TITLE: Genome-wide patterns of differentiation among house mouse subspecies

SUBMISSION TYPE: Investigation

AUTHORS: Megan Phifer-Rixey*, Matthew Bomhoff§, and Michael Nachman*

AUTHOR INSTITUTIONS AND AFFILIATIONS:
* Museum of Vertebrate Zoology, University of California, Berkeley, CA, USA
§ BIO5 Institute, University of Arizona, Tucson, AZ, USA

RUNNING TITLE: Genomic differentiation among subspecies of house mouse

KEYWORDS: Mus musculus, demography, speciation, selection, reproductive isolation, expression

CORRESPONDING AUTHOR: Megan Phifer-Rixey
The Museum of Vertebrate Zoology, University of California 3101 Valley Life Sciences Building Berkeley, CA 94720-3160. 510 642-3567 (phone), 510 643-8238 (fax), mrixey@berkeley.edu
ABSTRACT
One approach to understanding the genetic basis of speciation is to scan the genomes of recently diverged taxa to identify highly differentiated regions. The house mouse, *Mus musculus*, provides a useful system for the study of speciation. Three subspecies (*M. m. castaneus*, *M. m. domesticus*, and *M. m. musculus*) diverged ~350 KYA, are distributed parapatrically, show varying degrees of reproductive isolation in laboratory crosses, and hybridize in nature. We sequenced the testes transcriptomes of multiple wild-derived inbred lines from each subspecies to identify highly differentiated regions of the genome, to identify genes showing high expression divergence, and to compare patterns of differentiation among subspecies that have different demographic histories and exhibit different levels of reproductive isolation. Using a sliding-window approach, we found many genomic regions with high levels of sequence differentiation in each of the pairwise comparisons among subspecies. In all comparisons, the X chromosome was more highly differentiated than the autosomes. Sequence differentiation and expression divergence were greater in the *M. m. domesticus - M. m. musculus* comparison than in either pairwise comparison with *M. m. castaneus*, consistent with laboratory crosses which show greatest reproductive isolation between *M. m. domesticus* and *M. m. musculus*. Coalescent simulations suggest that differences in estimates of effective population size can account for many of the observed patterns. However, there was an excess of highly differentiated regions relative to simulated distributions under a wide range of demographic scenarios. Overlap of some highly differentiated regions with previous results from QTL mapping and hybrid zone studies points to promising candidate regions for reproductive isolation.
INTRODUCTION

Understanding the genetic basis of speciation is a fundamental goal of evolutionary biology. This problem has primarily been approached in two ways: through laboratory studies using crosses and through studies of genetic variation in natural populations. Laboratory studies control for genetic background and environment, and they make it possible to connect genotype and phenotype. These types of studies have produced some spectacular successes including the identification of individual genes underlying postzygotic isolation in Drosophila (e.g. Ting et al. 1998; Presgraves et al. 2003; Masly et al. 2006; Brideau et al. 2006), Arabidopsis (Bomblies et al. 2007; Bikard et al. 2009), Mus (Mihola et al. 2009) and others (reviewed in Presgraves 2010, Nosil and Schluter 2011).

Studies of natural populations rely on the idea that regions of the genome that are important in reproductive isolation may be more differentiated than other regions of the genome. Therefore, by studying patterns of differentiation, one may gain insight into the genomic regions that underlie isolation. The idea that the genomes of closely related species are mosaics of differentiated and less differentiated regions is not new and first emerged in the literature on hybrid zones (e.g., Key 1968; Harrison 1986; Tucker et al. 1992; Reiseberg et al. 1999; reviewed in Harrison 2012). The advent of genomic methods has fueled a renewed interest in studying patterns of differentiation between closely related species, including work on mosquitoes (Turner et al. 2005; Lawiczak et al. 2010; Neafsey et al. 2010), mice (Harr 2006), Drosophila (Kulathinal et al. 2009), Heliconius butterflies (Nadeau et al. 2013), flycatchers (Ellegren et al. 2012), crickets (Andrés et al. 2013), sunflowers (Renaut et al. 2013), and others.

Despite their appeal, genome scans present a number of challenges. One is correctly identifying genomic regions that show unexpectedly high levels of differentiation. This has typically been done either by specifying an appropriate null demographic model against which an observed distribution can be compared or by simply identifying extreme values as potential candidate regions. Another challenge is interpreting the biological meaning of a genomic region showing a high level of differentiation. Shared polymorphism can result from retained ancestral variation or from gene flow; conversely, differentiation can result from sorted ancestral variation (due to drift or selection) or from absence of gene flow. Charlesworth (1998) pointed out that reduced variation within a population will inflate estimates of differentiation, such as $F_{st}$, that are based on both within- and between-population components of variation. As a result, background selection (Charlesworth et al. 1993) and genetic hitchhiking (Maynard Smith and Haigh 1974) may lead to localized high values of $F_{st}$ even for regions that are not involved in reproductive isolation (Cruickshank and Hahn 2014). Therefore, genomic “islands of differentiation” may reflect (1) stochastic variation in lineage sorting, (2) regions of reduced gene flow, (3) regions in which the effects of selection at linked sites are more pronounced, regardless of involvement with reproductive isolation, or (4) some combination of these processes.

House mice provide a valuable system for the study of speciation. *Mus musculus* consists of three subspecies which are distributed parapatrically: *M. m. domesticus* in Western
Europe, *M. m. musculus* in Eastern Europe and Northern Asia, and *M. m. castaneus* in Southeast Asia. These subspecies are believed to have diverged in allopatry at roughly the same time ~350,000 years ago (Bonhomme et al. 2007; Geraldes et al. 2008, 2011; White et al. 2009) and come into secondary contact much more recently (e.g. Cucchi 2005; Duvaux et al. 2011). Each subspecies meets and hybridizes with the other two species where their ranges come into contact (e.g. Tucker et al. 1992; Boursot et al. 1993; Duvaux et al. 2011), though only the hybrid zone between *M. m. musculus* and *M. m. domesticus* is well-studied. Differences in estimated effective population size ($N_e$) provide an opportunity to investigate the effects of demography on patterns of differentiation. While estimates of effective population size ($N_e$) are large for *M. m. castaneus* (200,000-733,000), estimates are smaller for *M. m. domesticus* (58,000-200,000) and *M. m. musculus* (25,000-120,000; Salcedo et al. 2007; Geraldes et al. 2008, 2011; Halligan et al. 2010).

The degree of reproductive isolation differs in pairwise comparisons among house mouse subspecies. While significant reductions of F1 male fertility are seen in crosses between *M. m. domesticus* and *M. m. musculus* (e.g. Good et al. 2008; White et al. 2011), significant infertility is not observed until the F2 in crosses between *M. m. castaneus* and *M. m. domesticus*, and the degree of infertility is not as severe (White et al. 2012). There are no published studies documenting reduced fertility in lab crosses between *M. m. castaneus* and *M. m. musculus*. In fact, a cross between *M. m. castaneus* and *M. m. musculus* was used for a recombination mapping study and infertility was not observed (Dumont and Payseur 2011).

In this study, we used short-read sequencing of the testis transcriptomes of wild derived inbred lines of *M. m. castaneus*, *M. m. domesticus*, and *M. m. musculus* to characterize genome-wide patterns of sequence differentiation in pairwise comparisons between each subspecies. Although this study was primarily designed to investigate sequence differentiation, we also investigated patterns of differential gene expression among the subspecies.

MATERIALS AND METHODS

Samples. All mice came from wild-derived inbred strains (Supporting Information Table 1). We sequenced eight lines of *M. m. castaneus*, seven of *M. m. domesticus* and eight of *M. m. musculus*. Wild-derived laboratory strains of *Mus spretus* and *Mus caroli* were included for use as outgroups (She et al. 1990; Suzuki et al. 2004; Tucker, Sandstedt, and Lundrigan 2005). Mice were euthanized and testes were dissected under RNAse free conditions. Testes samples were kindly provided by François Bonhomme, Polly Campbell, Courtney Clayton, Matt Dean and Annie Orth. Testes were placed in RNA later at 4°C overnight and then transferred to -80°C for storage. RNA was extracted from frozen tissue using Quiagen’s RNAeasy Plus Mini Kit.

