium in the center of the hybrid zone suggest that isolation may be caused by multiple epistatic interactions between sets of genes. These data highlight the complex genetic architecture underlying speciation even at early stages of divergence and point to some of the biological processes that may govern this architecture.

[Supplemental material is available online at www.genome.org.]

The genetic basis of speciation is a central problem in evolutionary biology. Because reproductive isolation delineates the point at which emerging species will no longer freely exchange genes, considerable speciation research has focused on the genetics of reproductive isolation. In principle, reproductive isolation between taxa might be due to premating isolation, postmating but prezygotic isolation, or postzygotic isolation. Postzygotic isolation may be due to either inviability or sterility of hybrid individuals (either F1's or subsequent generations), or to some combination of these factors. Moreover, these fitness effects could be due to intrinsic genetic incompatibilities and/or they might depend heavily on ecological context (Coyne and Orr 2004). Historically, there have been two major approaches to studying the genetic basis of reproductive isolation. The first approach relies on crosses in the laboratory between species that display partial prezygotic or postzygotic reproductive isolation. The second approach utilizes naturally occurring hybrid populations to make inferences about the genetics of reproductive isolation.

Hybrid zone studies offer several advantages not inherent in studies based on laboratory crosses. They consider all aspects of hybrid fitness, including intrinsic genetic incompatibilities, in

addition to extrinsic ecological effects. They often provide many more generations of recombination than are obtainable in the lab, making it possible to pinpoint more precisely the chromosomal regions that are involved in reproductive isolation. In addition, even in cases where the specific phenotype involved in reproductive isolation is unknown or unclear, it may still be possible to detect the effects of selection through geographic patterns of changes in allele frequencies. For example, genes contributing to reproductive isolation are expected to introgress less than neutral markers (Hunt and Selander 1973; Harrison 1990; Tucker et al. 1992; Rieseberg et al. 1999; Buerkle and Rieseberg 2001). Thus, with reasonable coverage of markers throughout the genome, a hybrid zone can be used as a natural mapping experiment to identify the number and location of genomic regions that may be important in maintaining isolation between recently diverged taxa. Conversely, hybridization may be a source of beneficial alleles (Arnold et al. 1991; Rieseberg 1991, 1997). Advantageous alleles that are introduced from one species into another are expected to introgress faster and thus further (due to selection) than neutral genes.

House mice belonging to the *Mus musculus* species complex provide an excellent system for speciation research: They show intermediate levels of reproductive isolation, they can be crossed in the lab, they hybridize in nature, and a wealth of genetic and genomic tools are available. House mice are variably referred to in the literature as either distinct species or subspecies of *Mus musculus* and include *M. domesticus* (*M. musculus domesticus*) of

⁵Present address: Biology Department, Northern Michigan University, Marquette, Michigan 49855, USA.

⁶Corresponding author.

E-mail ptuck@umich.edu; fax (734) 763-4080.

Article published online before print. Article and publication date are at <http://www.genome.org/cgi/doi/10.1101/gr.6757907>.

Western Europe, North Africa, and the Middle East; *M. musculus* (*M. m. musculus*) of Eastern Europe and northern Asia; and *M. castaneus* (*M. m. castaneus*) of southeastern Asia. These three taxa are thought to have diverged from a common ancestor in west central Asia (Prager et al. 1998) or further east (Boursot et al. 1993, 1996) 0.35–0.9 million years ago (Mya) (She et al. 1990; Boursot et al. 1996; Suzuki et al. 2004). Through laboratory crosses between *M. musculus*, *M. domesticus*, and *M. castaneus* or strains derived from these taxa, some components of reproductive isolation have been identified. Most notably, there is clear evidence of hybrid male sterility from multiple studies (Forejt and Ivanyi 1975; Forejt 1996; Alibert et al. 1997; Oka et al. 2004, 2007; Storchova et al. 2004; Britton-Davidian et al. 2005; Trachtulec et al. 2005; Vyskocilova et al. 2005). There is evidence for limited female sterility in some crosses (Britton-Davidian et al. 2005) but not in others (Forejt and Ivanyi 1975; Britton-Davidian et al. 2005). There is also evidence for weak premating isolation in some crosses (Laukaitis et al. 1997; Smadja and Ganem 2002; Smadja et al. 2004) but not in others (Smadja and Ganem 2005). Finally, several studies have revealed higher parasite loads in hybrid mice in nature (Sage et al. 1986a) and in the lab (Mouliat et al. 1993).

A hybrid zone between *M. musculus* and *M. domesticus* formed after the movement of *M. domesticus* into Western Europe within the last 3000 yr (Cucchi et al. 2005). This hybrid zone has been studied in six transects along its 2400-km length from Denmark to the Transcaucasus (Ursin 1952; Hunt and Selander 1973; van Zegeren and van Oortmerssen 1981; Schnell and Selander 1981; Sage et al. 1986a,b; Vanlerberghe et al. 1986, 1988a,b; Nance et al. 1990; Tucker et al. 1992; Prager et al. 1993; Munclinger et al. 2002; Macholán et al. 2003, 2007; Payseur et al. 2004; Bozikova et al. 2005; Britton-Davidian et al. 2005; Dod et al. 2005; Raufaste et al. 2005). Using several phenotypic traits, allozymes, and other molecular markers, these studies have documented clinal patterns of variation and have shown that the majority of mice in the hybrid zone have recombinant genotypes.

Previous studies have demonstrated limited introgression of markers on the X chromosome (Tucker et al. 1992; Dod et al. 1993; Payseur et al. 2004; Macholán et al. 2007), which supports the hypothesis that the X chromosome harbors a large number of genes causing hybrid sterility or inviability (Muller 1940, 1942; Hagen and Scriber 1989; Turelli and Orr 1995; Orr 1997; Turelli and Begun 1997; Presgraves and Orr 1998; Saetre et al. 2003). However, there are several reasons for expecting autosomal genes to play an important role in reproductive isolation as well. First, the autosomes contain 94.3% of the mouse genome. Previous studies have suggested that many genes may contribute to reproductive isolation in these taxa (Raufaste et al. 2005; Macholán et al. 2007); thus, it seems likely that at least some of these genes will be autosomal. Second, Haldane's Rule, the sterility or inviability of the heterogametic sex (Haldane 1922), seems to be due to epistatic interactions between recessive X-linked genes and autosomal dominant genes (Turelli and Orr 1995). Crosses between *Mus musculus* and *Mus domesticus* obey Haldane's Rule (males preferentially show reduced fertility), and thus hybrids are expected to have autosomal loci involved in incompatibilities. Finally, several laboratory crosses using inbred lines have implicated a few autosomal regions in reproductive isolation between these taxa (Forejt and Ivanyi 1975; Forejt 1996; Britton-Davidian et al. 2005; Oka et al. 2007).

Despite the expected importance of autosomal genes in spe-

ciation, there have been no comprehensive studies of reproductive isolation in wild hybrid mice using markers across the genome. Here, we report on the differential introgression of loci across a hybrid zone in Bavaria, Germany using markers located on all mouse autosomes. Our goals are to (1) identify the number and location of autosomal regions showing reduced introgression as a means of finding genomic regions contributing to reproductive isolation, (2) identify genomic regions showing unusually high levels of introgression as a means of finding genomic regions that may contain beneficial alleles, and (3) use functional annotations of genes in regions of broad versus limited introgression to get a first look at biological processes associated with both patterns.

