University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

Jay F. Storz Publications | Papers in the Biological Sciences

2008

# Molecular evolution of cytochrome *b* in high- and low-altitude deer mice (genus *Peromyscus*)

E.J. Gering
*University of Texas, 1 University Station C0930, Austin, TX*, eben@mail.utexas.edu

J. C. Opazo
*Instituto de Ecologia y Evolucion, Facultad de Ciencias, Universidad Austral de Chile, Casilla 567, Valdivia, Chile*

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

Follow this and additional works at: https://digitalcommons.unl.edu/bioscistorz

 Part of the Genetics and Genomics Commons

Gering, E.J.; Opazo, J. C.; and Storz, Jay F., "Molecular evolution of cytochrome *b* in high- and low-altitude deer mice (genus *Peromyscus*)" (2008). *Jay F. Storz Publications*. 23.
https://digitalcommons.unl.edu/bioscistorz/23

This Article is brought to you for free and open access by the Papers in the Biological Sciences at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Jay F. Storz Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Published in *Heredity* (2008), doi: 10.1038/hdy.2008.124 Copyright © 2008 The Genetics Society; published by Macmillan/Nature Publishing Group. Used by permision. http://www.nature.com/hdy/index.html

Submitted March 19, 2008; revised November 11, 2008; accepted November 21, 2008; published online December 24, 2008.

# Molecular evolution of cytochrome *b* in high- and low-altitude deer mice (genus *Peromyscus*)

E. J. Gering,* J. C. Opazo,† and J. F. Storz

School of Biological Sciences, University of Nebraska–Lincoln, Lincoln, NE, USA

* *Corresponding author* – current address: Section of Integrative Biology, University of Texas, 1 University Station C0930, Austin, TX 78712, USA; email eben@mail.utexas.edu
† Current address: Instituto de Ecologia y Evolucion, Facultad de Ciencias, Universidad Austral de Chile, Casilla 567, Valdivia, Chile.

**Abstract**
Patterns of amino-acid polymorphism in human mitochondrial genes have been interpreted as evidence for divergent selection among populations that inhabit climatically distinct environments. If similar patterns are mirrored in other broadly distributed mammalian species, then adaptive modifications of mitochondrial protein function may be detected in comparisons among locally adapted populations of a single wide-ranging species, or among closely related species that have adapted to different environments. Here, we test for evidence of positive selection on cytochrome *b* variation within and among species of the ecologically diverse rodent genus *Peromyscus*. We used likelihood-based comparisons of synonymous and nonsynonymous substitution rates to test for evidence of divergent selection between high- and low-altitude haplogroups of the deer mouse, *Peromyscus maniculatus*. We also tested for evidence of divergent selection among different species of *Peromyscus* that inhabit different thermal environments. In contrast to the purported evidence for positive selection on mitochondrial proteins in humans and other nonhuman mammals, results of our tests suggest that the evolution of cytochrome *b* in *Peromyscus* is chiefly governed by purifying selection.