Sequencing. Single-end 76 base pair reads were sequenced from the mRNA of each individual on an Illumina GAIIX. For most lines, between 0.80 and 1.68 Gb of sequence was obtained (Supporting Information Table 2). One wild derived inbred line of *Mus musculus* and the two wild derived lines selected from outgroup taxa were sequenced at
higher coverage (1.92 to 3.64 Gb). Reads containing fewer than 20 high quality bases (phred $\geq 20$) were removed prior to mapping. Sequence data can be accessed via NCBI BioProject PRJNA252743. TopHatv1.2 (Trapnell et al. 2009) was used with default settings to map reads to the C57BL/6 reference genome and only reads that mapped uniquely were retained. Finally, sites with a depth less than 6X of high quality sequence (phred $\geq 20$) were removed from the analysis. These filters left between 12.01 and 22.70 Mb of sequence per line. Genomic sequence data (>20X) were available from the Wellcome Trust Mouse Genomes Project for two of the lines included in our study (SPRET/EiJ and PWK/PhJ; Yalcin et al. 2012), and genomic sequence data was used to augment transcriptome sequencing. We compared genotype calls between our data and the Wellcome Trust data in regions of overlap (Supporting Information Table 3). Although both datasets were obtained via shotgun sequencing, the higher coverage of the Wellcome Trust data allows for an assessment of the possible risks of sequencing and mapping using our lower coverage approach. Mismatches were rare, occurring at rates ranging from 1 in $\sim$325,000 to 1 in $\sim$420,000 coding sites (Supporting Information Table 3). In addition, we included data from two lines (CAST/EiJ and WSB/Eiij) sequenced only by the Wellcome Trust Mouse Genomes Project (Yalcin et al. 2012).

Previous analyses have shown that wild-derived inbred lines can contain large introgressed segments from other subspecies (Yang et al. 2011). STRUCTURE analysis (Pritchard et al. 2000) was used to test for admixture in the wild-derived inbred lines sequenced in this study. We found that two lines of $M. m. castaneus$ and one line of $M. m. musculus$ were highly admixed, and we excluded them from the study (Supporting Information Table 1). After excluding these lines, the remaining subspecies formed three distinct groups corresponding to the three subspecies, and each line was assigned with most support to the expected subspecies (Supporting Information Table 4). After removing the admixed lines and including the lines sequenced by the Wellcome Trust, six $M. m. castaneus$, eight $M. m. domesticus$, and seven $M. m. musculus$ were used in all analyses.

SAMtools was used with default settings to call bases and all SNPs within and among subspecies were identified using custom PERL scripts (Li et al. 2009; Supporting File 1). Inbred lines are expected to be homozygous. Observed heterozygosity may reflect true residual heterozygosity in inbred lines or errors in sequencing; distinguishing between these two possibilities is difficult with low-coverage data. When heterozygosity was inferred using SAMtools, the site was masked and not included in further analyses. In addition, indels and sites with more than two segregating alleles were excluded. After filtering, we identified $>32,000$ SNPs within and among subspecies of $M. m. musculus$ from 4,705 genes with an average of 6.12 SNPs per gene. This represents $\sim$20% of the genes in the genome.

Measures of sequence differentiation. There are many statistics for measuring differentiation, and these capture different aspects of the data. Here, we calculated $F_{st}$ (Hudson et al. 1992), $D_{xy}$ (Nei 1987), and $\delta$, the absolute value of the difference in allele frequencies (see Renaut et al. 2010; Gagnaire et al. 2012; equations given in Supporting File 2) for each SNP for each pairwise comparison: $M. m. castaneus$-$M. m. domesticus$
(hereafter CD), \textit{M. m. castaneus-M. m. musculus} (hereafter CM), and \textit{M. m. domesticus-M. m. musculus} (hereafter DM). To account for unequal sample sizes of the subspecies, we sub-sampled five lines per subspecies 10,000 times at each site with sufficient data, and the average value of a given statistic was used for all subsequent analyses. Measures of differentiation were highly correlated in our dataset (Table 1), thus we chose to use $\delta$ for subsequent analyses, although similar results are obtained using other measures. We defined SNPs as highly differentiated if average re-sampled values of $\delta$ per site were $\geq 0.8$. In such cases, the two subspecies are one allele or less from fixation of alternate nucleotides. We defined highly differentiated as $\delta \geq 0.8$ rather than using an approach based on the distributions of statistics (for example, the upper 5% of values) because it allowed us to compare among the three pairwise analyses. We then asked how many sites were fixed in a single subspecies but polymorphic in both of the other two subspecies. Finally, we identified all fixed, derived sites in each subspecies using comparisons to the outgroup taxa, \textit{M. spretus} and \textit{M. caroli}.

\textit{Sliding windows}. To identify genomic regions with groups of sites that are highly differentiated, we performed two kinds of sliding window analyses. First, sliding window analyses of $\delta$ over 100 kb windows with a 25 kb step size were used to identify regions of the genome that were highly differentiated among subspecies. We defined regions as highly differentiated if average values of $\delta$ were $\geq 0.8$ across all sites in a window. All SNPs were included in the analyses, and windows were only evaluated when there were three or more SNPs in the window.

Private polymorphisms (i.e. those segregating in just one species) can lower the average level of differentiation in a region as measured by $F_{st}$, $D_{xy}$, and $\delta$. The presence of private polymorphisms does not mean that such regions are not potentially relevant to speciation, only that they are less likely, on average, to have experienced recent coalescent events. Analyses that do not distinguish between shared and private polymorphisms may fail to identify many regions that are fully sorted. To address this problem, we tracked the ratio of fixed differences to shared polymorphisms plus fixed differences over 100 kb windows with a 25 kb step size. This ratio can take on values between zero and one and is defined for all regions that contain an informative site. High values indicate reciprocally monophyletic gene genealogies while low values indicate populations that harbor ancestral polymorphism or are experiencing gene flow (e.g. Carneiro \textit{et al.} 2013). For this window analysis, we only included windows with at least three topologically informative SNPs.

In both sliding window approaches, regions were delimited by joining overlapping or adjacent windows with the same classification, and overall levels of diversity and differentiation were estimated by averaging across all SNPs in a delimited region. The average number of SNPs in these regions is given in Table 2. We also adopted a third approach to defining regions of differentiation by delimiting runs of fixed differences uninterrupted by shared polymorphisms. The results of these analyses were very similar to the two sliding window analyses and are given in Supporting Information File 2, Supporting Information Figure 1, and Supporting Information Tables 5 and 6.
Demographic simulations. We used coalescent simulations (Hudson 2002) to compare observed patterns of differentiation to those expected under different demographic scenarios (Supporting Information Table 7). Parameters in these models included divergence time, current and ancestral \( N_e \), and migration rates in each direction, and were based on maximum likelihood estimates obtained using IM (Nielsen and Wakeley 2001) in a previous study (Geraldes et al. 2011). We assumed no recombination within loci and free recombination among loci. This assumption is reasonable given that LD decays over distances of 10-50 kb in house mice (Laurie et al. 2007). For each pairwise split, we simulated 100,000 gene genealogies given five chromosomes from each subspecies and assumed a scaled \( \theta \) value of 1.33 based on estimates of the mutation rate (4 x 10^{-9}; Waterston et al. 2002), ancestral population size, and the approximate average number of sites surveyed per locus. Because ms (Hudson 2002) simulates individual loci, we then compared the distribution of \( \delta \) from the simulations to the observed measures across individual genes in our dataset. We also explored a wider range of demographic parameters to better match simulated distributions to observed values (see Results).

Recombination and inversions. Regions of low recombination are expected to be more highly differentiated than other regions of the genome (Noor et al. 2001; Reiseberg 2001; Nachman and Payseur 2012). For example, a recent study in sunflowers showed that regions of greater differentiation were strongly associated with reduced recombination (Renaut et al. 2013). We used the revised genetic map to estimate the recombination rate for 5 Mb intervals of the mouse genome (Shifman et al. 2006; Cox et al. 2009). We defined low recombination regions as intervals with recombination rates falling in the bottom 10% of the genome and high recombination regions as intervals falling in the top 10% of the genome. We then asked whether levels of differentiation differed between regions of low and high recombination. One limitation of this approach is that the genetic map derives from \( M. m. \) domesticus and there is some evidence that recombination rate varies among subspecies (Dumont and Payseur 2011, Dumont et al. 2011). This likely limits the power to detect differences if they exist. To compare these results with those from a previous study (Geraldes et al. 2011), we repeated the analyses estimating recombination rates over 10 Mb intervals. We also repeated the analyses defining high and low recombination regions as those falling in the upper or lower 5%, 15% and 20% of the distribution of recombination rate.