Results

Cline analysis

Scatter plots for all single nucleotide polymorphism (SNP) markers are shown in Supplemental Figures 1 and 2. Most changes in allele frequency occurred over a short distance; the median cline width was only 28.6 km (Table 1). Despite this small median cline width, there was enormous variation among loci in the level of introgression (Table 1; Supplemental Table 1). Three autosomal markers on three different chromosomes had cline widths <10 km: marker 5.097 (width 6.5 km), marker 9.075 (width 6.4 km), and marker 16.014 (width 7.0 km) (Table 1). These were the lowest autosomal cline widths observed. They are similar to the two-parameter estimates of cline widths for all but one of the 13 X-linked markers studied in Payseur et al. (2004). (The average two-parameter estimate of cline width for 12 X-linked markers [excluding marker X.100, with a cline width of 270.8 km] is 9.3 [range = 3.4–42.0 km, data not shown].) At the other extreme, eight markers had cline widths >100 km: markers 3.007, 4.057, 7.083, 8.078, 10.045, 12.031, 13.029, and 18.028. The average cline width for these eight markers was 165 km, or more than five times the median value for all markers. As extreme outliers, they are reasonable candidates for harboring genes under positive selection.

There was a strong positive correlation between cline center and cline width ($P < 0.001$ Spearman's rank correlation test; Fig. 1). Wider clines had centers more to the east along the transect, indicating a pattern of asymmetric introgression, from *M. domesticus* to *M. musculus*. Evidence for introgression from *M. musculus* to *M. domesticus* was minimal and mainly limited to mice collected from the Augsburg Zoo. The presence of *M. musculus* alleles in this westernmost locality is likely the result of passive (long-distance) transport with zoo animals.

Genomic attributes of marker regions

We investigated several genomic features to identify attributes of regions showing high or low levels of introgression. The number of genes in the 10-Mb windows centered on each marker ranged from 46 to 231, and, excluding marker 13.029, there was a significant negative correlation between gene density and cline width (Fig. 2; $R^2 = 0.102$, $P = 0.050$). We also calculated the mean value of d_N/d_S between mouse and rat for all genes in each 10-Mb window. There was a significant positive correlation between mean d_N/d_S and cline width, although this pattern was driven by a single marker (13.029) with a window showing an unusually high mean d_N/d_S value. When this marker was removed from the analysis, the regression was no longer significant. The correlation between GC content and cline width was not significant, even

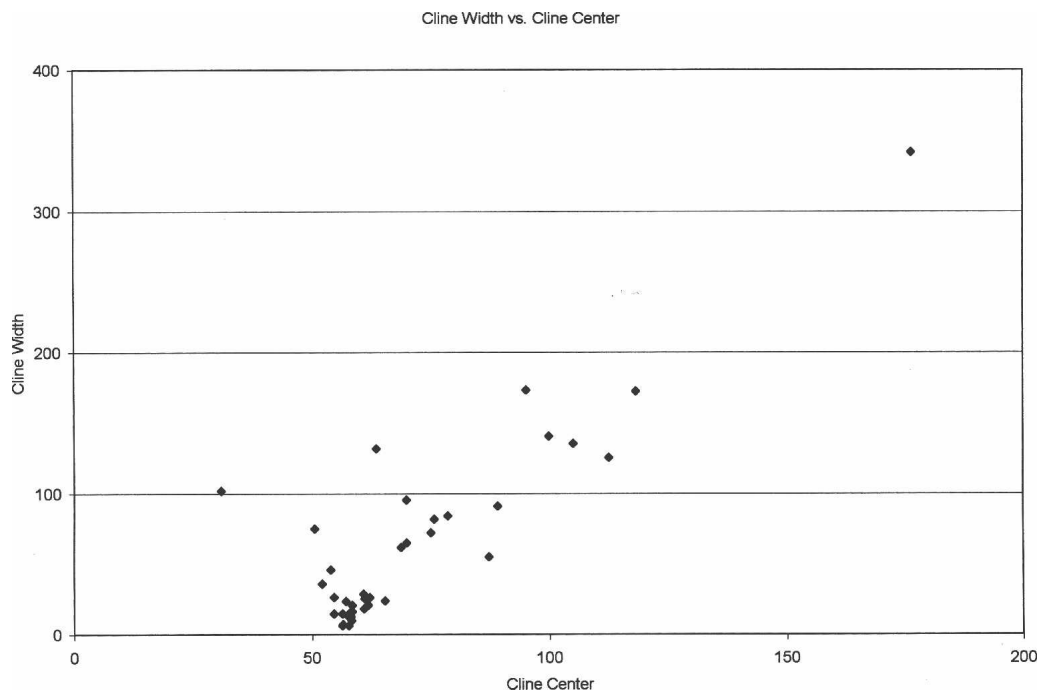


Figure 1. Plot of cline width vs. cline center. These data were generated from two-parameter models of cline shape for 39 autosomal markers.

including the marker 13.029 ($R^2 = 0.099$, $P = 0.051$). Surprisingly, no significant correlation was observed between the local rate of recombination and cline width, although this may be due partly to the fact that recombination rates vary relatively little in the mouse (Shifman et al. 2006).

Linkage disequilibrium (LD)

Theory predicts that reproductive isolation is most likely to be due to epistatic interactions (Dobzhansky–Muller incompatibilities), rather than to the effects of single underdominant genes

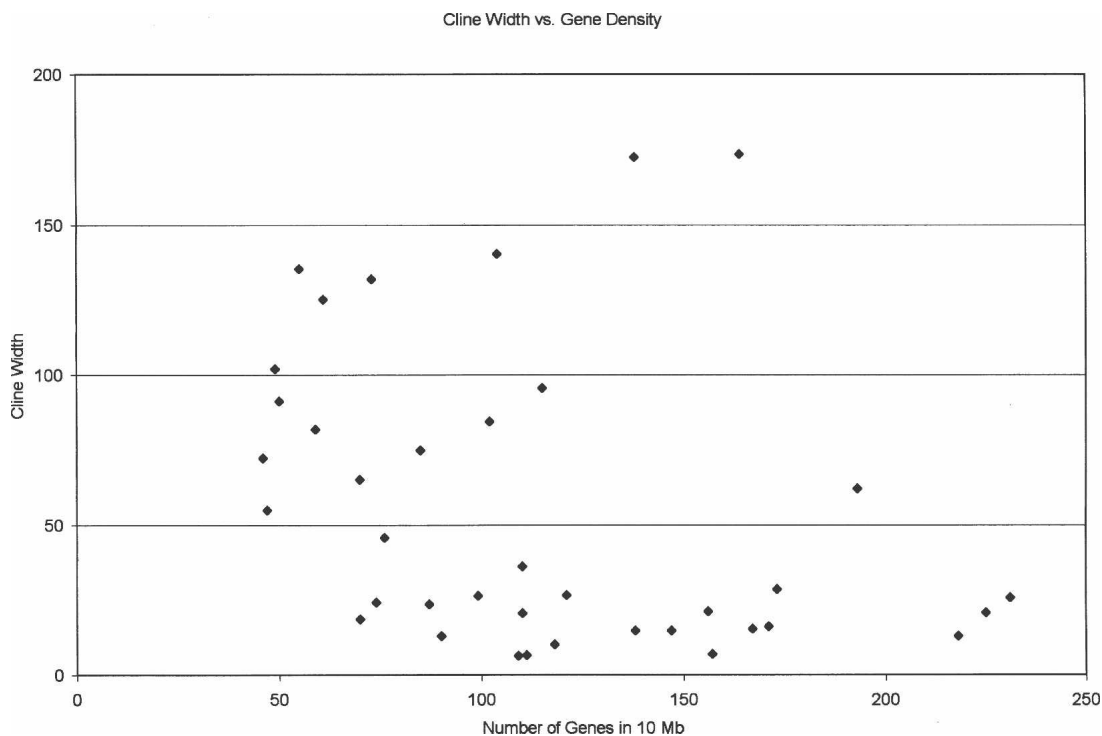


Figure 2. Plot of cline width versus gene density.

