February 2007

THE STUDY OF ADAPTATION AND SPECIATION IN THE GENOMIC ERA

Jay F. Storz
University of Nebraska - Lincoln, jstorz2@unl.edu

Hopi E. Hoekstra
University of California San Diego, La Jolla, CA

Journal of Mammalogy, 88(1):1–4, 2007

THE STUDY OF ADAPTATION AND SPECIATION IN THE GENOMIC ERA

JAY F. STORZ* AND HOPI E. HOEKSTRA

School of Biological Sciences, University of Nebraska, Lincoln, NE 68588, USA (JFS)
Division of Biological Sciences, University of California San Diego, La Jolla, CA 92093, USA (HEH)

The availability of complete genome sequences and genetic linkage maps for a growing number of mammalian species is opening up exciting new opportunities for studies of evolutionary change in natural populations. For example, multilocus mapping approaches hold the promise of identifying the specific genetic changes that underlie ecological adaptation and reproductive isolation. The fact that many of the genomic resources that have been developed for *Mus* and *Rattus* are transferable to other muroid rodents means that roughly 25% of all mammalian species can now be considered ''genome-enabled'' study organisms to varying degrees. The transferability of genomic resources between model organisms and their more ecologically interesting kin should usher in a renaissance period of research on adaptation and speciation in mammals.

Key words: adaptation, genomics, hybrid zones, *Mus*, natural selection, *Peromyscus*, quantitative trait locus mapping, *Rattus*, reproductive isolation, speciation

Recent advances in mammalian genomics provide many exciting new opportunities to dissect functional genetic variation that underlies adaptation and reproductive isolation. The complete sequencing of the human genome was soon followed by similar efforts to sequence and annotate the genomes of the house mouse (*Mus musculus*) and the brown Norway rat (*Rattus norvegicus*—Gibbs et al. 2004; Waterston et al. 2002). At the time of this writing, efforts are underway to sequence the genomes of taxa that are representative of most mammalian orders. The availability of these genomic data will open up new research directions that promise to illuminate mechanisms and processes of evolutionary change in ways that were probably difficult to imagine back in the early to mid-1900s when people like Lee R. Dice, W. Frank Blair, Francis B. Sumner, and others were engaged in their pioneering studies of adaptation and speciation in natural populations of small mammals.

Because of the wealth of genomic tools and resources that have been developed for model organisms such as *Mus* and *Rattus*, these 2 rodent taxa are especially well-suited to studies of functional genetic variation in an evolutionary context. In recent years, a number of researchers have capitalized on this potential to address important evolutionary questions about the genetic basis of adaptation and reproductive isolation (Harr 2006; Ihle et al. 2006; Kohn et al. 2000, 2003; Payseur and Hoekstra 2005; Payseur et al. 2004; Payseur and Nachman 2005). For example, Payseur et al. (2004) screened patterns of variation at X-linked DNA markers across a European hybrid zone between *Mus domesticus* and *M. musculus* in an effort to identify regions of the X-chromosome that contribute to reproductive isolation between these 2 recently diverged taxa. This study was motivated by 2 well-established patterns of postzygotic reproductive isolation in animals: reproductive isolation typically arises from disrupted epistatic interactions between 2 or more loci, and a disproportionate number of these genic incompatibilities map to the X-chromosome. By analyzing patterns of clinal variation at X-linked DNA markers with known map positions, Payseur et al. (2004) identified a region of the X-chromosome that was characterized by a significantly reduced level of introgression across the hybrid zone, suggesting that this chromosomal region may harbor 1 or more genes involved in reproductive isolation between *M. domesticus* and *M. musculus*. DNA markers within this region of reduced introgression also exhibited a shift in cline position to the west (toward the *M. domesticus* side of the hybrid zone), suggesting that the reproductive incompatibilities may be attributable to disrupted interactions between the *M. domesticus* X-chromosome and other (unidentified) regions of the *M. musculus* genome. Interestingly, an independent survey of nucleotide variation between wild-derived inbred strains of *M. domesticus* and *M. musculus* revealed that this same region of the X-chromosome is characterized by an unusually high level of differentiation compared to the genome-wide average (Harr 2006).