Inversions may suppress recombination. Inversion data for \( M. m. \) castaneus and \( M. m. \) musculus relative to the reference mouse genome (C57BL/6) are available from the Wellcome Trust (Yalcin et al. 2012). The reference strain used for determining structural variants was C57BL/6. This classical inbred strain is primarily of \( M. m. \) domesticus origin but contains small introgressions from other subspecies. We used the Mouse Phylogeny Viewer (Wang et al. 2012) to eliminate regions not of \( M. m. \) domesticus origin from the inversion data for \( M. m. \) castaneus and \( M. m. \) musculus. The location of inversions between \( M. m. \) castaneus and \( M. m. \) musculus was determined by identifying and removing inversions in both lines that overlap and therefore represented inversions relative to \( M. m. \) domesticus. Many inversions were identified between each pair of subspecies, but most were small (CD: \( \bar{x} = 1762 \), range = 99-19,005, n=398; CM: \( \bar{x} = 1907 \), range=63-19,752, n=620; DM: \( \bar{x} = 1749 \), range=63-23,239, n=479). Therefore,
variant SNPs in inversions were rare in our data set. However, many runs of fixed
differences spanned inversions. To investigate the relationship between inversions and
differentiation, we asked whether runs of fixed differences were more likely to overlap
with inversions than expected by chance. To determine the expected overlap, we
randomly generated the same number of regions across the genome sampled with
replacement from the same size distribution as the runs data and determined the overlap
with inversions. We did this 10,000 times and determined the percentile rank of the
observed data.

Gene expression differences. The primary motivation for generating transcriptome data
was to provide a set of common loci at which patterns of sequence differentiation could
be analyzed. Nonetheless, these data also provide an opportunity to study gene
expression differences and to compare expression divergence with sequence
differentiation.

All mice were unmated, reproductively mature males, but they differed in age. In addition,
*M. m. domesticus* individuals were reared in one facility while *M. m. castaneus* and *M. m.
*musculus* individuals were reared in another. We used two approaches to assess whether
differences in rearing conditions might bias expression analysis. First, we calculated the
average count of transcripts mapped for each gene in each subspecies correcting for
differences in sequencing effort. Pairwise comparisons between subspecies showed that
expression patterns were highly correlated (Pearson's correlation, $r_{CP}>0.99$, df=15,123,
$P<10^{-15}$; $r_{CM}>0.99$, df=15,123, $P<10^{-15}$; $r_{DM}>0.99$, df=15,123, $P<10^{-15}$). Second, we
compared our results to those of another study on gene expression in *M. m. domesticus*
and *M. m. musculus* (Nachman, unpublished data). In that study, testis transcriptomes
were sequenced for three individuals of one inbred line of *M. m. domesticus* (LEWES)
and three individuals of one inbred line of *M. m. musculus* (PWK). All mice were
unmated, reproductively mature males of the same age and all were housed in the same
room of a single animal care facility. Average mean counts of transcripts mapped per
gene for each subspecies after normalization were highly correlated in the two data sets
($r_{DOM Base Mean}=0.98$, df=11,671, $P<10^{-15}$; $r_{MUS Base Mean}=0.98$, df=11,671, $P<10^{-15}$).
In addition, the log$_2$ fold change in expression for each gene between the two subspecies
was significantly correlated between the two studies ($\eta_{log_2 fold change}=0.71$, df=11,671, $P<10^{-15}$). Although the power of the two studies differ due to design, the majority of genes
(~80%) identified in the smaller study as having significantly different expression
between the two subspecies after correction for multiple testing ($\alpha=0.01$) were identified
as significantly differently expressed in this study after correction for multiple testing
given a less conservative cut-off ($\alpha=0.05$). These analyses suggest that expression
patterns in this study were not strongly biased by differences in rearing conditions.

We identified genes that were differentially expressed in each pairwise comparison of
subspecies. Given results from TopHat (see above), HTseq (Anders et al. 2014) was used
to create tables of counts of reads mapped for all represented genes. All genes with an
average read count of $\leq 10$ in more than one subspecies were removed from the analysis.
The DESeq package in R (Anders and Huber 2010) was used to further filter the data and
identify genes with significant differential expression using a binomial test. We first
normalized counts to account for differences in sequencing effort among individuals. We then filtered out the bottom 20% of the data based on sums of the counts across all subspecies. This left 12,098 genes with sufficient data for analysis. We estimated dispersions for each subspecies and then used a binomial test to identify differentially expressed genes between each pair of subspecies. P-values were adjusted via a Benjamini-Hochberg correction with a FDR=0.01 (Benjamini and Hochberg 1995). Sites were filtered in each pairwise test if the average normalized read count was ≤10 across both subspecies.

We estimated the correlation between measures of sequence variation on a gene-by-gene basis and the absolute value of the log2-fold change in expression for each pair of subspecies. Sequence divergence and expression divergence might be correlated, particularly if differences in expression are due to sequence changes at or near the gene itself (i.e. cis-regulatory changes). In another study, patterns of allele specific expression in the testes of F1 hybrids of *M. m. domesticus* and *M. m. musculus* were used to identify genes in which differences in expression between the two subspecies were due to cis regulatory changes (Nachman, unpublished data). We used those data and repeated the correlation analysis in the DM comparison restricting it to genes identified as having cis regulatory changes. Similar data were not available for the other two pairwise comparisons.

Testis specific expression. Genes involved in reproduction are known to evolve quickly (e.g. Begun 2000; Wyckoff, Wang and Wu 2000; Good and Nachman 2005) and genes that are tissue-specific have higher rates of evolution than others in mammals (Duret and Mourchiroud 2000). We asked whether regions containing testis-specific genes were more highly differentiated than other regions in the window analysis based on all SNPs. Genes with testis-specific expression were identified using data from Su et al. (2004) available via BioGPS (Wu et al. 2009). Expression data were reduced to those tissues with support for independent expression, and expression values were averaged over the available measurements for a given tissue (Winter et al. 2004). Genes were considered testis-specific if the proportion of total expression in the testis compared to overall expression was ≥ 0.1 (Winter et al. 2004; Supporting File 3). Measures of differentiation for all regions containing testis-specific genes were then compared to measures for all other regions using 1-sided t-tests.

Identifying candidate regions for reproductive isolation. One approach to identifying candidate genes for reproductive incompatibilities is to look for overlap between the results of genomic scans and other methods such as QTL mapping studies and cline analyses in hybrid zones. There are many reasons to expect that the results of such studies will not overlap: QTL analyses focus on specific traits, hybrid zone data may track more recent processes, and genomic scans of differentiation will identify many regions that do not contribute to reproductive isolation. Nevertheless, intervals that are identified consistently across different methods are good candidates for additional study.

There are no published QTL mapping data of traits relevant to reproductive isolation for crosses between *M. m. castaneus* and *M. m. musculus* nor are there any detailed hybrid
zone studies between these taxa. However, there are published QTL mapping results for the other two pairwise comparisons (White et al. 2011, 2012) and many studies of the hybrid zone between *M. m. domesticus* and *M. m. musculus* (e.g. Vanlerberghe et al. 1986, 1988; Tucker et al. 1992; Prager et al. 1993; Munclinger et al. 2002; Macholan et al. 2007; Teeter et al. 2008, 2010; Janoušek et al. 2012). For QTL, we identified overlap between 1.5-LOD intervals associated with sterility phenotypes and highly differentiated regions in the window analysis based on all SNPs (White et al. 2011, 2012). We combined QTL into a single region for analysis if the QTL 1.5-LOD intervals were overlapping. For comparison to patterns in the *musculus*-*domesticus* hybrid zone, we used the study by Janoušek et al. (2012) in which candidate Bateson-Dobzhansky-Muller incompatibility (BDMI) loci were identified from patterns of introgression and epistasis in two different transects. We identified overlap between a 2 MB window centered on the candidate BDMI loci of Janoušek et al. (2012) and highly differentiated regions in the window analysis based on all SNPs. For both types of comparisons, we compared the observed overlap to a distribution created using 10,000 simulated datasets of genomic regions from the same size distribution as those identified as highly differentiated in our study. We repeated all analyses identifying overlap between genes that were differentially expressed in our data and the results of previous studies. For these analyses, we compared the observed overlap to a distribution created using 10,000 simulated datasets. Simulations sampled with replacement from among those genes for which there were expression data.

We used the coordinates of all SNPs in the Ensembl mouse genome assembly GRCm38 to identify genes found in overlapping regions and we used the Mouse Genome Database to identify phenotypes in laboratory lines associated with mutations in these genes (Eppig et al. 2012). When highly differentiated regions were flanked by regions with fewer than three SNPs, we expanded the query regions to the next region with data or 2MB whichever was smaller. We did not include regions from the X chromosome as the QTL associated with male infertility in both crosses encompassed most of the chromosome.