Table 1. Genomic location, cline widths, and genomic attributes for 39 SNP markers surveyed across the *Mus musculus*–*Mus domesticus* hybrid zone

SNP position (Chr.bp)	Abbrev. marker name	Cline width (km)	Cline center (km)	Number of genes	Rec. rate	%CG	d_N/d_S	d_N	d_S
1.14,026,920	1.014	18.5	60.9	70	0.36	43	0.16	0.04	0.22
1.46,402,091	1.046	81.9	75.7	59	0.50	43	0.16	0.03	0.19
1.158,689,899	1.159	36.1	52.1	110	0.43	44	0.21	0.04	0.20
2.30,409,503	2.030	20.7	58.4	225	0.73	50	0.17	0.03	0.20
2.78,369,208	2.078	74.9	50.5	85	0.70	45	0.16	0.04	0.19
2.164,903,540	2.165	21.1	61.7	156	0.74	49	0.18	0.05	0.25
3.7,368,144	3.007	101.9	30.9	49	0.18	44	0.18	0.05	0.24
3.140,416,764	3.140	65.1	69.9	70	0.93	43	0.21	0.07	0.25
4.57,302,353	4.057	140.4	99.9	104	1.26	44	0.20	0.04	0.19
4.129,074,493	4.129	13.0	57.5	218	0.86	50	0.14	0.03	0.22
5.6,836,406	5.007	45.8	53.9	76	0.43	44	0.16	0.03	0.22
5.96,799,978	5.097	6.5	57.7	111	0.86	44	0.22	0.05	0.23
6.88,319,986	6.088	15.4	57.9	167	0.51	48	0.19	0.03	0.18
6.112,706,955	6.113	26.3	62.0	99	0.69	47	0.16	0.03	0.18
7.83,170,743	7.083	172.4	118.2	138	0.68	45	0.18	0.03	0.19
7.126,200,262	7.126	62.0	68.7	193	0.93	49	0.19	0.03	0.17
8.77,724,048	8.078	173.4	95.1	164	0.50	51	0.15	0.04	0.25
8.100,982,853	8.101	12.9	58.1	90	0.59	47	0.19	0.04	0.22
9.52,309,582	9.052	14.8	56.3	138	0.50	46	0.17	0.03	0.19
9.75,141,481	9.075	6.4	56.3	109	0.81	44	0.19	0.04	0.19
10.44,829,603	10.045	131.9	63.6	73	0.50	45	0.22	0.05	0.26
10.54,662,870	10.055	24.1	65.2	74	0.44	44	0.13	0.02	0.21
11.52,567,023	11.053	28.6	60.7	173	0.59	47	0.21	0.04	0.19
11.88,642,186	11.089	10.1	58.2	118	0.67	46	0.16	0.04	0.20
12.30,640,920	12.031	125.2	112.5	61	0.62	45	0.16	0.03	0.18
12.99,182,498	12.099	23.6	57.1	87	0.78	46	0.26	0.05	0.21
13.28,996,604	13.029	341.8	176.4	87	0.74	42	0.30	0.07	0.21
13.55,641,923	13.056	26.5	54.6	121	0.71	49	0.14	0.03	0.24
14.30,598,525	14.031	95.6	69.9	115	0.43	46	0.17	0.04	0.21
14.73,920,455	14.074	84.4	78.5	102	0.75	46	0.18	0.04	0.20
15.64,824,507	15.065	91.2	89.1	50	0.53	46	0.21	0.03	0.16
15.99,541,517	15.099	25.7	61.0	231	1.10	50	0.12	0.03	0.23
16.13,815,994	16.014	7.0	56.5	157	0.94	49	0.19	0.04	0.20
17.46,219,311	17.046	14.7	54.6	147	1.07	48	0.18	0.04	0.22
17.90,955,861	17.091	54.9	87.2	47	0.54	46	0.18	0.05	0.25
18.28,302,931	18.028	135.4	105.0	55	0.72	46	0.16	0.03	0.21
18.63,609,404	18.064	20.6	61.7	110	0.99	46	0.12	0.03	0.22
19.44,127,651	19.044	16.2	58.4	171	0.77	47	0.16	0.03	0.22
19.52,101,649	19.052	72.2	75.0	46	1.23	47	0.15	0.04	0.26
AVERAGES		62.5	70.4	114	0.70	46	0.18	0.04	0.21
MEDIANS		28.6	61.0	109	0.70	46	0.18	0.04	0.21

Attributes are measured for a 10-Mb window centered on each marker.

(Coyne and Orr 2004). We studied nonrandom associations of alleles at different loci in one hybrid population near the middle of the hybrid zone (Neufahrn) to identify genomic regions that might be involved in such epistatic interactions. Observed values of R^2 from the Neufahrn population ranged from 0 to 0.321, with a mean of 0.036. In the absence of selection and in a randomly mating population at equilibrium, the expected value of R^2 is near 0. Slightly but significantly more than half (603) of the 1127 tests showed associations among conspecific alleles ($P = <10^{-6}$; binomial test), indicating a bias toward conspecific, two-locus genotypes. After accounting for multiple tests, 68 SNP pairs showed significant linkage disequilibrium, and conspecific alleles were associated in 42 of these cases (Table 2). Since different chromosomes assort independently each generation, and since the average R^2 in this population is near 0, these exceptions might be the result of epistatic incompatibilities. Some SNPs showed LD with several genomic regions (Table 2), suggesting the possibility of complex interactions involving multiple loci. However, the 68 locus pairs showing significant LD did not show a greater proportion of conspecific/heterospecific associations when compared with the 1059 nonsignificant locus pairs

($P = 0.528$; Fisher's exact test). Additionally, there was no evidence for a bias toward X–autosome pairs (vs. autosome–autosome pairs) among significant tests ($P = 0.451$; Fisher's exact test).

PANTHER analysis of genes

We classified genes into functional categories in order to search for functions that were overrepresented in genomic regions of high or low introgression. In total, 12,437 genes found within 15 Mb of each marker used in the study were classified into 240 different categories of biological processes. Of these genes, 2315 (18.6%) were not recognized by PANTHER, 6133 (49.3%) were classified into one or more biological process categories, and the remaining 3989 (32.1%) were unclassified.

In the analysis of genes in 10-Mb windows, four biological processes were significantly associated with wide clines: G-protein-mediated signaling, pheromone response, chromatin packaging and remodeling, and mesoderm development (Table 3). In the analysis of genes in the 30-Mb windows, the first three of the biological processes listed above were associated with

Table 2. Locus pairs from different chromosomes that show significant conspecific linkage disequilibrium

SNP 1 ^a	SNP 2	R ²	Q ^b
1.014	15.065	0.183	0.022
1.014	X.160	0.153	<0.001
1.046	5.007	0.195	<0.001
1.046	6.113	0.165	0.034
1.046	19.052	0.167	0.042
1.159	2.078	0.145	0.034
1.159	15.099	0.148	<0.001
1.159	X.107	0.136	0.022
1.159	X.132	0.122	0.045
2.078	6.113	0.187	0.045
2.078	17.046	0.321	<0.001
2.078	19.052	0.134	0.022
2.078	X.107	0.173	<0.001
2.078	X.129	0.138	0.034
4.129	14.031	0.167	<0.001
5.007	11.089	0.212	<0.001
6.113	19.052	0.227	<0.001
6.113	X.107	0.156	0.022
7.126	16.014	0.237	<0.001
9.052	14.031	0.130	0.045
9.052	18.064	0.171	<0.001
9.052	X.052	0.167	0.022
9.052	X.064	0.243	<0.001
9.052	X.070	0.179	<0.001
9.052	X.089	0.152	0.022
9.052	X.099 ^a	0.251	<0.001
9.052	X.099 ^b	0.131	0.045
9.075	12.099	0.130	0.049
9.075	19.052	0.147	0.034
10.045	15.099	0.127	0.042
10.055	X.033	0.136	0.042
11.053	X.089	0.128	0.045
11.089	15.065	0.084	0.045
11.089	X.107	0.128	0.049
11.089	X.132	0.138	0.045
12.099	X.129	0.192	<0.001
14.031	X.033	0.129	0.049
14.031	X.160	0.158	0.034
17.046	19.052	0.162	0.049
18.064	X.099 ^a	0.182	0.034
18.064	X.160	0.130	0.049
19.052	X.129	0.126	0.042

^aAutosomal markers are sorted according to chromosome as in Table 1. X-linked markers are from Payseur and Nachman (2005) with positions updated using Build 36.