* *Correspondent: jstorz2@unl.edu*

Surveys of gene flow across hybrid zones have been conducted in many different taxa (Barton and Hewitt 1985, 1989). However, the study by Payseur et al. (2004) represents the 1st such survey in a species that has a completely sequenced genome. This type of study holds the promise of identifying and characterizing the specific genetic changes involved in speciation. Because laboratory crosses suggest that reproductive isolation between *M. domesticus* and *M. musculus* is primarily attributable to hybrid male sterility, genes that are expressed exclusively in the male germ line represent the most obvious candidate genes for reproductive isolation. Accordingly, Payseur and Nachman (2005) identified 3 such genes (*Tktl1, Halapx*, and *Tex11*) that mapped to the general region of the X-chromosome that exhibited reduced introgression in the *M. domesticus–M. musculus* hybrid zone.

An example of how genomic information can be used to obtain insights into the genetic basis of adaptation is provided by a multilocus mapping study of warfarin resistance in natural populations of *R. norvegicus* (Kohn et al. 2000, 2003). By measuring nonrandom patterns of association among positionally mapped DNA markers in a warfarin-resistant rat population, Kohn et al. (2000) were able to localize the genomic position of *Rw,* a gene that plays a well-documented role in resistance to anticoagulant rodenticides. By comparing patterns of variation at *Rw*-linked markers with patterns at unlinked markers, Kohn et al. (2003) were then able to assess variation in the response to selection for anticoagulant resistance among rat populations that were subject to different warfarin treatment regimes.

At present, opportunities for dissecting the genetic basis of ecologically important traits in nonhuman mammals are largely restricted to model organisms such as *Mus* and *Rattus*. The irony of studying ''ecologically important traits'' in *Mus* and *Rattus* is that neither of these commensal organisms have an ecology that is independent of human activity. Fortunately, *Mus* and *Rattus* are both nested within a highly diverse clade of rodents—superfamily Muroidea—that comprises nearly 25% of all mammalian species. The fact that many of the genomic resources that have been developed for *Mus* and *Rattus* are transferable to their muroid kin means that ecologically interesting taxa such as *Peromyscus*, *Neotoma*, *Apodemus*, and others can now be considered ''genome-enabled'' study organisms to varying degrees. As study organisms for research in evolutionary biology, differences between *M. musculus* and a species such as *Peromyscus maniculatus* are in many ways similar to the differences between *Drosophila melanogaster* and *D. pseudoobscura* (which served as the main study organism for much of the pioneering work of Theodosius Dobzhansky on the genetics of speciation). Like *D. pseudoobscura*, *P. maniculatus* does not offer all the advantages of being a true model organism, but for certain research questions, any shortcomings related to experimental tractability are at least partially offset by the fact that it has a more interesting and well-studied ecology than its commensal cousins.

In a similar way, now that a draft sequence of the dog genome is available (Lindblad-Toh et al. 2005; Ostrander and Wayne 2005), the door is open for wild canids such as wolves, coyotes, and jackals to join the ranks of genome-enabled taxa. It will be extremely interesting to see if the genetic changes that underlie behavioral, physiological, and morphological differences among domestic dog breeds are similar to those that distinguish different species or subspecies of wild canids.

Over the last few decades, a major focus of evolutionary biology has been the detection of adaptive genetic change in natural populations. This area of research has been invigorated by the increasing number of genome sequences and genetic linkage maps that have become available in recent years. In principle, inferences about the incidence and nature of adaptive change in genomes can be accomplished by using either ''bottom-up'' or ''top-down'' approaches. The bottom-up approach focuses on detailed study of variation in single genes that are connected to a fitness-related phenotype (Eanes 1999; Watt and Dean 2000). With the increasing availability of genomic data and genetic mapping resources, the top-down approach to studying functional genetic variation is now becoming increasingly feasible for a growing number of organisms. This approach focuses on finding genes that underlie traits of interest by means of multilocus mapping methods, such as quantitative trait locus mapping. This approach involves crossing phenotypically distinct parental types to create 1st generation (F_1) hybrids that themselves are crossed to produce a 2nd generation (F_2) mapping population. The F_2 progeny are then genotyped at a genome-wide panel of DNA markers to identify chromosomal regions that are statistically associated with variation in the phenotype of interest. This general approach has been widely employed in efforts to identify genes that underlie traits of agronomic importance in cattle and other domestic animals, but only recently has the approach been used to study the genetic architecture of ecologically important traits in natural populations of nonhuman mammals.