**RESULTS**

*Measures of sequence differentiation.* Over 20,000 SNPs were segregating in each pairwise comparison, but many were private, segregating at low to moderate frequency within a single subspecies in a given comparison (Table 1). Different measures of differentiation (*F_{st}, D_{xy},* and *δ* ) were highly correlated (Table 1). Average levels of differentiation varied among pairwise comparisons, and all measures of differentiation were consistently higher in DM than in either of the other two pairwise comparisons. In addition, all measures of differentiation were higher on the X chromosome than on the autosomes (Table 1). Non-synonymous sites showed slightly lower levels of differentiation than synonymous sites in each pairwise comparison (Supporting Information Table 8).

There were 9,529 individual SNPs that were highly differentiated (*δ* ≥0.80) in at least one of the three pairwise comparisons (CD: 4,223 from 1,970 genes; CM: 5,338 from 2,343 genes; DM: 7,423 from 2,783 genes). Fewer SNPs were fixed in *M. m. castaneus* but segregating in both of the other lines (744 from 437 genes) than in either *M. m.*
domesticus (1,084 from 651 genes) or M. m. musculus (1,396 from 881 genes). In addition, many fewer sites represented derived states relative to the other subspecies and both outgroups in M. m. castaneus (372 from 263 genes) than in either M. m. domesticus (770 from 554 genes) or M. m. musculus (1,570 from 1099 genes).

Sliding window analyses. The first sliding window approach (based on windows with average δ ≥0.80) included ~385-403 Mb in each pairwise comparison after filtering (~15% of the genome). We identified many windows with an average value of δ ≥0.80 in each pairwise comparison between subspecies (Figure 1; Table 2). Highly differentiated regions were characterized by higher measures of $F_{st}$ and $D_{xy}$ and lower measures of within subspecies variation than other regions (Table 2). Strikingly, regions of high differentiation represent a much larger part of the total surveyed transcriptome in the DM comparison than in either of the other two comparisons (Table 2). In each pairwise comparison, over 65% of genes sampled in highly differentiated regions contained at least one fixed difference. Approximately half of those genes contained at least one non-synonymous fixed difference (CD 57.4% of genes; CM 46.6% of genes; DM 42.4% of genes). Although this implies that approximately half of these genes have no non-synonymous fixed differences, it is important to bear in mind that coverage was incomplete for many genes and thus some non-synonymous changes may be missed.

The second sliding window approach (based on the ratio of fixed differences to shared polymorphisms plus fixed differences), after filtering, included ~89-146 Mb in each pairwise comparison. This analysis required at least three topologically informative SNPs per window and thus covered much less of the genome than the first sliding window approach. We identified many fully sorted windows in each pairwise comparison (Table 3; Supporting Information Figures 2a, b, c). Even when including private polymorphisms, autosomal regions that were fully sorted were characterized by higher measures of $F_{st}$, $D_{xy}$, and δ and lower measures of within subspecies variation than other regions (Table 3). Notably, all regions on the X chromosome were fully sorted in all pairwise comparisons. Fully sorted regions represent a much larger part of the total surveyed transcriptome in the DM comparison than in either of the other two comparisons (Table 3). As expected, a much higher proportion of the surveyed genome was identified as fully sorted in this analysis than was identified as highly differentiated in the analysis averaging across all variable sites in a window.

Demographic simulations. We conducted coalescent simulations based on demographic parameters estimated in a previous study (Geraldes et al. 2011; Supporting Information Table 7). The simulations predicted the greatest overall levels of differentiation in the DM comparison and the lowest overall levels in the CD comparison (Figure 2), a pattern consistent with the observed data. However, for all three pairwise comparisons, the observed distribution of δ was flatter than the simulated distribution, resulting from a larger than predicted proportion of genes with extreme values. Differences between the observed and simulated distributions were significant in all pairwise comparisons (Kolmogorov-Smirnov Test; CD: $D_{2-sided}=0.23$, $P<1 \times 10^{-10}$; CM: $D_{2-sided}=0.21$, $P<1 \times 10^{-10}$; DM: $D_{2-sided}=0.19$, $P<1 \times 10^{-10}$). This pattern was most pronounced in the CD and CM comparisons. An excess of genes with low values of differentiation is consistent
with higher than simulated rates of gene flow, whereas an excess of genes with high levels of differentiation is consistent with longer than simulated divergence times, lineage specific positive selection and/or barriers to gene flow. However, on average, the simulated loci had more variable sites than were surveyed in the observed data (Supporting Information Table 7). We repeated the simulations choosing values of current and ancestral $N_e$ and divergence time to try to match more closely the number of SNPs in the simulated and observed data (Supporting Information Table 9). Overall patterns were similar under both demographic scenarios (Supporting Information Figures 3 and 4) with more loci falling in the extremes of the distribution than expected based on the simulations.

We further explored the simulation parameter space to determine if increasing gene flow and divergence times could generate patterns similar to those observed in the data. We started with the original parameter values (Supporting Information Table 7) but with a common divergence time of 325 KYA. We then explored different values of gene flow for each pairwise comparison until the proportion of genes with low values of differentiation ($0 < \delta \leq 0.2$) in the simulations matched the proportion observed in the data. The levels of gene flow required were high, ranging from 7-15 times the values originally simulated (Supporting Information Table 10). In all cases, increasing gene flow to the level required resulted in an even larger excess of highly differentiated genes in the observed data relative to the simulations (Supporting Information Figure 4). Next, we tested whether increasing divergence times in tandem with gene flow could result in a distribution more similar to the one observed. We increased the divergence time first to 425 KYA and then to 825 KYA. Increasing divergence time had little effect on the proportion of genes falling in the extreme tails of the distribution for both the CD and CM comparison (Supporting Information Figures 5, 6). Increasing divergence times given such high levels of gene flow tended to increase the proportion of simulations for which the average value of $\delta$ was low with little effect on the proportion of highly differentiated loci. In the DM comparison, increasing divergence time did increase the proportion of the distribution that was highly differentiated to levels close to or exceeding those observed (Supporting Information Figures 5, 6). However, in both simulations with older divergence times, the proportion of simulated loci with low values of differentiation was much smaller than observed. We did not exhaustively explore the effects of uncertainty in estimates of effective population size, recombination rate, or mutation rate, any of which might affect the expected distribution of differentiation. Nonetheless, taken together, the simulations suggest that differences in demography can account for some of the observed patterns, such as increased differentiation in the DM comparison, but also that some regions of high differentiation result from either lineage specific positive selection or barriers to gene flow.

Recombination and inversions. Levels of differentiation were generally higher in regions of low recombination than in regions of high recombination, but the difference was only significant in the CM comparison. Results were similar among different measures of differentiation; we report results for $\delta$ (Table 4). Repeating the analysis with 10 Mb windows or with different cutoff values for high and low recombination regions yielded
qualitatively similar results (data not shown). We found no evidence of greater than expected overlap between inversions and runs of fixed differences in the CM and CD comparisons, but we did find significant overlap in the DM comparison, with the observed overlap falling in the extreme tail of the simulated distribution (P = 0.015; Supporting Information Table 11).

**Gene expression differences.** We identified many more significantly differentially expressed genes in the DM comparison that in either of the other two comparisons (CD: 594 out of 12,078; CM: 685 out of 12,078; DM: 1,049 out of 12,081; Supporting File 4). Average δ per gene and log₂-fold change in expression were significantly positively correlated in all pairwise comparisons of subspecies, although the correlation coefficients were small (CD: $r_{\delta,\text{abs}(\text{log}_2 \text{fold change})} = 0.05$, df=2,483, $P = 0.01$; CM: $r_{\delta,\text{abs}(\text{log}_2 \text{fold change})} = 0.06$, df=2,471, $P = 0.001$; DM: $r_{\delta,\text{abs}(\text{log}_2 \text{fold change})} = 0.05$, df=2,356, $P = 0.02$). Results were very similar for other measures of differentiation (data not shown). Restricting the data to genes identified in another study as being significantly differentially transcribed due to cis-regulatory changes (Nachman, unpublished data) strengthened the correlation in the DM comparison though the significance was reduced given less power ($r_{\delta,\text{abs}(\text{log}_2 \text{fold change})} = 0.09$, df=439, $P = 0.05$).