^bQ-value assuming a false discovery rate of 5% across all tests.

wide clines, as well as three additional biological processes: cell-adhesion-mediated signaling, cell adhesion, and olfaction. Cell adhesion, however, showed an opposite but nonsignificant trend in the analysis of the 10-Mb windows. There is a cluster of 60 protocadherin genes located 5–15 Mb upstream of marker 18.028, which has a cline width of 135.4 km. These genes are not in the 10-Mb windows, which accounts for the difference in the trends between the 10-Mb and 30-Mb analyses. Finally, mesoderm development showed a nonsignificant trend toward a wide cline in the 30-Mb analysis. Thus, there is good agreement between 10-Mb and 30-Mb windows in the identification of processes significantly associated with wide clines.

In the analysis of biological processes that were significantly associated with narrow clines, three processes were significant using 30-Mb windows: MHCI-mediated immunity, steroid hormone metabolism, and cell structure. The latter two processes showed similar but nonsignificant trends in the analysis of 10-Mb windows.

Discussion

We documented patterns of introgression across the mouse genome and documented regions of both extensive and restricted gene flow between recently diverged species of house mice. An earlier study of this same transect (Payseur et al. 2004) identified a region in the center of the X chromosome with an especially narrow cline. This same region was also identified as harboring a locus involved in hybrid male sterility between *M. musculus* and *M. domesticus* mice (Oka et al. 2004; Storchova et al. 2004), suggesting that cline shape can be used to pinpoint regions of the genome involved in reproductive isolation. Here, we find similarly narrow clines for multiple autosomal markers, and these are also likely found in genomic regions involved in reproductive isolation. In fact, one marker, 11.089, with a cline width of 10.1 km, is in close proximity to a recently identified QTL interacting with the center of the X chromosome to produce hybrid male sterility in these taxa (Oka et al. 2007). The fact that cline width is negatively associated with gene density suggests that selection, rather than random effects, determines the degree of introgression for each marker. It also suggests that the number of loci underlying reproductive isolation is rather large (Payseur and Nachman 2005).

Harr (2006) identified ten genomic regions, two on the X and eight on autosomes, of elevated differentiation between wild derived strains of *M. musculus* and *M. domesticus* and suggested they are likely to harbor genes associated with hybrid unfitnes. Most of the strains used by Harr came from within the ranges of *musculus* and *domesticus*, not from near the hybrid zone. Four of our markers, 1.159, 8.078, 10.045, and 10.055, fall within the regions of elevated differentiation identified by Harr. Two of these markers, 1.159 and 10.055, have relatively narrow cline widths of 36.1 km and 24.1 km, respectively. However, the other two markers, 8.078 and 10.045, have relatively wide cline widths of 173.4 km and 131.9 km, respectively. In addition, both chromosome 10 markers, with very different patterns of introgression despite being located ~10 Mb apart, fall within a single region of elevated differentiation that is 31 Mb in size (Harr 2006). Thus, while there is some concordance between the results presented here and those of Harr (2006), it is clear that regions showing reduced introgression in the hybrid zone are not always the same as those showing increased *F_{st}* between samples taken outside of the hybrid zone.

Reasoning that selection against heterospecific combinations of alleles that cause inviability or sterility can maintain nonrandom associations between loci in hybrid populations (Gardner et al. 2002; Payseur and Hoekstra 2005), LD among SNP pairs was measured. Several patterns emerge that provide evidence for natural selection. First, despite the fact that there is little overall LD in the middle of the hybrid zone, a number of SNP pairs show significant associations after correcting for multiple tests. Many of these significant associations show a bias against heterospecific genotypes, suggesting that selection tends to remove these interlocus combinations in the center of the hybrid zone. Second, these SNP pairs map to different chromosomes, demonstrating that linkage is not responsible for the maintenance of these associations. Third, several SNPs show strong associations with multiple, unlinked genomic regions. Because chance (and false-positive) associations are not expected to involve the same loci, this pattern also suggests the presence of functional interactions. Furthermore, these associations involving the same marker with two or more other markers suggest that

Table 3. Biological processes with nonrandom distribution of cline widths in the Bavaria transect

Biological process	30-Mb windows				10-Mb windows			
	# of genes	Mean cline width		Q-value	# of genes	Mean cline width		Q-value
		Within category	Other genes			Within category	Other genes	
Narrow clines								
MHCI-mediated immunity	35	15.74	57.03	1.15E-04^a	0			
Steroid hormone metabolism	27	32.37	56.91	0.0047^a	4	38.75	59.71	1
Cell structure	302	49.49	57.16	0.0074^a	106	50.22	60.14	1
Wide clines								
Cell-adhesion-mediated signaling ^b	186	78.92	56.17	6.99E-09^a	37	45.96	59.90	1
G-protein-mediated signaling	612	70.86	55.35	4.26E-06^a	137	100.06	57.08	1.98E-04^a
Cell adhesion	266	67.47	56.36	7.01E-05^a	61	49.94	59.94	0.5350
Pheromone response	36	115.03	56.50	7.01E-05^a	19	122.54	59.14	1.98E-04^a
Chromatin packaging and remodeling	134	130.52	55.27	2.06E-04^a	62	156.95	56.94	2.18E-04^a
Olfaction	300	58.81	56.72	6.25E-04^a	22	51.81	59.75	0.6850
Mesoderm development	88	120.90	55.93	0.2165	43	196.98	57.02	8.82E-05^a

^aBoldface indicates significant Q-values at 5% FDR. P -value ≤ 0.05 using Bonferroni correction for 240 categories.

^bGenes involved with cell-adhesion-mediated signaling are associated with wide clines in the 30-Mb window analysis and with narrow clines in the 10-Mb window analysis. See Discussion.

selection may be acting on complex epistatic interactions (i.e., involving three or more genes), even in early stages of speciation (Coyne and Orr 2004). It is also worth noting that some markers with narrow cline widths show strong associations with each other. For example, 9.052, which has a cline width of 14.8 km, is in linkage disequilibrium with X.089, a marker with very little introgression (Payseur et al. 2004). Because patterns of LD and introgression provide independent ways of identifying incompatibility candidates, these marker pairs seem worthy of further attention.

Hybrid populations of laboratory mice have also been surveyed for linkage disequilibrium between unlinked SNPs (Payseur and Hoekstra 2005; Payseur and Place 2007). The regions surveyed in these studies are generally not the same as those surveyed in the present study. Nevertheless, of the 23 individual autosomal SNP markers involved in extreme associations between conspecific alleles in laboratory mice (Payseur and Hoekstra 2005), 10 were <10 Mb from SNP markers surveyed in our analysis. Interestingly, eight of these markers show narrow cline widths ranging in size from 6.4 to 26.3 km. The remaining two markers were furthest from the SNPs identified in Payseur and Hoekstra (2005) and had the widest cline widths, 45.8 and 62 km, respectively. Agreement between these data sets in the identification of regions potentially involved in reproductive isolation reinforces the notion that reproductive isolation between these taxa involves disrupted interactions between multiple autosomal loci, "Dobzhansky-Muller incompatibilities" (Dobzhansky 1937; Muller 1942; Coyne and Orr 2004).