**Keywords:** adaptation, cytochrome *b*, deer mouse, McDonald–Kreitman test, positive selection, *Peromyscus*

## Introduction

The mitochondrial cytochrome *b* gene encodes an integral membrane protein component of the cytochrome *bc*1 complex, which catalyzes the redox transfer of electrons from ubiquinone to cytochrome *c* in the mitochondrial electron transport chain. As the efficiency of the electron transport chain governs key aspects of aerobic energy metabolism (reviewed by Rolfe and Brown, 1997), several investigators have suggested that functional modifications of redox proteins, such as cytochrome *b*, may be involved in physiological adaptation to different thermal environments (Mishmar *et al.*, 2003; Ruiz-Pesini *et al.*, 2004; Fontanillas *et al.*, 2005). For example, Mishmar *et al.* (2003) and Ruiz-Pesini *et al.* (2004) argued that patterns of amino-acid polymorphism in human mitochondrial genes reflect a history of local adaptation to different climatic regimes. Global surveys of nucleotide variation in human mitochondrial genes show that ratios of nonsynonymous to synonymous substitution rates are significantly higher in European/temperate and Asian (Siberian)/Arctic haplogroups than in the African/tropical haplogroups (Mishmar *et al.*, 2003; Ruiz-Pesini *et al.*, 2004; but see also Sun *et al.*, 2007). This pattern is consistent with positive selection or a relaxation of purifying selection in the high-latitude populations (Elson *et al.*, 2004; Kivisild *et al.*, 2006). A second pattern that has emerged from phylogenetic analysis of human mtDNA variation is that physicochemically radical amino-acid substitutions are generally more common in terminal branches of the haplotype genealogies than in the older, internal branches (Moilanen and Majamaa, 2003). This pattern further suggests a history of purifying selection against mildly deleterious amino-acid variants and is consistent with a commonly observed excess of low-frequency amino-acid polymorphisms in mitochondrial genes of humans and other animals (Nachman *et al.*, 1996; Hasegawa *et al.*, 1998; Nachman, 1998; Rand and Kann, 1996, 1998; Fry, 1999). However, in human mtDNA haplogroups that are characteristic of temperate and Arctic regions, a significantly disproportionate number of physicochemically radical amino-acid substitutions were concentrated in the internal branches of the haplotype genealogies, a pattern consistent with a history of lineage-specific positive selection (Ruiz-Pesini *et al.*, 2004). In cytochrome *b* haplogroups characteristic of north-temperate Eurasia, Ruiz-Pesini *et al.* (2004) identified physicochemically radical substitutions in highly conserved coenzyme Q (CoQ)-binding sites. As the CoQ-binding sites of cytochrome *b* play a critical role in the complex III Q cycle, these substitutions may influence the allocation of energy derived from oxidative phosphorylation to metabolic heat production.

If patterns of amino-acid polymorphism in human mitochondrial genes reflect a history of divergent selection among populations that inhabit climatically distinct environments, as suggested by Mishmar *et al.* (2003) and Ruiz-Pesini *et al.* (2004), then it stands to reason that parallel patterns of

adaptive variation may be observed in other broadly distributed mammalian species that inhabit a similar diversity of thermal environments. Modifications of mitochondrial protein function may be evident in comparisons between locally adapted populations of a wide-ranging species, or between closely related species that have adapted to different environments.

Here, we test for evidence of positive selection on cytochrome *b* variation within and among species of the ecologically diverse rodent genus, *Peromyscus*. Mice in the genus *Peromyscus* inhabit an extremely diverse range of habitat types, ranging from alpine and subarctic environments to subtropical cloud forests in Central America (Hall, 1981; Carleton, 1989). An appreciable fraction of this environmental heterogeneity is also contained within the range of the most widely distributed species in the genus, *Peromyscus maniculatus* (Hall, 1981). As might be expected, this ecological variation within and among *Peromyscus* species is mirrored by extensive variation in basal metabolic rates and thermal tolerances (MacMillen and Garland, 1989).

If the highly conserved functions of the electron transport chain have undergone adaptive modifications in human populations that have colonized cold, high-latitude environments within the last 10,000–50,000 years, then similar modifications of redox proteins, such as cytochrome *b*, may be evident in *Peromyscus* species or subspecies that have inhabited high altitude or high latitude environments for much longer periods of evolutionary time. In addition to differences in the likely timescale of environmental adaptation, species like *P. maniculatus* are characterized by much larger effective population sizes than humans, so weak selection on amino-acid variation is more likely to dominate drift.

To test for evidence of divergent selection on cytochrome *b* variation in deer mice, we compared ratios of synonymous and nonsynonymous substitution among the branches of species- and genus-level cytochrome *b* phylogenies. To test for evidence of divergent selection within *P. maniculatus*, we also made use of neutrality tests that compare ratios of polymorphism to divergence at both synonymous and nonsynonymous sites. Finally, we used an alignment of cytochrome *b* from 632 mammalian species together with structural models of the *bc*1 complex to ascertain the functional importance of amino-acid substitutions occurring within *P. maniculatus* and between *Peromyscus* species.