**Testis specific expression.** Approximately 30% of regions surveyed in the analyses including all SNPs contained at least one testis-specific gene (Supporting Information Table 12). Overall, testis-specific genes were significantly more common in highly differentiated regions than in other regions ($\chi^2 = 3.74$, n=9,822, $P_{1\text{-tailed}} = 0.03$). Regions containing testis-specific genes were more differentiated than other regions in the CM and DM comparisons, but these differences were only consistently significant in the DM comparison (Supporting Information Table 12).

**Identifying candidate regions for reproductive isolation between M. m. castaneus and M. m. domesticus.** In comparisons between M. m. castaneus and M. m. domesticus, we did not observe significant overlap between differentially expressed genes and QTL associated with hybrid male infertility. However, we did observe significant overlap between highly differentiated regions and QTL (White et al. 2012; Figure 3a). Of the nine QTL intervals, seven contain peaks of high differentiation, and the observed overlap ranked in the 97th percentile of the simulated distribution.

Regions of overlap on the autosomes contained 221 protein-coding regions (Supporting Information Table 13). Across all 221 autosomal genes in regions of overlap, 20 were testis-specific (BC049635, Bps9, Catsper2, Ccdc53, Ccl27a, Ccl27b, Eif3j1, Euf1, Gm13306, Lin7a, Lrre57, Nup37, Parpfb, Psmc3, Sord, Sycp3, Tex26, Trim69, Tsc22d4, Ttk2), 11 had functional annotations and/or phenotypes associated with male fertility (Arhgap1, Bps9, Cdkn2c, Celf1, Duox2, Ehd4, Igf1, Ilra1, Nr1h3, Pmch, and Pparg), and three genes were both testis-specific and had mutational variants associated with male infertility (Catsper2, Sord, Sycp3). Nine genes in regions of overlap showed significant expression differences ($p_{\text{adj}} < 0.05$; 2700089E24Rik, Atg7, B2m, Capn3, Cdkn2c, Igf1, Nup37, Ppip5k1, Shf). Two of those, Cdkn2c and Igf1, are associated with male fertility and one, Nup37, is testis-specific.
In general QTL are large, while regions of high differentiation are relatively small, potentially helping to narrow QTL intervals. For example, one QTL on chromosome nine associated with amorphous sperm heads encompasses 26.7 Mb and contains ~220 protein coding genes (White et al. 2012). It overlaps with only one highly differentiated region that contains only one protein-coding gene, Bbs9. Bbs9 has gene ontology terms relating to cilia and is testis-specific. In humans, Bbs9 mutations are associated with Bardet-Biedl syndrome, a disease with multiple phenotypic effects including reduced testis size. Expression levels at Bbs9 were different in M. m. castaneus and M. m. domesticus, but this difference was not significant after correction for multiple testing ($p<0.025, P_{adj}=0.16$). We observed three silent and no replacement fixed differences between M. m. castaneus and M. m. domesticus at Bbs9. Not all sites were surveyed, and thus observed patterns of differentiation may reflect linkage to functionally important unsurveyed sites. It is also important to bear in mind that not all genes in QTL intervals were surveyed.

Identifying candidate regions for reproductive isolation between M. m. domesticus and M. m. musculus. In comparisons between M. m. domesticus and M. m. musculus, the overlap between highly differentiated regions or differentially expressed genes and QTL associated with male infertility (White et al. 2011) was not more than expected by chance. The overlap between differentially expressed genes and candidate BDMIs loci from the hybrid zone study of Janoušek et al. (2012) was also not more than expected by chance. However, the overlap between highly differentiated regions and the candidate BDMI loci ranks in the 92$^{nd}$ percentile of simulated data (Janoušek et al. 2012). Regions of overlap between all three kinds of studies (QTL, candidate BDMI loci from the hybrid zone, and regions of high differentiation) are particularly promising candidates for reproductive isolation. Importantly, six autosomal regions were identified in which candidate BDMIs and regions of high differentiation overlap precisely or are contiguous and fall within QTL intervals (Figure 3b). These regions collectively contain 242 genes and the number of genes found in each specific region ranges from 17-97 (Supporting Information Table 14).

Two regions fall in relatively small QTL intervals. The first is on chromosome 4. This QTL is associated with relative testis weight (White et al. 2011) and contains only 16 genes (Figure 4). Of these 16 genes, four are testis-specific (4921539E11Rik, Mier1, Tctex1d1 and Wdr78) and two (Ins15 and Dab1) are associated with male infertility. We found that three genes in this interval are differentially expressed between M. m. domesticus and M. m. musculus after correction for multiple testing ($\alpha = 0.05$; C8b, Dab1, and Prkaa2).

The second small QTL interval with precise overlap is found on chromosome 5. This QTL is associated with both total abnormal sperm and distal bent tail phenotypes. This interval contains 97 genes and is relatively well sampled in our study. There are 14 testis-specific genes in this interval (Fbxo24, Mcm7, Mepce, Muc3, Myl10, Ppp1r35, Rabl5, Srrt, Stag3, Taf6, Tmem184a, Tsc22d4, Znhit1 and Zscan21). Two of these genes (Stag3 and Zscan21) have GO annotations and/or phenotypes relating to male fertility. Myl10
and Rabl5 were differentially expressed between *M. m. domesticus* and *M. m. musculus* after correction for multiple testing ($P_{adj} < 0.05$). There are ten additional genes in the region with known functional annotations and/or phenotypes relating to male fertility (*Ache, Cnpy4, Cux1, Fam20c, Pdgfa, Smok3a, Smok3b, Sun1, Vgf, and Zan*).

**DISCUSSION**

*Genome-wide patterns of sequence differentiation.* We used a transcriptomic approach to characterize genome-wide patterns of differentiation between the three subspecies of house mice and discovered many highly differentiated regions in all pairwise comparisons. By comparing three subspecies that split from one another at approximately the same time but that have different estimated effective population sizes, we were able to study the influence of demography on the early stages of speciation and divergence. In this case, we found higher levels of sequence differentiation between *M. m. domesticus* and *M. m. musculus* than between the other pairs of subspecies. This result is consistent with estimates of the demographic history of these species; both *M. m. domesticus* and *M. m. musculus* are believed to have undergone significant bottlenecks resulting in current $N_e$ of $\approx 1/10$-$1/2$ of the ancestral $N_e$. *M. m. castaneus*, on the other hand, is estimated to have a population size very similar to the ancestral population (Geraldes *et al.* 2011). Lineage specific changes observed in this study support those expectations. The fewest lineage specific changes were assigned to *M. m. castaneus*, the subspecies with the highest $N_e$, and the most were assigned to *M. m. musculus*, the subspecies with the smallest $N_e$. More generally, the coalescent simulations performed here recapitulated the broad patterns of differentiation seen among the three subspecies, suggesting that many of the observed patterns can be explained by differences in $N_e$ and levels of gene flow (Figure 2).

At the same time, greater reproductive isolation is seen in laboratory crosses between *M. m. domesticus* and *M. m. musculus* than between *M. m. castaneus* and *M. m. domesticus* or *M. m. castaneus* and *M. m. musculus* (Dumont and Payseur 2011; White *et al.* 2011, 2012). This observation, by itself, leads to the prediction of greater differentiation in the DM comparison than in the other two comparisons. Because this pattern is also predicted by demographic differences among the subspecies, it is difficult to disentangle the relative contribution of differences in demography and differences in reproductive isolation to the observed patterns. It is also unclear whether differences in demography are the cause of the differing levels of reproductive isolation. For example, if most BDMI alleles were neutral on their own genetic background, then subspecies with smaller $N_e$ would be expected to accumulate more BDMI differences due to drift and would therefore show greater reproductive isolation. However, most BDMI genes in other systems seem to show evidence of positive selection, suggesting that drift is not the predominant process fixing alleles involved in BDMI (Coyne and Orr 2004; Presgraves 2010).

*Differentiation on the X chromosome.* The X chromosome was significantly more differentiated than the autosomes in all pairwise comparisons (Table 1). In principle, this pattern is expected for two reasons. First, the smaller estimated $N_e$ of the X chromosome should result in faster lineage sorting. Second, this pattern is consistent with the large X
effect, the disproportionate accumulation of reproductive incompatibilities on the X chromosome (e.g. Coyne and Orr 1989). In this case, the greater level of differentiation appears to be more than can be explained by differences in the X to autosome ratio of $N_e$. At migration-drift equilibrium, assuming constant bi-directional migration, a sex ratio of one, and equal migration of males and females, $F_{st} = 1/(4N_m + 1)$ for the autosomes and $F_{st} = 1/(3N_m + 1)$ for the X chromosome. If $Nm$ is approximately 0.1 (Supporting Information Table 7) then the expected X:autosome ratio of $F_{st}$ is 1.08 and the observed ratios are 1.45 (CD), 1.35 (CM) and 1.21 (DM; Table 1). This model is clearly overly simplistic. For example, there is some evidence of male biased dispersal in this system (Pocock et al. 2005). However, these rough calculations suggest that differences in $N_e$ alone cannot account for the greater differentiation seen on the X chromosome.