The only study to explicitly map multiple partners of incompatibilities that underlie reproductive isolation phenotypes in mice is Oka et al. (2007). In this study, a region from the middle of the X chromosome of *M. molossinus* (a mixture of Asian *M. musculus* and *M. castaneus* genomes) caused hybrid male sterility when combined with regions on chromosomes 1, 9, and 11 from C57BL6 (an inbred strain primarily descended from *M. domesticus*). The central part of the X chromosome also shows significant conspecific associations with SNPs from chromosomes 1, 9, and 11 in the hybrid zone (Table 2). Although the large size of the mapped regions and low density of markers surveyed in the

hybrid zone preclude strong conclusions, the agreement between these two studies suggests that some loci showing strong linkage disequilibrium are linked to incompatibilities associated with hybrid male sterility.

The PANTHER analysis of 10- and 30-Mb windows provides preliminary information on the biological processes of genes that are nonrandomly associated with regions of both reduced and extensive introgression. The 30-Mb analysis has greater statistical power than the 10-Mb analysis due to the larger number of genes, resulting in a larger number of significant associations. However, the biological meaning of these associations is less clear at the larger window size. Nevertheless, these data suggest that reproductive isolation may be driven by multiple kinds of interactions. These include ecological responses such as immune function. Interestingly, higher loads of intestinal parasites have been documented in mice from the hybrid zone, indicating possible reduced viability of hybrids (Sage et al. 1986a; Moulia et al. 1991, 1993). In addition, genes associated with physiological and behavioral aspects of reproduction involving steroid hormone metabolism appear to be associated with reduced introgression. Inspection of the expression patterns (SymAtlas v1.2.4; Su et al. 2002) for cell structure genes associated with the nine narrowest clines suggests involvement of a wide range of tissues, but with the majority being associated with sensory apparatuses such as the tongue, snout epidermis, digits, and epidermis.

Across the full set of markers, introgression is found to be strongly asymmetric. If a marker has extensive introgression across the hybrid zone, it most often occurs only from *M. domesticus* to *M. musculus*. With the exception of the Czech transect (Macholán et al. 2007), this same pattern has been documented in other studies of the *Mus* hybrid zone (Vanlerberghe et al. 1986; Tucker et al. 1992; Dod et al. 1993, 2005; Payseur et al. 2004; Raufaste et al. 2005). There are four possible explanations for this pattern, which are not mutually exclusive. The first possibility is that there are intrinsic genomic incompatibilities between *M. domesticus* alleles and *M. musculus* alleles, with many more *M. musculus* alleles being incompatible with a *M. domesticus* genetic background than the converse. The second possibility is that mate preference is asymmetric between the two species. The

third possibility is the dispersal patterns differ between the two species, and the fourth possibility is that the hybrid zone has shifted over time.

There is evidence for asymmetric mating preferences between *M. domesticus* and *M. musculus*. A preference for conspecific urine signals has been shown in *M. musculus*, but no such preference is apparent in *M. domesticus* (Smadja and Ganem 2002; Smadja et al. 2004). There also seems to be a trend toward conspecific preferences in *M. musculus* for salivary androgen-binding proteins (Bimova et al. 2005). A lack of conspecific mate preference in *M. domesticus* could contribute to the observed asymmetric pattern of introgression, assuming that *M. domesticus* simply are less choosy about their mates, and more likely to mate with hybrid mice and *M. musculus*.

There is also behavioral evidence that *M. domesticus* is dominant to *M. musculus* in male–male competition (van Zegeren and van Oortmerssen 1981), and that the *M. domesticus* males tend to be more aggressive. The behavioral dominance of *M. domesticus* could result in a shift of the position of the hybrid zone to the east (moving from *M. domesticus* toward *M. musculus*).

Preliminary analyses based on partial annotation of the mouse genome indirectly corroborate the laboratory studies on aggression and mating behavior, as genes involved in pheromone response are nonrandomly associated with genome regions of extensive introgression. These include regions housing vomeronasal receptor genes and genes coding for mouse urinary protein (*Mup*). There are five *Mup* genes located within 5 Mb of marker 4.057, with a cline width of 140 km. They bind a variety of pheromones affecting mouse physiology and behavior including estrus, puberty, and intermale aggression (Novotny 2003). Quite possibly, *M. domesticus* alleles at these loci confer a fitness advantage over *M. musculus* alleles.

The other processes nonrandomly associated with wide clines include genes broadly associated with cell signaling and with olfaction. The latter process includes a large number ($n = 300$) of olfactory receptor loci that, upon closer inspection, are associated with either narrow or wide clines. These rapidly evolving loci are obvious targets of selection as they play important roles in survival and reproduction and likely reflect responses to the environment (Lane et al. 2001).

Conclusion

This study represents the most detailed genetic survey of a vertebrate hybrid zone. With markers situated on all mouse autosomes, we have used patterns of introgression to map genomic regions contributing to the maintenance of genetic isolation between recently diverged species. Patterns of LD corroborate that isolation is likely caused by epistatic interactions between sets of parental alleles. We find multiple regions of the genome with narrow clines, and these regions probably house genes involved in reproductive isolation. They include genes associated with a variety of biological processes including reproductive physiology and behavior and physiological and sensory responses to the environment. Autosomal regions with wide asymmetric clines have different effects on the fitness of hybrid mice. Interestingly, these regions also include genes involved in reproductive behavior and sensory responses to the environment. A more fully annotated genome coupled with denser sampling of the genome for patterns of introgression along with comparisons of introgression across multiple transects will provide a more complete un-

derstanding of the genetic underpinnings of reproductive isolation in this system.

Methods

Sampling

Four hundred forty-nine mice used in this study were collected by R.D. Sage from a transect through the hybrid zone in the German state of Bavaria, and western Austria. Collecting for this transect was performed by R.D. Sage in 1984, 1985, and 1992. The location of the hybrid zone and the transect are shown in Payseur et al. (2004), and information on collecting localities and numbers of mice from each locality is found in Supplemental Table 2. Sampling was performed in a roughly linear, east–west manner, and transect distances (in kilometers) were calculated from the western end of the transect.

Development and scoring of molecular markers

Single nucleotide markers for autosomes were identified using the SNP database described by Lindblad-Toh et al. (2000), <http://www.broad.mit.edu/snp/mouse/>, and the later SNP database from the mouse genome project, <http://www.ncbi.nlm.nih.gov/projects/SNP/MouseSNP.cgi>. To identify SNPs with fixed differences between *M. domesticus* and *M. musculus*, markers were sequenced in 10 allopatrically distributed mice from each species (Supplemental Table 3). The genomic location of markers was originally checked against the UCSC Mouse Genome database (build 33, May 2004, <http://genome.ucsc.edu>). Markers were chosen with the goal of having a relatively even marker density throughout the genome. Three markers were developed for the largest chromosomes, 1 and 2, and two markers per chromosome were developed for the remaining autosomes with the exception of chromosome 16. Markers fixed for alternate alleles in *M. domesticus* and *M. musculus* were used to design probes for TaqMan genotyping, through Applied Biosystems Assays-By-Design service. Approximately 500 bp of sequence centered on the SNP was checked for repeat regions using RepeatMasker (<http://www.repeatmasker.org/>). Reactions for TaqMan genotyping were set up using 5 μ L of TaqMan master mix, 40 μ L of the custom fluorescently labeled assay mix, \sim 0.01 μ g of DNA in a 3- μ L volume, and 1.75 μ L of sterile dH₂O in 96-well plates. Genotyping plates were run on a real-time PCR machine. Thirteen X-linked markers were taken from Payseur et al. (2004) and reanalyzed here to obtain two-parameter estimates of cline shape. Genotyping data for all markers is found in Supplemental Table 4.