## Materials and methods

### *Taxon sampling*

For the analysis of species-level variation (within the *P. maniculatus* species complex), we used cytochrome *b* sequences from a total of 311 specimens collected from 34 localities across North America (Table S1, Supplementary material Online). We sequenced cytochrome *b* from a total of 143 specimens and we obtained the remaining 168 sequences from GenBank (NCBI). Four sequences from *P. keeni* and two sequences from *P. melanotis* were also included in the alignment, as previous studies suggested that these taxa may be part of a 'species complex' that includes several genetically distinct lineages that are at present regarded as subspecies of *P. maniculatus* (Carleton, 1989; Hogan *et al*., 1993; Dragoo *et al*., 2006).

For analyses of cytochrome *b* variation among different species of *Peromyscus*, we used a total of 38 sequences representing 26 nominal species with diverse altitudinal and latitudinal ranges (Table S2, Supplementary material online). This set of species was selected with the goal of maximizing phylogenetic coverage, while maintaining a relatively even distribution of branch lengths in the reconstructed tree. In the cases of *P. betae*, *P. boylii*, *P. maniculatus,* and *P. truei*, we included specimens that were representative of multiple, genetically distinct subspecies because physiological traits that may be influenced by efficiency of the electron transport chain are known to vary at the intraspecific level (Cook and Hannon, 1954; Chappell and Snyder, 1984; MacMillen and Garland, 1989; Storz, 2007).

As explained below (see Results), the species-level and genus-level phylogeny reconstructions both showed that *P. melanotis* was basal to a monophyletic group that comprised the entire *P. maniculatus* species complex. We therefore sequenced cytochrome *b* from 13 additional *P. melanotis* specimens (Supplementary Table S1) for the purpose of making outgroup comparisons with *P. maniculatus* in McDonald–Kreitman tests (McDonald and Kreitman, 1991).

### *PCR and sequencing*

Genomic DNA was isolated from tissue samples with the DNeasy Tissue Kit (Qiagen Inc., Valencia, CA, USA) according to the manufacturer's specifications. We amplified the cytochrome *b* gene using the primers Cytochrome-*b* 15334L (5′-CTTCATTTTTGGTTTACAAGAC-3′) and L14724 (5′-TGATATGAAAAACCATCGTTG-3′) with a 5 min denaturing period (95 °C) followed by 30 cycles of denaturing (94 °C ×30 s), annealing (55 °C ×30 s) and extension (72 °C ×60 s). PCRs were performed using 2.5 µl of 10 ×Buffer II (Applied Biosystems, Foster City, CA, USA), 2.5 µl of 25 mM $MgCl_2$, 2.5 vl of 10 mM dNTPs (2.5 µM each, dATP, dGTP, dCTP, dTTP), approximately 100 ng template DNA (0.5–2 µl DNeasy elution), 1 U of Ampli-*Taq* Gold polymerase (Applied Biosystems) and $ddH_2O$ to make a total volume of 25 µl. PCR products were purified using ExoSAP-IT (USB Corporation, Cleveland, OH, USA), cycle-sequenced with Big-Dye chemistry and run on an ABI 3730 sequencer (Applied Biosystems). The resulting chromatograms were visually inspected and aligned using Sequencher (Genecodes Corporation, Ann Arbor, MI, USA), and amino-acid translations of DNA sequence were conducted with MacClade (Maddison and Maddison, 1989).

### *Phylogenetic reconstruction*

We used Bayesian and maximum likelihood methods to reconstruct phylogenetic relationships among nonredundant cytochrome *b* haplotypes of *P. maniculatus*, and among species of *Peromyscus*. Sequences for *Reithrodontomys raviventer* and *P. leucopus* obtained from Genbank were used as outgroups for genus-level (*Peromyscus*) and species-level (*P. maniculatus*) data sets, respectively (for accession numbers, refer to Supplementary Tables S1 and S2). Owing to incomplete sequences from some representatives of the *P. man-*

*iculatus* species complex, the species-level analysis was restricted to codons 19–374 of the cytochrome *b* gene. The genus-level data set included only *Peromyscus* species for which complete sequences were available, and analyses were therefore performed on the complete 381 codon sequence of the gene.