On the other hand, our observations are consistent with previous work suggesting a large X effect. For example, hybrid zone studies of *M. m. domesticus* and *M. m. musculus* indicate reduced introgression on the X (Tucker et al. 1992; Dod et al. 1993) and IM analysis of a limited number of loci in all three subspecies suggest that gene flow on the X chromosome is lower than that of autosomes (Geraldes et al. 2008, 2011). In laboratory crosses, hybrid male sterility phenotypes map to the X chromosome in crosses between *M. m. domesticus* and *M. m. musculus* (Storchová et al. 2004; Good et al. 2008; White et al. 2011) and *M. m. castaneus* and *M. m. domesticus* (White et al. 2012). Our findings here demonstrate that elevated differentiation on the X chromosome is a general pattern, is observed in allopatric populations, and extends to all pairs of subspecies.

**Recombination.** Several recent models suggest that chromosomal rearrangements may lead to reduced gene flow via their effect on suppressing recombination (Noor et al. 2001; Reiseberg 2001; Navarro and Barton 2003). Recombination can also influence differentiation by amplifying the effects of genetic hitchhiking and background selection within lineages (Maynard Smith and Haigh 1976; Charlesworth et al. 1993), reducing variation within lineages, and thus leading to increased differentiation between lineages. We found weak support for a negative relationship between differentiation and recombination, consistent with these models. However, the power of this approach may be limited by the absence of data on recombination rate variation across the genome in all three subspecies.

**Testis-specific expression.** Regions of high differentiation contained a significantly higher proportion of testis-specific genes than other regions. This result is unexpected if highly differentiated regions simply reflect stochastic variation in lineage sorting. In contrast, this result is consistent with (1) reduced gene flow due to BDMIs associated with testis-specific genes, (2) hitchhiking effects associated with positive selection at testis-specific genes, (3) or both. Importantly, this observation suggests that some proportion of highly differentiated regions is associated with functional differences within or between nascent species.

**Candidate regions for reproductive incompatibilities.** It would be incorrect to claim that all regions of high differentiation contribute to reproductive isolation when many such regions are expected simply as a consequence of drift in small populations. In addition,
some regions of high differentiation are likely the result of lineage specific selection that may not contribute to reproductive isolation. Nonetheless, the observed data differed from demographic simulations in one major way: the distribution of differentiation statistics was flatter, with more values in the extremes of the distribution. This is consistent with more gene flow and more differentiation than expected. Even when exploring a broad range of values for gene flow and divergence time, we were unable to find a demographic scenario that recapitulated observed patterns. Therefore, some highly differentiated regions likely result either from lineage specific positive selection and/or from barriers to gene flow at loci underlying incompatibilities.

One approach to prioritizing candidate reproductive isolation loci is to identify overlap between the results of genome scans, laboratory crosses and hybrid zone analyses. We identified several areas of overlap between QTL associated with male sterility in a cross of M. m. castaneus and M. m. domesticus (White et al. 2012) and highly differentiated regions identified in our study. The overlap was more than expected by chance, but in most cases, QTL were large, making the overlap difficult to interpret (e.g. chromosome five, Figure 3a). However, in other cases, the QTL intervals were narrower, there was reasonable coverage in our dataset, and relatively few genes were found in the overlap. For example, on chromosome 9, there is just one gene in a region of high differentiation that falls in a moderately-sized QTL. Even though more than 200 genes fall in all of the areas of overlap, this number is considerably smaller than the total number of genes that fall in QTL intervals (White et al. 2012). Moreover, less than three dozen of those genes have known phenotypes or gene ontology terms that relate to male fertility and/or are testis-specific as might be expected if they affect male sterility phenotypes measured in the QTL analyses. Of course, gene ontology annotation and documentation of phenotypes associated with mutations or knockouts in mice are far from complete and additional genes in these regions may be related to fertility.

In the DM comparison, overlap between our results and QTL analyses was considerable, but not more than expected by chance. More promisingly, there was meaningful overlap between candidate BMDIs (Janoušek et al. 2012) and our results. In particular, there were six cases in which regions identified as highly differentiated in our study were directly overlapping or contiguous with a candidate BDMI and fell in a QTL interval. In two of those cases, the overlap was relatively precise and the region of overlap contains a short list of genes that are testis-specific, differentially expressed, and/or related to male infertility. While there is still much work to be done, the intersection of results from multiple studies is encouraging and highlights the promise of this approach for narrowing QTL intervals.

ACKNOWLEDGEMENTS
This research was funded by NIH grant R01 GM074245 to M.W.N. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1053575. We thank F. Bonhomme, P. Campbell, C. Clayton, M. Dean and A. Orth for providing tissue samples and J. MacDonald for providing access to the source code for SLIDER. We also thank
members of the Nachman Lab, D. Begun, D. Matute, and one anonymous reviewer for their thoughtful comments on the manuscript.
LITERATURE CITED


Cruickshank, T.C., Hahn, M.W., 2014 Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. Mol. Ecol. In press.

Dod B., Jermiin L. S., Boursot P., Chapman V. H., Nielsen J. T., et al., 1993


Table 1. Summary statistics describing patterns of differentiation at all SNPs in pairwise comparisons of the subspecies of *Mus musculus*.

<table>
<thead>
<tr>
<th>SUBSPECIES</th>
<th>SUBSPECIES</th>
<th>CHR.</th>
<th># OF SNPS</th>
<th>$F_{st}$ (SD)</th>
<th>$D_{xy}$ (SD)</th>
<th>$\delta$ (SD)</th>
<th>$r_{F_{st},D_{xy}}$</th>
<th>$r_{F_{st},\delta}$</th>
<th>$r_{D_{xy},\delta}$</th>
<th>% FIXED DIFF.</th>
<th>% PRIV. POLY.</th>
<th>% SHARED POLY.</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>M. m.</em> castaneus</td>
<td><em>M. m.</em> domesticus</td>
<td>Autosomes</td>
<td>24,136</td>
<td>0.22 (0.32)</td>
<td>0.41 (0.28)</td>
<td>0.40 (0.29)</td>
<td>0.99</td>
<td>0.98</td>
<td>0.98</td>
<td>9.14</td>
<td>83.66</td>
<td>7.20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>X</td>
<td>226</td>
<td>0.32 (0.41)</td>
<td>0.53 (0.34)</td>
<td>0.53 (0.35)</td>
<td>0.99</td>
<td>0.99</td>
<td>1.00</td>
<td>23.89</td>
<td>73.91</td>
<td>2.21</td>
</tr>
<tr>
<td><em>M. m.</em> castaneus</td>
<td><em>M. m.</em> musculus</td>
<td>Autosomes</td>
<td>23,709</td>
<td>0.26 (0.38)</td>
<td>0.44 (0.30)</td>
<td>0.43 (0.31)</td>
<td>0.97</td>
<td>0.98</td>
<td>0.99</td>
<td>12.50</td>
<td>81.55</td>
<td>5.95</td>
</tr>
<tr>
<td></td>
<td></td>
<td>X</td>
<td>237</td>
<td>0.35 (0.39)</td>
<td>0.54 (0.33)</td>
<td>0.53 (0.33)</td>
<td>0.97</td>
<td>0.98</td>
<td>0.99</td>
<td>18.14</td>
<td>73.00</td>
<td>8.86</td>
</tr>
<tr>
<td><em>M. m.</em> domesticus</td>
<td><em>M. m.</em> musculus</td>
<td>Autosomes</td>
<td>21,598</td>
<td>0.38 (0.41)</td>
<td>0.53 (0.35)</td>
<td>0.52 (0.35)</td>
<td>0.98</td>
<td>0.99</td>
<td>0.99</td>
<td>24.01</td>
<td>70.96</td>
<td>5.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>X</td>
<td>246</td>
<td>0.46 (0.44)</td>
<td>0.57 (0.38)</td>
<td>0.57 (0.38)</td>
<td>0.99</td>
<td>0.99</td>
<td>1.00</td>
<td>30.08</td>
<td>68.29</td>
<td>1.63</td>
</tr>
</tbody>
</table>
Table 2. Summary statistics describing regions identified as highly differentiated with a sliding window analysis vs. all other regions in pairwise comparisons of the subspecies of *Mus musculus*.