For example, Hoekstra et al. (2006) conducted an association study to assess the contribution of the melanocortin-1 receptor gene (*Mc1r*) to pigmentation differences between light-colored beach mice that inhabit Gulf Coast barrier islands (*Peromyscus polionotus leucocephalus*) and dark-colored conspecifics from the mainland (*P. polionotus subgriseus*). Classic work by Sumner (1929a, 1929b) demonstrated that this genetically based coat-color variation between the coastal and mainland forms is attributable to geographically localized selection for crypsis in environments that are characterized by different substrate colors. Hoekstra et al. (2006) conducted a reciprocal F_2 intercross between phenotypically distinct beach mice and their mainland counterparts and found that a single charge-changing amino acid mutation in *Mc1r* explained 9.8–36.4% of the variance in 7 different pigmentation phenotypes that are involved in background-matching. This study demonstrates that single nucleotide changes can have surprisingly large effects on quantitative trait variation. This study also suggests that structural variation (i.e., changes in the amino acid sequence of a protein) may have played a more important role than regulatory variation (i.e., changes in developmental timing or tissue specificity of gene expression) in the evolution of the light-colored phenotype. Interestingly, nucleotide variation in

the *Mc1r* gene was not associated with light coloration in beach mice from the Atlantic coast of Florida, which demonstrates that functionally similar phenotypes may have different genetic underpinnings in different populations. A similar pattern was revealed by studies of coat-color variation in rock pocket mice (*Chaetodipus intermedius*), as a combination of 4 amino acid polymorphisms in the *Mc1r* gene contributed to adaptive melanism in some populations but not in others (Hoekstra and Nachman 2003; Nachman et al. 2003).

The top-down approach can also be used to find genes that underlie fitness-related traits (irrespective of function) by using genome scans of DNA polymorphism to detect footprints of recent selection (Luikart et al. 2003; Storz 2005). This approach has been applied to multilocus data from natural populations of house mice (Ihle et al. 2006) and deer mice (Storz and Dubach 2004; Storz and Nachman 2003), as well as breeds of domestic dog (Pollinger et al. 2005). It also is possible to combine complementary approaches that integrate phenotype-based mapping approaches that are based on laboratory crosses and genome scans for selection that are based on samples from natural populations (Vasemagi and Primmer 2005). In principle, this integrative approach can provide confirmatory evidence that quantitative trait loci underlying a particular trait have contributed to a past response to selection (Storz 2005).

Our knowledge of genome structure in mammals is far more complete than it is for any other vertebrate taxon. As a result, mammalogists are uniquely positioned to press genomic data into the service of evolutionary studies. With the resources that are now available, mammals should be at center stage in this exciting new growth area in biology.

The accompanying papers (Gatesy and Swanson 2007; Storz 2007; Vrana 2007) represent a small sampling of case studies in which genetic data provide insights into different aspects of evolutionary change in mammals. Vrana (2007) discusses the interesting phenomenon of genomic imprinting and its possible role as a mechanism of reproductive isolation in mammals. Storz (2007) provides an example of a bottom-up approach to the study of adaptive genetic variation in natural populations. Because hemoglobin is known to play a critical role in the physiological response to hypoxic stress, detailed study of variation in globin genes provides a unique opportunity to elucidate the molecular underpinnings of adaptation to high-altitude environments. Finally, Gatesy and Swanson (2007) provide evidence for a history of positive selection on a gene that encodes the fertilization protein, acrosin. As is the case with many other reproductive proteins, acrosin is characterized by an accelerated rate of amino acid substitution that may be attributable to sexual selection. Because of the relatively rapid rate of evolution of reproductive proteins, the authors argue that genes such as acrosin may be especially useful for resolving phylogenetic relationships among taxa that have undergone rapid radiations.

The coming flood of mammalian genome sequence data will eventually have a democratizing effect on research in mammalian evolutionary biology because it will facilitate studies of functional genetic variation in taxa other than *Mus* and *Rattus*. For this reason, the coming years should usher in a renaissance period for studies of mammalian adaptation and speciation under natural conditions.

Literature Cited

Barton, N. H., and G. M. Hewitt. 1985. Analysis of hybrid zones. Annual Review of Ecology and Systematics 16:113–148.

Barton, N. H., and G. M. Hewitt. 1989. Adaptation, speciation, and hybrid zones. Nature 341:497–503.

Eanes, W. F. 1999. Analysis of selection on enzyme polymorphisms. Annual Review of Ecology and Systematics 30:301–326.