We used Modeltest (Posada and Crandall, 1998) to select substitution models for our likelihood analyses. For both genus-level and species-level alignments, the best-fitting model was the parameter-rich GTR+G+I. We implemented this same model in our Bayesian analyses to allow for direct comparisons of tree topologies, and because overparameterization is unlikely to introduce type-one or type-two errors into Bayesian phylogenetic reconstructions (Lemmon and Moriarty, 2004).

Bayesian phylogeny reconstructions were conducted using MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003). We fit the GTR+G+I model to each codon position and conducted two simultaneous independent runs of 10,000,000 generations with four simultaneous chains, sampling every 1,000th tree. Burn-in values were varied according to the time required for likelihood values to converge. To obtain 50% consensus trees, we discarded the first 1,500 sampled trees in the species-level analysis and the first 500 trees in the genus-level analysis.

Maximum likelihood reconstructions were conducted using TreeFinder (Jobb *et al.*, 2004). We pruned the species-level data set to 40 sequences by selecting representative haplotypes from each well-supported clade in the Bayesian phylogeny. We used 1,000 pseudoreplicates to compute bootstrap support values.

Summary statistics of DNA polymorphism and divergence were calculated using the program DnaSP (Rozas *et al.*, 2003), and levels of nucleotide divergence between *P. maniculatus* clades were estimated using MEGA 3.1 (Kumar *et al.*, 2004).

### *Codon-based analysis of selection pressure*

To test for natural selection on cytochrome *b* coding sequence, we used a maximum likelihood approach to examine variation in the ratio of nonsynonymous and synonymous substitutions, $d_N/d_S$ (= $\omega$) (Goldman and Yang, 1994; Nielsen and Yang, 1998; Yang and Nielsen, 2000; Bielawski and Yang, 2004). Positive selection is implicated when $\omega > 1$, whereas purifying selection is implicated when $\omega < 1$. Analyses of the species-level and genus-level cytochrome *b* alignments were conducted with the codeml program in the PAML 3.15 package (Yang, 1997; http://abacus.gene.ucl.ac.uk/software/paml.html). All analyses were run three times with different starting values for the $\omega$ ratio (0.5, 1 and 2). Maximum likelihood scores for nested models were compared using likelihood ratio tests (Felsenstein, 1981).

For the species-level *P. maniculatus* data set, we first tested for heterogeneity in $\omega$ ratios among branches of the inferred phylogeny by comparing log-likelihood scores for a model that allowed a single $\omega$ ratio (model M0) vs a model that allowed separate $\omega$ ratios for each branch in the phylogeny (model M1). We then tested for evidence of positive selection (as indicated by $\omega > 1$) in specific branches of the *P. maniculatus* phylogeny. We estimated $\omega$ for distinct haplogroups that were characteristic of the Great Plains (the low-altitude haplogroup) and montane environments of the Southern Rocky Mountains (the high-altitude haplogroup). We included 1–3 representative sequences from each branch of the 50% consensus Bayesian phylogeny in these analyses, and compared log-likelihood scores for three evolutionary models. The first model allowed two $\omega$ ratios, one for the pooled sample of high- and low-altitude haplogroups, and a second $\omega$ ratio for the remaining 'background' branches of the intraspecific gene genealogy. The second model allowed $\omega$ ratios to vary among haplogroups by assuming three separate $\omega$ ratios: one for the high-altitude haplogroup, one for the low-altitude haplogroup and a third one for the remaining background branches. The third model allowed five separate $\omega$ ratios for the following partitions of the *P. maniculatus* gene genealogy: (i) the stem of the high-altitude haplogroup; (ii) the crown of the high-altitude haplogroup; (iii) the stem of the low-altitude haplogroup; (iv) the crown of the low-altitude haplogroup; and (v) the remaining background branches of the phylogeny. Here and elsewhere, the stem of a given clade refers to the ancestral branch leading to the most recent common ancestor of the set of sequences under consideration, whereas the crown group includes the most recent common ancestor and all descendent sequences. By allowing separate $\omega$ ratios for the stem and the crown of a given clade, it may be possible to detect an early acceleration of the nonsynonymous substitution rate that was eventually superseded by purifying selection (Goodman, 1982).