<table>
<thead>
<tr>
<th>SUBSPECIES</th>
<th>SUBSPECIES</th>
<th>CHR.</th>
<th>WINDOW TYPE&lt;sup&gt;a&lt;/sup&gt;</th>
<th>n&lt;sup&gt;b&lt;/sup&gt;</th>
<th>AVG SIZE OF REGION IN BPS&lt;sup&gt;c&lt;/sup&gt; (SD)</th>
<th>AVG. # OF SNPS (SD)</th>
<th><em>F</em>&lt;sub&gt;ST&lt;/sub&gt; (SD)</th>
<th><em>D</em>&lt;sub&gt;xy&lt;/sub&gt; (SD)</th>
<th><em>δ</em> (SD)</th>
<th>π&lt;sub&gt;1&lt;/sub&gt; (SD)</th>
<th>π&lt;sub&gt;2&lt;/sub&gt; (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>M. m. castaneus</em></td>
<td><em>M. m. domesticus</em></td>
<td>Autosomes</td>
<td>highly differentiated</td>
<td>63</td>
<td>133,333&lt;sup&gt;*&lt;/sup&gt; (37,567)</td>
<td>5.13&lt;sup&gt;*&lt;/sup&gt; (2.85)</td>
<td>0.80&lt;sup&gt;*&lt;/sup&gt; (0.10)</td>
<td>0.87&lt;sup&gt;*&lt;/sup&gt; (0.06)</td>
<td>0.87&lt;sup&gt;*&lt;/sup&gt; (0.06)</td>
<td>0.11&lt;sup&gt;*&lt;/sup&gt; (0.12)</td>
<td>0.05&lt;sup&gt;*&lt;/sup&gt; (0.07)</td>
</tr>
<tr>
<td>Autosomes</td>
<td>all others</td>
<td>1651</td>
<td>235,933&lt;sup&gt;*&lt;/sup&gt; (122,354)</td>
<td>13.74&lt;sup&gt;*&lt;/sup&gt; (13.67)</td>
<td>0.21&lt;sup&gt;*&lt;/sup&gt; (0.14)</td>
<td>0.40&lt;sup&gt;*&lt;/sup&gt; (0.12)</td>
<td>0.39&lt;sup&gt;*&lt;/sup&gt; (0.13)</td>
<td>0.26&lt;sup&gt;*&lt;/sup&gt; (0.11)</td>
<td>0.16&lt;sup&gt;*&lt;/sup&gt; (0.09)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>highly differentiated</td>
<td>5</td>
<td>160,000 (33,541)</td>
<td>4.60 (2.30)</td>
<td>0.84 (0.15)</td>
<td>0.90 (0.09)</td>
<td>0.90 (0.09)</td>
<td>0.11 (0.15)</td>
<td>0.03 (0.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>all others</td>
<td>27</td>
<td>169,444 (37,553)</td>
<td>5.11 (2.87)</td>
<td>0.31 (0.23)</td>
<td>0.45 (0.20)</td>
<td>0.45 (0.20)</td>
<td>0.19 (0.17)</td>
<td>0.13 (0.11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>All</td>
<td>1746</td>
<td>230,985 (121,119)</td>
<td>13.27 (13.45)</td>
<td>0.24 (0.18)</td>
<td>0.42 (0.15)</td>
<td>0.41 (0.16)</td>
<td>0.25 (0.12)</td>
<td>0.15 (0.10)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Highly differentiated regions defined as average *δ* ≥ 0.8

<sup>b</sup> Number of delimited regions

<sup>c</sup> Regions were delimited by joining overlapping or contiguous windows with the same classification. The resolution of individual windows is limited to the sliding window increment of 25,000.

<sup>d</sup> *π<sub>1</sub>* and *π<sub>2</sub>* refer to nucleotide diversity (Nei and Li 1979) in the first and second subspecies, respectively.

<sup>*</sup> *P* < 10<sup>-10</sup> in 2-sided t-tests comparing highly differentiated autosomal regions and all other autosomal regions in each pairwise comparison.
Table 2. cont’d.

<table>
<thead>
<tr>
<th>SUBSPECIES</th>
<th>SUBSPECIES</th>
<th>CHR. TYPE</th>
<th>WINDOW TYPE&lt;sup&gt;a&lt;/sup&gt;</th>
<th>n&lt;sup&gt;b&lt;/sup&gt;</th>
<th>AVG SIZE OF REGION IN BPS&lt;sup&gt;c&lt;/sup&gt; (SD)</th>
<th>AVG. # OF SNPS (SD)</th>
<th>( \overline{F}_{st} ) (SD)</th>
<th>( \overline{D}_{xy} ) (SD)</th>
<th>( \overline{\delta} ) (SD)</th>
<th>( \pi_1 ) (SD)</th>
<th>( \pi_2 ) (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>M. m. castaneus</em></td>
<td><em>M. m. musculus</em></td>
<td>Autosomes</td>
<td>highly differentiated</td>
<td>105</td>
<td>129,762&lt;sup&gt;*&lt;/sup&gt; (35,878)</td>
<td>5.31&lt;sup&gt;*&lt;/sup&gt; (2.84)</td>
<td>0.80 (0.09)</td>
<td>0.87&lt;sup&gt;*&lt;/sup&gt; (0.06)</td>
<td>0.87 (0.06)</td>
<td>0.11 (0.10)</td>
<td>0.05 (0.07)</td>
</tr>
<tr>
<td>Autosomes</td>
<td>all others</td>
<td>1625</td>
<td>229,708&lt;sup&gt;*&lt;/sup&gt; (114,947)</td>
<td>13.55&lt;sup&gt;*&lt;/sup&gt; (13.46)</td>
<td>0.24&lt;sup&gt;*&lt;/sup&gt; (0.15)</td>
<td>0.43&lt;sup&gt;*&lt;/sup&gt; (0.12)</td>
<td>0.41&lt;sup&gt;*&lt;/sup&gt; (0.13)</td>
<td>0.27&lt;sup&gt;*&lt;/sup&gt; (0.11)</td>
<td>0.14&lt;sup&gt;*&lt;/sup&gt; (0.11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>all others</td>
<td>6</td>
<td>154,167&lt;sup&gt;*&lt;/sup&gt; (33,229)</td>
<td>3.67 (1.03)</td>
<td>0.86 (0.17)</td>
<td>0.92 (0.09)</td>
<td>0.92 (0.09)</td>
<td>0.03 (0.08)</td>
<td>0.12 (0.14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>all others</td>
<td>26</td>
<td>168,269&lt;sup&gt;*&lt;/sup&gt; (43,335)</td>
<td>5.38 (3.54)</td>
<td>0.32 (0.18)</td>
<td>0.49 (0.14)</td>
<td>0.48 (0.15)</td>
<td>0.21 (0.17)</td>
<td>0.16 (0.12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>All</td>
<td>1762</td>
<td>222,588&lt;sup&gt;*&lt;/sup&gt; (113,626)</td>
<td>12.90 (13.14)</td>
<td>0.28 (0.20)</td>
<td>0.46 (0.16)</td>
<td>0.45 (0.17)</td>
<td>0.25 (0.17)</td>
<td>0.13 (0.11)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Highly differentiated regions defined as average \( \delta \) \( \geq \) 0.8

<sup>b</sup> Number of delimited regions

<sup>c</sup> Regions were delimited by joining overlapping or contiguous windows with the same classification. The resolution of individual windows is limited to the sliding window increment of 25,000.

<sup>d</sup> \( \pi_1 \) and \( \pi_2 \) refer to nucleotide diversity (Nei and Li 1979) in the first and second subspecies, respectively.

* \( P < 10^{-10} \) in 2-sided t-tests comparing highly differentiated autosomal regions and all other autosomal regions in each pairwise comparison.
Table 2. cont’d.