Cline analysis

Cline shape was estimated individually for each marker. As the goal of these analyses was to allow comparison between markers, a relatively simple two-parameter model was used to estimate the center and width of the cline for each marker, similar to the analyses in Bozikova et al. (2005). This model describes the relationship between allele frequency and geographic distance for individuals along a transect. The position of the cline center and the cline width were used in comparisons between markers. This two-parameter model consists of a hyperbolic tangent function that describes a sinusoidal curve (Equation 1):

$$p = \frac{1}{2} \left[1 + \tanh\left(\frac{2[x - c]}{w}\right) \right]. \quad (1)$$

In this equation, p = allele frequency, x = geographic distance along a transect, c = the center of the hybrid zone, defined

as the point of steepest slope, and $w = 1/\text{slope}$ of the cline at the center. All cline analyses were performed using ClineFit software, available from Adam Porter (Porter et al. 1997) at <http://www-unix.oit.umass.edu/~aporter/software/index.html>. ClineFit estimates are based on data from individuals, using a Metropolis algorithm (Metropolis et al. 1953) to numerically estimate the parameters that describe the cline shape. Because the estimates are based on individuals, rather than populations, populations with small samples are weighted less than populations with large samples. We conducted some analyses excluding the populations with smallest samples sizes, and the results were unchanged. Two-unit support limits are derived from the likelihood searches for each parameter. These support limits are roughly analogous to 95% confidence intervals (Edwards 1992).

The data set presented here was also analyzed using six-parameter models (see Supplemental Methods). Spearman non-parametric rank correlation tests were used to detect correlations between the estimated cline parameters. Correlation tests were performed for two different types of comparisons: (1) comparing cline widths and centers for each marker, and (2) comparing the estimated cline widths with the local gene density. These tests were performed in SPSS 11.0 for Macintosh OS X.

Genomic attributes of marker regions

Genomic attributes were calculated for 10-Mb windows centered on each marker using Build 36 of the ENSEMBL mouse sequence. Previous work in the hybrid zone suggested that linkage disequilibrium extends beyond this distance (Payseur et al. 2004), so that windows of this size are reasonable for tracking the effects of linked sites. Gene density was measured by counting the number of known and predicted genes. The mean of d_N (number of non-synonymous substitutions per nonsynonymous site), d_S (number of synonymous substitutions per synonymous site), and d_N/d_S for all genes in each window were calculated by comparing mouse and rat sequences (M. Dean and J. Good, unpubl.). Outliers were excluded after the examination of scatter plots between d_N , d_S , d_N/d_S , and gene position in each window. Recombination rates were estimated based on the relationship between the genetic and physical maps of the *Mus musculus* genome, using data from 2293 heterogeneous stock (HS) mice (http://gscan.well.ox.ac.uk/#genetic_map; Shifman et al. 2006). Genetic positions are for 9904 single nucleotide polymorphisms (SNPs) typed in 2293 HS animals. Recombination rates were calculated as the slope of a linear regression comparing the genetic positions of SNPs (in centiMorgans, cM) against their physical position (in megabase pairs, Mb) for each 10-Mb window.

Linkage disequilibrium analysis

Linkage disequilibrium was measured in 63 mice from a single locality, Neufahrn bei Freising, in the center of the transect. Linkage disequilibrium was estimated using two-locus genotype counts (Weir et al. 2004). Specifically, genotypes at each locus were recoded as 0, 1, or 2 for autosomal loci and X-linked loci in females, and 0 or 2 for X-linked loci in males. Linkage disequilibrium was estimated as the squared Pearson's correlation coefficient (genotypic R^2) between these recoded genotypes. The probability of each observed R^2 value was estimated by comparison to 1000 R^2 values obtained by randomizing genotypes across individuals. We adjusted statistical significance to achieve a 5% false discovery rate (FDR) (Storey and Tibshirani 2003) across all tests.

Two markers, 17.091 (which was fixed for the *M. domesticus* allele in this locality) and 3.007 (which harbored a single *M. domesticus* allele), were not included in these analyses. To ensure

independent assortment among loci in each generation, linkage disequilibrium was measured for pairs of loci residing on different chromosomes. Because the X chromosome is involved in reproductive isolation between house mouse species (Oka et al. 2004; Storchova et al. 2004), we included 13 X-linked markers (Payseur et al. 2004) in our survey of linkage disequilibrium.

PANTHER analysis of genes

Using the PANTHER database (Thomas et al. 2003; Mi et al. 2005), genes found within 30 Mb of each marker used in the study were classified into different categories of biological processes. As biological process groups are not mutually exclusive, some genes belong to more than one group. In addition, some biological process groups are subsets of others (e.g., "Pheromone response" is a subset of "Sensory perception").

For each of the biological processes identified, genes with PANTHER classifications were divided into two groups: (1) those associated with the biological process in question and (2) all others. Genes unclassified or unrecognized by PANTHER were excluded from this analysis, but including these genes does not significantly change the results (not shown). Genes within 10-Mb and 30-Mb windows were ranked according to the cline widths for the Bavaria transect. Genes associated with more than one marker were included once for each associated marker. A nonrandom distribution of cline widths for genes in a particular biological process was identified by comparing groups 1 and 2. If genes for a particular biological process tend to be associated with narrow (or wide) clines, this tendency can be identified by a significant P -value from a two-tailed Mann-Whitney U-test. Because multiple biological processes were tested, it is necessary to correct for multiple testing. Q -values were calculated (Storey and Tibshirani 2003), and those values significant at a 5% false discovery rate (FDR) are reported. A Bonferroni correction for multiple tests was also used. This correction is especially conservative because some biological processes are subsets of others.

Acknowledgments

Richard Sage collected all of the animals from the Bavaria transect and generously provided DNA samples and tissues to both the Tucker and Nachman laboratories. Michael Hammer provided some DNAs from *M. musculus*. Matt Dean, Milos Macholán, and Jaroslav Piálek provided helpful comments on an earlier draft. Research was supported by NSF DEB0212667 to P.K.T., NSF DEB0213013 to M.W.N., and MEC/Fulbright grant from Secretaría de Estado de Universidades e Investigación from the Spanish Ministerio de Educación y Ciencia to M.A.S.-F.

References

- Alibert, P., Fel-Clair, F., Manolakou, K., Britton-Davidian, J., and Auffray, J.C. 1997. Developmental stability, fitness, and trait size in laboratory hybrids between European subspecies of the house mouse. *Evolution Int. J. Org. Evolution* **51**: 1284–1295.
- Arnold, S.J., Buckner, C.M., and Robinson, J.J. 1991. Pollen-mediated introgression and hybrid speciation in Louisiana irises. *Proc. Natl. Acad. Sci.* **88**: 1398–1402.
- Bimova, B., Karn, R.C., and Pialek, J. 2005. The role of salivary androgen-binding protein in reproductive isolation between two species of house mouse: *Mus musculus musculus* and *Mus musculus domesticus*. *Biol. J. Linn. Soc. Lond.* **84**: 349–361.
- Boursot, P., Auffray, J.C., Britton-Davidian, J., and Bonhomme, F. 1993. The evolution of house mice. *Amu. Rev. Ecol. Syst.* **24**: 119–152.
- Boursot, P., Din, W., Anand, R., Darviche, D., Dod, B., Von Deimling, F., Talwar, G.P., and Bonhomme, F. 1996. Origin and radiation of the house mouse: Mitochondrial DNA phylogeny. *J. Evol. Biol.* **9**: 391–415.