Using the genus-level data set, we designed two intermediate nested models to investigate variation in the $\omega$ ratio during the diversification of the *P. maniculatus* species complex. The first model allowed one $\omega$ ratio for the *P. maniculatus* clade and a second $\omega$ ratio for the remaining background branches of the phylogeny. The second model allowed three $\omega$ ratios: one for the stem of the *P. maniculatus* clade, a second for the crown and a third for the remaining background branches of the phylogeny. One limitation of the intermediate nested models is that each assumes a single $\omega$ ratio for all amino-acid sites in the cytochrome *b* alignment, and may therefore fail to detect positive selection on a small subset of sites. We therefore fitted a branch-site model (Yang and Nielsen, 2002; Zhang *et al.*, 2005) to the *Peromyscus* data set that allowed us to test for rate heterogeneity across sites within the *P. maniculatus* clade. Model A allowed four classes of sites: site class 0 included codons that were subject to stringent functional constraints with $0 < \omega_0 < 1$ estimated from the data, site class 1 included codons that were unconstrained with $\omega_1 = 1$ fixed and site classes 2a and 2b included codons that were conserved or neutral on the background branches but were allowed to undergo a shift to positive directional selection in the foreground lineage, with $\omega_2$ estimated as a free parameter. We used model A to construct a likelihood ratio test in which the null hypothesis was a simplified version of model A with $\omega_2 = 1$ fixed, following Zhang *et al.* (2005). This null model allowed a shift from purifying selection on the background branches to relaxed constraint on the foreground branches. Consequently, the likelihood ratio test involving the comparison of model A *vs.* the null model allowed us to test whether an accelerated rate of non-