Gatesy, J., and W. J. Swanson. 2007. Adaptive evolution and phylogenetic utility of *Acr* (acrosin), a rapidly evolving mammalian fertilization gene. Journal of Mammalogy 88:32–42.

Gibbs, R. A., et al. 2004. Genome sequence of the brown Norway rat yields insights into mammalian evolution. Nature 428:493–521.

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

Hoekstra, H. E., R. J. Hirschmann, R. A. Bundey, P. A. Insel, and J. P. Crossland. 2006. A single amino acid mutation contributes to adaptive beach mouse color pattern. Science 313:101–104.

Hoekstra, H. E., and M. W. Nachman. 2003. Different genes underlie adaptive melanism in different populations of rock pocket mice. Molecular Ecology 12:1185–1194.

Ihle, S., I. Ravaoarimanana, M. Thomas, and D. Tautz. 2006. An analysis of signatures of selective sweeps in natural populations of the house mouse. Molecular Biology and Evolution 23:790–797.

Kohn, M. H., H. J. Pelz, and R. K. Wayne. 2000. Natural selection mapping of the warfarin-resistance gene. Proceedings of the National Academy of Sciences of the United States of America 97:7911–7915.

Kohn, M. H., H. J. Pelz, and R. K. Wayne. 2003. Locus-specific genetic differentiation at *Rw* among warfarin-resistant rat (*Rattus norvegicus*) populations. Genetics 164:1055–1070.

Lindblad-Toh, K., et al. 2005. Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature 438:803–819.

Luikart, G., P. R. England, D. Tallmon, S. Jordan, and P. Taberlet. 2003. The power and promise of population genomics: from genotyping to genome typing. Nature Reviews. Genetics 4:981–994.

Nachman, M. W., H. E. Hoekstra, and S. L. D'Agostino. 2003. The genetic basis of adaptive melanism in pocket mice. Proceedings of the National Academy of Sciences of the United States of America 100:5268–5273.

Ostrander, E. A., and R. K. Wayne. 2005. The canine genome. Genome Research 15:1706–1716.

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

Payseur, B. A., J. G. Krenz, and M. W. Nachman. 2004. Differential patterns of introgression across the X chromosome in a hybrid-zone between two species of house mice. Evolution 58:2064–2078.

Payseur, B. A., and M. W. Nachman. 2005. The genomics of speciation: investigating the molecular correlates of X chromosome introgression across the hybrid zone between *Mus domesticus* and *Mus musculus*. Biological Journal of the Linnean Society 84: 523–534.

Pollinger, J. P., C. D. Bustamante, A. Fledel-Alon, S. Schmutz, M. M. Gray, and R. K. Wayne. 2005. Selective sweep mapping of genes with large phenotypic effects. Genome Research 15:1809–1819.

STORZ, J. F. 2005. Using genome scans of DNA polymorphism to infer adaptive population divergence. Molecular Ecology 14:671–688.

STORZ, J. F. 2007. Hemoglobin function and physiological adaptation to hypoxia in high-altitude mammals. Journal of Mammalogy 88:24–31.

STORZ, J. F., AND J. M. DUBACH. 2004. Natural selection drives altitudinal divergence at the albumin locus in deer mice, *Peromyscus maniculatus*. Evolution 58:1342–1352.

STORZ, J. F., AND M. W. NACHMAN. 2003. Natural selection on protein polymorphism in the rodent genus *Peromyscus*: Evidence from interlocus contrasts. Evolution 57:2628–2635.

SUMNER, F. B. 1929a. The analysis of a concrete case of intergradation between two subspecies. Proceedings of the National Academy of Sciences of the United States of America 15:110–120.

SUMNER, F. B. 1929b. The analysis of a concrete case of intergradation between two subspecies. II. Additional data and interpretations. Proceedings of the National Academy of Sciences of the United States of America 15:481–493.

VASEMAGI, A., AND C. R. PRIMMER. 2005. Challenges for identifying functionally important genetic variation: the promise of combining complementary research strategies. Molecular Ecology 14:3623–3642.

VRANA, P. B. 2007. Genomic imprinting as a mechanism of reproductive isolation in mammals. 88:5–23.

WATERSTON, R. H., ET AL. 2002. Initial sequencing and comparative analysis of the mouse genome. Nature 420:520–562.

WATT, W. B., AND A. M. DEAN. 2000. Molecular-functional studies of adaptive genetic variation in prokaryotes and eukaryotes. Annual Review of Genetics 34:593–622.

Special Feature Editor was Edward J. Heske.