- Bozikhov, E., Munclinger, P., Teeter, K.C., Tucker, P.K., Macholán, M., and Piálek, J. 2005. Mitochondrial DNA in the hybrid zone between *Mus musculus musculus* and *Mus musculus domesticus*: A comparison of two transects. *Biol. J. Linn. Soc. Lond.* **84**: 363–378.
- Britton-Davidian, J., Fel-Clair, F., Lopez, J., Alibert, P., and Boursot, P. 2005. Postzygotic isolation between the two European subspecies of the house mouse: Estimates from fertility patterns in wild and laboratory-bred hybrids. *Biol. J. Linn. Soc. Lond.* **84**: 379–393.
- Buerkle, C.A. and Rieseberg, L.H. 2001. Low intraspecific variation for genomic isolation between hybridizing sunflower species. *Evolution Int. J. Org. Evolution* **55**: 684–691.
- Coyne, J.A. and Orr, H.A. 2004. *Speciation*. p. 545. Sinauer Associates, Inc., Sunderland, MA.
- Cucchi, T., Vigne, J.D., and Auffray, J.C. 2005. First occurrence of the house mouse (*Mus musculus domesticus* Schwarz and Schwarz, 1943) in the Western Mediterranean: A zooarchaeological revision of subfossil occurrences. *Biol. J. Linn. Soc. Lond.* **84**: 429–445.
- Dobzhansky, T. 1937. *Genetics and the origin of species*, p. 364. Columbia University Press, New York.
- Dod, B., Jermiin, L.S., Boursot, P., Chapman, V.H., Nielsen, J.T., and Bonhomme, F. 1993. Counterselection on sex-chromosomes in the *Mus musculus* European hybrid zone. *J. Evol. Biol.* **6**: 529–546.
- Dod, B., Smadja, C., Karn, R.C., and Boursot, P. 2005. Testing for selection on the androgen-binding protein in the Danish mouse hybrid zone. *Biol. J. Linn. Soc. Lond.* **84**: 447–459.
- Edwards, A.W.F. 1992. *Likelihood*, p. 275. Johns Hopkins University Press, Baltimore, MD.
- Forejt, J. 1996. Hybrid sterility in the mouse. *Trends Genet.* **12**: 412–417.
- Forejt, J. and Ivanyi, P. 1975. Genetic studies on male sterility of hybrids between laboratory and wild mice (*Mus musculus* L.). *Genet. Res.* **24**: 189–206.
- Gardner, K., Buerkle, A., Whitton, J., and Rieseberg, L.H. 2002. Inferring epistasis in wild sunflower hybrid zones. In *Epistasis and the evolutionary process* (eds. J.B. Wolf et al.), pp. 264–279. Oxford University Press, New York.
- Hagen, R.H. and Scriber, J.M. 1989. Sex-linked diapause, color, and allozyme loci in *Papilio glaucus*: Linkage analysis and significance in a hybrid zone. *J. Hered.* **80**: 179–185.
- Haldane, J.B.S. 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, R.G. 1990. Hybrid zones: Windows on evolutionary process. *Oxf. Surv. Evol. Biol.* **7**: 69–128.
- Hunt, W.G. and Selander, R.K. 1973. Biochemical genetics of hybridisation in European house mice. *Heredity* **31**: 11–33.
- Lane, R.P., Cutforth, T., Young, J., Athanasiou, M., Friedman, C., Rowen, L., Evans, G., Axel, R., Hood, L., and Trask, B.J. 2001. Genomic analysis of orthologous mouse and human olfactory receptor loci. *Proc. Natl. Acad. Sci.* **98**: 7390–7395.
- Laukaitis, C.M., Critser, E.S., and Karn, R.C. 1997. Salivary androgen-binding protein (ABP) mediates sexual isolation in *Mus musculus*. *Evolution Int. J. Org. Evolution* **51**: 2000–2005.
- Lindblad-Toh, K., Winchester, E., Daly, M.J., Wang, D.G., Hirschhorn, J.N., Lavolette, J.P., Ardlie, K., Reich, D.E., Robinson, E., Sklar, P., et al. 2000. Large-scale discovery and genotyping of single-nucleotide polymorphisms in the mouse. *Nat. Genet.* **24**: 381–386.
- Macholán, M., Krytufek, E., and Vohralík, V. 2003. The location of the *Mus musculus* / *M. domesticus* hybrid zone in the Balkans: Clues from morphology. *Acta Theriol. (Warsz.)* **48**: 177–188.
- Macholán, M., Munclinger, P., Sugerková, M., Dufkova, P., Bimova, B., Zima, J., and Piálek, J. 2007. Genetic analysis of autosomal and X-linked markers across a mouse hybrid zone. *Evolution Int. J. Org. Evolution* **61**: 746–771.
- Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., and Teller, E. 1953. Equation of state calculations by fast computing machines. *J. Chem. Phys.* **21**: 1087–1092.
- Mi, H., Lazareva-Ulitsky, B., Loo, R., Kejarawal, A., Vandergriff, J., Rabkin, S., Guo, N., Muruganujan, A., Doremieux, O., Campbell, M.J., et al. 2005. The PANTHER database of protein families, subfamilies, functions and pathways. *Nucleic Acids Res.* **31**: 334–341.
- Mouliá, C.J., Aussel, P., Bonhomme, F., Boursot, P., Nielsen, J.T., and Renaud, F. 1991. Wormy mice in a hybrid zone: A genetic control of susceptibility to parasite infection. *J. Evol. Biol.* **4**: 679–687.
- Mouliá, C.J., LeBrun, N., Dallas, J., Orth, A., and 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, H.J. 1940. Bearing of the *Drosophila* work on systematics. In *The new systematics* (ed. J.S. Huxley), pp 185–268. Clarendon, Oxford, UK.
- Muller, H.J. 1942. Isolating mechanisms, evolution, and temperature. *Biol. Symp.* **6**: 71–125.
- Munclinger, P., Božíková, E., Sugerková, M., Piálek, J., and Macholán, M. 2002. Genetic variation in house mice (*Mus*, Muridae, Rodentia) from the Czech and Slovak Republics. *Folia Zool. (Brno)* **51**: 81–92.
- Nance, V., Vanlerberghe, F., Nielsen, J.T., Bonhomme, F., and Britton-Davidian, J. 1990. Chromosomal Introgression in house mice from the hybrid zone between *M. m. domesticus* and *M. m. musculus* in Denmark. *Biol. J. Linn. Soc. Lond.* **41**: 215–227.
- Novotny, M.V. 2003. Pheromones, binding proteins and receptor responses in rodents. *Biochem. Soc. Trans.* **31**: 117–122.
- Oka, A., Mita, A., Sakurai-Yamatani, N., Yamamoto, H., Takagi, N., Takano-Shimizu, T., Toshimori, K., Moriwaki, K., and Shiroishi, T. 2004. Hybrid breakdown caused by substitution of the X chromosome between two mouse subspecies. *Genetics* **166**: 913–924.
- Oka, A., Aoto, T., Totsuka, Y., Takahashi, R., Ueda, M., Mita, A., Sakurai-Yamatani, N., Yamamoto, H., Kuriki, S., Takagi, N., et al. 2007. Disruption of genetic interaction between two autosomal regions and the X chromosome causes reproductive isolation between mouse strains derived from different subspecies. *Genetics* **175**: 185–197.
- Orr, H.A. 1997. Haldane's rule. *Annu. Rev. Ecol. Syst.* **28**: 195–218.
- Payseur, B.A. and Hoekstra, H.E. 2005. Signatures of reproductive isolation in patterns of single nucleotide diversity across inbred strains of mice. *Genetics* **171**: 1905–1916.
- Payseur, B.A. and Nachman, M.W. 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. Lond.* **84**: 523–534.
- Payseur, B.A. and Place, M. 2007. Searching the genomes of inbred mouse strains for incompatibilities that reproductively isolate their wild relatives. *J. Hered.* **98**: 115–122.
- Payseur, B.A., Krenz, J.G., and Nachman, M.W. 2004. Differential patterns of introgression across the X chromosome in a hybrid zone between two species of house mice. *Evolution Int. J. Org. Evolution* **58**: 2064–2078.
- Porter, A.H., Wenger, R., Geiger, H., Scholl, A., and Shapiro, A.M. 1997. The *Pontia daplidice-edusa* hybrid zone in Northwestern Italy. *Evolution Int. J. Org. Evolution* **51**: 1561–1573.
- Prager, E.M., Sage, R.D., Gyllensten, U., Thomas, W.K., Hubner, R., Jones, C.S., Noble, L., Searle, J.B., and Wilson, A.C. 1993. Mitochondrial DNA sequence diversity and the colonization of Scandinavia by house mice from East Holstein. *Bio. J. Linn. Soc. Lond.* **50**: 85–122.
- Prager, E.M., Orrego, C., and Sage, R.D. 1998. Genetic variation and phylogeography of central Asian and other house mice, including a major new mitochondrial lineage in Yemen. *Genetics* **150**: 835–861.
- Presgraves, D.C. and Orr, H.A. 1998. Haldane's rule in taxa lacking a hemizygous X. *Science* **282**: 952–954.
- Raufaste, N., Orth, A., Belkhir, K., Senet, D., Smadja, C., Baird, S.J.E., Bonhomme, F., Dod, B., and Boursot, P. 2005. Inferences of selection and migration in the Danish house mouse hybrid zone. *Biol. J. Linn. Soc. Lond.* **84**: 593–616.
- Rieseberg, L.H. 1991. Homoploid reticulate evolution in *Helianthus* (Asteraceae): Evidence from ribosomal genes. *Am. J. Bot.* **78**: 1218–1237.
- Rieseberg, L.H. 1997. Hybrid origin of plant species. *Annu. Rev. Ecol. Syst.* **28**: 359–389.
- Rieseberg, L.H., Whitton, J., and Gardner, K. 1999. Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. *Genetics* **152**: 713–727.
- Saetre, G.P., Borge, T., Lindroos, K., Haavbie, J., Sheldon, B.C., Primmer, C., and Syvänen, A.C. 2003. Sex chromosome evolution and speciation in flycatchers. *Proc. R. Soc. Lond. B. Biol. Sci.* **270**: 53–59.
- Sage, R.D., Heyneman, D., Lim, K.C., and Wilson, A.C. 1986a. Wormy mice in a hybrid zone. *Nature* **324**: 60–63.
- Sage, R.D., Whitney III, J.B., and Wilson, A.C. 1986b. Genetic analysis of a hybrid zone between *domesticus* and *musculus* mice (*Mus musculus* complex) hemoglobin polymorphisms. *Curr. Top. Microbiol. Immunol.* **127**: 75–85.
- Schnell, G.D. and Selander, R.K. 1981. Environmental and morphological correlates of genetic variation in mammals. In *Mammalian population genetics* (eds. J. Joule and M.H. Smith), pp. 60–69. University of Georgia Press, Athens, GA.
- She, J.X., Bonhomme, F., Boursot, P., Thaler, L., and Catzeflis, F. 1990. Molecular phylogenies in the genus *Mus*: Comparative analysis of electrophoretic, scnDNA hybridization, and mtDNA RFLP data. *Biol. J. Linn. Soc. Lond.* **41**: 83–103.
- Shifman, S., Bell, J.T., Copley, R.R., Taylor, M.S., Williams, R.W., Mott, R., and Flint, J. 2006. A high-resolution single nucleotide polymorphism genetic map of the mouse genome. *PLoS Biol.*