<table>
<thead>
<tr>
<th>SUBSPECIES</th>
<th>SUBSPECIES</th>
<th>WINDOW TYPE&lt;sup&gt;a&lt;/sup&gt;</th>
<th>chr.</th>
<th>Window Type&lt;sup&gt;a&lt;/sup&gt;</th>
<th>n&lt;sup&gt;b&lt;/sup&gt;</th>
<th>AVG SIZE OF REGION IN BPS&lt;sup&gt;c&lt;/sup&gt; (SD)</th>
<th>AVG. # OF SNPS (SD)</th>
<th>F&lt;sub&gt;st&lt;/sub&gt; (SD)</th>
<th>D&lt;sub&gt;xy&lt;/sub&gt; (SD)</th>
<th>δ (SD)</th>
<th>π&lt;sub&gt;1&lt;/sub&gt; (SD)</th>
<th>π&lt;sub&gt;2&lt;/sub&gt; (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M. m. domesticus</td>
<td>M. m. musculus</td>
<td>Autosomes</td>
<td>Autosomes</td>
<td>highly differentiated</td>
<td>287</td>
<td>135,279*</td>
<td>5.98*</td>
<td>0.81*</td>
<td>0.87*</td>
<td>0.87*</td>
<td>0.07*</td>
<td>0.07*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Autosomes</td>
<td>Autosomes</td>
<td>all others</td>
<td>1561</td>
<td>216,944*</td>
<td>11.99*</td>
<td>0.34*</td>
<td>0.49*</td>
<td>0.48*</td>
<td>0.18*</td>
<td>0.15*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>highly differentiated</td>
<td>4</td>
<td>162,500</td>
<td>4.75*</td>
<td>0.90</td>
<td>0.93</td>
<td>0.93</td>
<td>0.04</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>all others</td>
<td>33</td>
<td>170,455</td>
<td>4.97*</td>
<td>0.38</td>
<td>0.50</td>
<td>0.50</td>
<td>0.11</td>
<td>0.15</td>
</tr>
<tr>
<td>All</td>
<td>All</td>
<td>All</td>
<td>All</td>
<td>1885</td>
<td></td>
<td>203,581</td>
<td>10.94*</td>
<td>0.41</td>
<td>0.55</td>
<td>0.54</td>
<td>0.16</td>
<td>0.14</td>
</tr>
</tbody>
</table>

<sup>a</sup> Highly differentiated regions defined as average δ ≥ 0.8

<sup>b</sup> Number of delimited regions

<sup>c</sup> Regions were delimited by joining overlapping or contiguous windows with the same classification. The resolution of individual windows is limited to the sliding window increment of 25,000.

<sup>d</sup> π<sub>1</sub> and π<sub>2</sub> refer to nucleotide diversity (Nei and Li 1979) in the first and second subspecies, respectively.

* P < 10<sup>-10</sup> in 2-sided t-tests comparing highly differentiated autosomal regions and all other autosomal regions in each pairwise comparison.
Table 3. Summary statistics describing patterns of differentiation in fully sorted regions vs. all other regions in pairwise comparisons of the subspecies of *Mus musculus*.

| SUBSPECIES       | ROLES                        | CHR. TYPE | WINDOW TYPE | n^b | AVG SIZE OF REGION IN BPS (SD) | AVG. # OF SNPS (SD) | FST (SD) | Dxy (SD) | δ (SD) | π₁ (SD) | π₂ (SD) |
|------------------|------------------------------|-----------|-------------|-----|----------|-------------------|--------|---------|--------|--------|--------|--------|
| M. m. castaneus  | M. m. domesticus             | Autosomes | fully sorted | 223 | 159,193 (38,840) | 15.05** (9.17) | 0.48 (0.16) | 0.59 (0.13) | 0.59 (0.13) | 0.15** (0.10) | 0.11*** (0.07) |
|                  |                              | Autosomes | all others | 321 | 157,243 (46,855) | 19.36** (12.56) | 0.20*** (0.13) | 0.43*** (0.10) | 0.39*** (0.11) | 0.29*** (0.09) | 0.20*** (0.09) |
|                  |                              | X        | fully sorted | 5   | 175,000 (0) | 5.20 (1.92) | 0.80 (0.12) | 0.85 (0.09) | 0.85 (0.09) | 0.06 (0.06) | 0.05 (0.04) |
| M. m. castaneus  | M. m. musculus               | Autosomes | fully sorted | 317 | 172,003 (52,573) | 16.54 (10.86) | 0.47*** (0.13) | 0.60*** (0.13) | 0.60*** (0.13) | 0.20*** (0.10) | 0.07*** (0.06) |
|                  |                              | Autosomes | all others | 287 | 154,355 (52,334) | 18.20 (12.42) | 0.19*** (0.14) | 0.42*** (0.11) | 0.38*** (0.13) | 0.28*** (0.09) | 0.21*** (0.10) |
|                  |                              | X        | fully sorted | 3   | 175,000 (0) | 3.67 (1.15) | 1.00 (0.00) | 1.00 (0.00) | 1.00 (0.00) | 0.00 (0.00) | 0.00 (0.00) |
| M. m. domesticus | M. m. musculus               | Autosomes | fully sorted | 624 | 178,766 (55,386) | 13.93 (9.50) | 0.57*** (0.17) | 0.66*** (0.13) | 0.66*** (0.13) | 0.12*** (0.07) | 0.08*** (0.06) |
|                  |                              | Autosomes | all others | 234 | 140,171 (51,005) | 14.50 (10.45) | 0.27*** (0.16) | 0.48*** (0.12) | 0.44*** (0.14) | 0.22*** (0.09) | 0.23*** (0.10) |
|                  |                              | X        | fully sorted | 10  | 16,000 (24,152) | 5.50 (1.27) | 0.70 (0.18) | 0.76 (0.15) | 0.76 (0.15) | 0.07 (0.06) | 0.05 (0.06) |

^a^Fully sorted regions are defined as those in which \# fixed differences \# fixed differences + \# shared polymorphisms = 1 (see Materials and Methods).

^b^Number of delimited regions

^c^π₁ and π₂ refer to nucleotide diversity (Nei and Li 1979) in the first and second subspecies, respectively.

*P<0.05, **P<0.01, ***P<0.001 in 2-sided t-tests comparing highly differentiated autosomal regions and all other autosomal regions in each pairwise comparison.
Table 4. Average sequence differentiation in regions of low and high recombination.

<table>
<thead>
<tr>
<th>SUBSPECIES</th>
<th>SUBSPECIES</th>
<th>$\bar{\delta}_{\text{low recombination}}$ (SD)</th>
<th>$\bar{\delta}_{\text{high recombination}}$ (SD)</th>
<th>$t$</th>
<th>$P_{t\text{-tailed}}$</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M. m. castaneus$</td>
<td>$M. m. domesticus$</td>
<td>0.29 (0.12)</td>
<td>0.28 (0.06)</td>
<td>0.22</td>
<td>0.83</td>
<td>88</td>
</tr>
<tr>
<td>$M. m. castaneus$</td>
<td>$M. m. domesticus$</td>
<td>0.36 (0.10)</td>
<td>0.31 (0.05)</td>
<td>2.72</td>
<td>&lt;0.01</td>
<td>88</td>
</tr>
<tr>
<td>$M. m. domesticus$</td>
<td>$M. m. musculus$</td>
<td>0.38 (0.13)</td>
<td>0.35 (0.07)</td>
<td>1.17</td>
<td>0.12</td>
<td>88</td>
</tr>
</tbody>
</table>
Figure 1. Sliding window analysis showing average values of $\delta$ throughout all chromosomes for each pairwise subspecies comparison (CD, *M. m. castaneus* vs. *M. m. domesticus*; CM, *M. m. castaneus* vs. *M. m. musculus*; DM, *M. m. domesticus* vs. *M. m. musculus*). Each dot marks the start of delimited region, and red dots represent regions for which the average value of $\delta$ is $\geq 0.8$. 
Figure 1. cont’d.
Figure 2. The distribution of $\delta$ values in the observed data and in the simulated data based demographic parameters from Geraldes et al. 2011.
Figure 3. Candidate regions for reproductive incompatibilities A) Overlap between autosomal regions identified as QTL associated with male sterility in a cross between *M. m. castaneus* and *M. m. domesticus* (White et al. 2012; blue bars) and regions identified as highly differentiated in our scan based on all SNPs (black dots). B) Overlap between QTL associated with male infertility (White et al. 2011; blue bars), regions identified as contributing to epistasis between *M. m. domesticus* and *M. m. musculus* in two hybrid zones in Central Europe (Janoušek et al. 2012; red dots) and regions identified as highly differentiated in our scan based all SNPs (black dots).
Figure 4. A region on chromosome four in which a QTL for relative testis weight (White et al. 2011), a candidate BDMI (Janoušek et al. 2012), and a highly differentiated region identified in this study overlap in the DM comparison. Testis specific genes are given in italics, differentially expressed genes are underlined, and genes known to be related to male fertility are given in bold.