- 4: e395. doi: 10.1371/journal.pbio.0040395.
- Smadja, C. and Ganem, G. 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. and Ganem, G. 2005. Asymmetrical reproductive character displacement in the house mouse. *J. Evol. Biol.* **18**: 1485–1493.
- Smadja, C., Catalan, J., and Ganem, G. 2004. Strong premating divergence in a unimodal hybrid zone between two subspecies of the house mouse. *J. Evol. Biol.* **17**: 165–176.
- Storchova, R., Gregorova, S., Buckiova, D., Kyselova, V., Divina, P., and Forejt, J. 2004. Genetic analysis of X-linked hybrid sterility in the house mouse. *Mamm. Genome* **15**: 515–524.
- Storey, J.D. and Tibshirani, R. 2003. Statistical significance for genome-wide studies. *Proc. Natl. Acad. Sci.* **100**: 9440–9445.
- Su, A.I., Cooke, M.P., Ching, K.A., Hakak, Y., Walker, J.R., Wiltshire, T., Orth, A.P., Vega, R.G., Sapinosa, L.M., Moqrich, A., et al. 2002. Large-scale analysis of the human and mouse transcriptomes. *Proc. Natl. Acad. Sci.* **99**: 4465–4470.
- Suzuki, H., Shimada, T., Terashima, M., Tsuchiya, K., and Aplin, K. 2004. Temporal, spatial and ecological modes of evolution of Eurasian *Mus* based on mitochondrial and nuclear gene sequences. *Mol. Phylogenet. Evol.* **33**: 626–646.
- Thomas, P.D., Campbell, M.J., Kejariwal, A., Mi, H., Kariak, B., Daverman, R., Diemer, K., Muruganujan, A., and Narechania, A. 2003. PANTHER: A library of protein families and subfamilies indexed by function. *Genome Res.* **13**: 2129–2141.
- Trachtulec, Z., Mihola, O., Vlcek, C., Himmelbauer, H., Paces, V., and Forejt, J. 2005. Positional cloning of the Hybrid sterility 1 gene: Fine genetic mapping and evaluation of two candidate genes. *Biol. J. Linn. Soc. Lond.* **84**: 637–641.
- Tucker, P.K., Sage, R.D., Warner, J.H., Wilson, A.C., and Eicher, E.M. 1992. Abrupt cline for sex chromosomes in a hybrid zone between two species of mice. *Evolution Int. J. Org. Evolution* **46**: 1146–1163.
- Turelli, M. and Begun, D.J. 1997. Haldane's rule and X-chromosome size in *Drosophila*. *Genetics* **147**: 1799–1815.
- Turelli, M. and Orr, H.A. 1995. The dominance theory of Haldane's rule. *Genetics* **140**: 389–402.
- Ursin, E. 1952. Occurrence of voles, mice, and rats (Muridae) in Denmark, with a special note on a zone of intergradation between two subspecies of the house mouse (*Mus musculus* L.). *Vid. Medd. Dansk Naturhist. Foren.* **114**: 217–244.
- van Zegeren, K. and van Oortmerssen, G.A. 1981. Frontier disputes between the West- and East-European house mouse in Schleswig-Holstein, West Germany. *Z. Saugetierkd.* **46**: 363–369.
- Vanlerberghe, F., Dod, B., Boursot, P., Bellis, M., and Bonhomme, F. 1986. Absence of Y-chromosome introgression across the hybrid zone between *Mus musculus domesticus* and *Mus musculus musculus*. *Genet. Res.* **48**: 191–197.
- Vanlerberghe, F., Boursot, P., and Catalan, J. 1988a. Analyse genetique de la zone d'hybridation entre les deux sous-espèces de souris *Mus musculus domesticus* et *Mus musculus musculus* en Bulgaire. *Genome* **30**: 427–437.
- Vanlerberghe, F., Boursot, P., Nielsen, J.T., and Bonhomme, F. 1988b. A steep cline for mitochondrial DNA in Danish mice. *Genet. Res.* **52**: 185–193.
- Vyskocilova, M., Trachtulec, Z., Forejt, J., and Piálek, J. 2005. Does geography matter in hybrid sterility in mice? *Biol. J. Linn. Soc. Lond.* **84**: 663–674.
- Weir, B.S., Hill, W.G., and Cardon, L.R. 2004. Allelic association patterns for a dense SNP map. *Genet. Epidemiol.* **27**: 442–450.

Received June 1, 2007; accepted in revised form September 23, 2007.