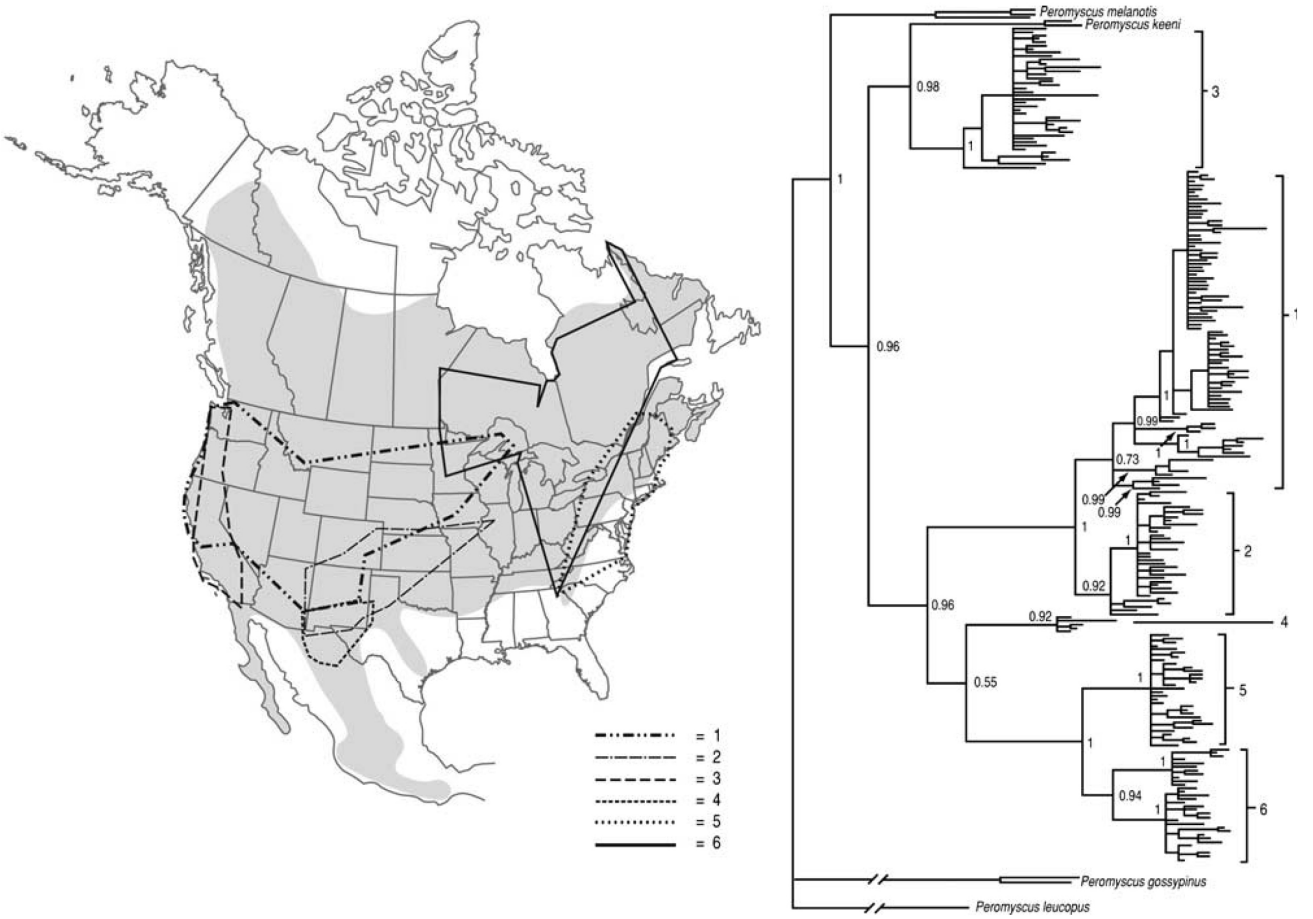


**Figure 1.** Bayesian phylogeny of cytochrome *b* sequences from 311 *P. maniculatus* specimens on the basis of a GTR+I+G model of nucleotide substitution. Bayesian posterior probabilities are given above each branch. A similar topology was recovered in a maximum likelihood phylogeny of 40 randomly selected sequences; all labeled nodes were supported by > 50% bootstrap values in 1,000 pseudo replicates. The inset map of North America shows the geographic distribution of *P. maniculatus* (shaded region) and the six major cytochrome *b* clades.

synonymous substitution in the foreground lineage could be attributed to positive selection or a relaxation of functional constraint.

## Results and discussion

### *Cytochrome b variation in the P. maniculatus species complex*

The phylogeographic structure of cytochrome *b* variation within *P. maniculatus* is largely concordant with results obtained in earlier studies of mtDNA variation in this species complex (Avise *et al.*, 1983; Lansman *et al.*, 1983; Dragoo *et al.*, 2006). Our phylogeny reconstruction of the 311 cytochrome *b* sequences from the *P. maniculatus* species complex revealed six distinct clades (Figure 1), which we designate as clades 1–6 (following Dragoo *et al.*, 2006). Whereas Dragoo *et al.* (2006) found *P. melanotis* to be nested within the *P. maniculatus* species complex, we found *P. melanotis* to be sister to a monophyletic group containing all *P. maniculatus* haplotypes and the nominal species *P. keeni*. Monophyly of the *maniculatus*+*keeni* clade was supported by 78.3% of maximum likelihood bootstrap replicates and by a Bayesian posterior probability of 0.96.

With the exception of clade 3, which is restricted to the southwestern United States, inferred relationships of all major cytochrome *b* clades were well supported in Bayesian and maximum likelihood trees. Net nucleotide divergences between clades ranged from 1.7 to 4.4% (Table 1). At the interface between the Great Plains and the Southern Rocky Mountains, we observed a clear phylogeographic break between a high-altitude haplogroup (contained within clade 1) and a low-altitude haplogroup (contained within clade 2; Figure 2).

Consistent with the findings of Dragoo *et al.* (2006), all haplogroups were characterized by an excess of low-frequency polymorphisms, as reflected by negative values for Tajima's *D* and Fu and Li's $D^*$ (Table 2). With the exception of clade 4, which was represented in our data set by only six individuals, the values for Tajima's *D* and Fu and Li's $D^*$ were significantly negative in each of the haplogroups from the Great Plains and western North America. This same excess of low-frequency polymorphisms was also observed at multiple, unlinked autosomal loci in a large sample of mice that are representative of clades 1 and 2 (Storz *et al.*, 2007; Storz and Kelly, 2008). The consistently negative skews in the site-frequency distributions of nuclear and mitochondrial loci most likely reflect a recent history of population expansion.

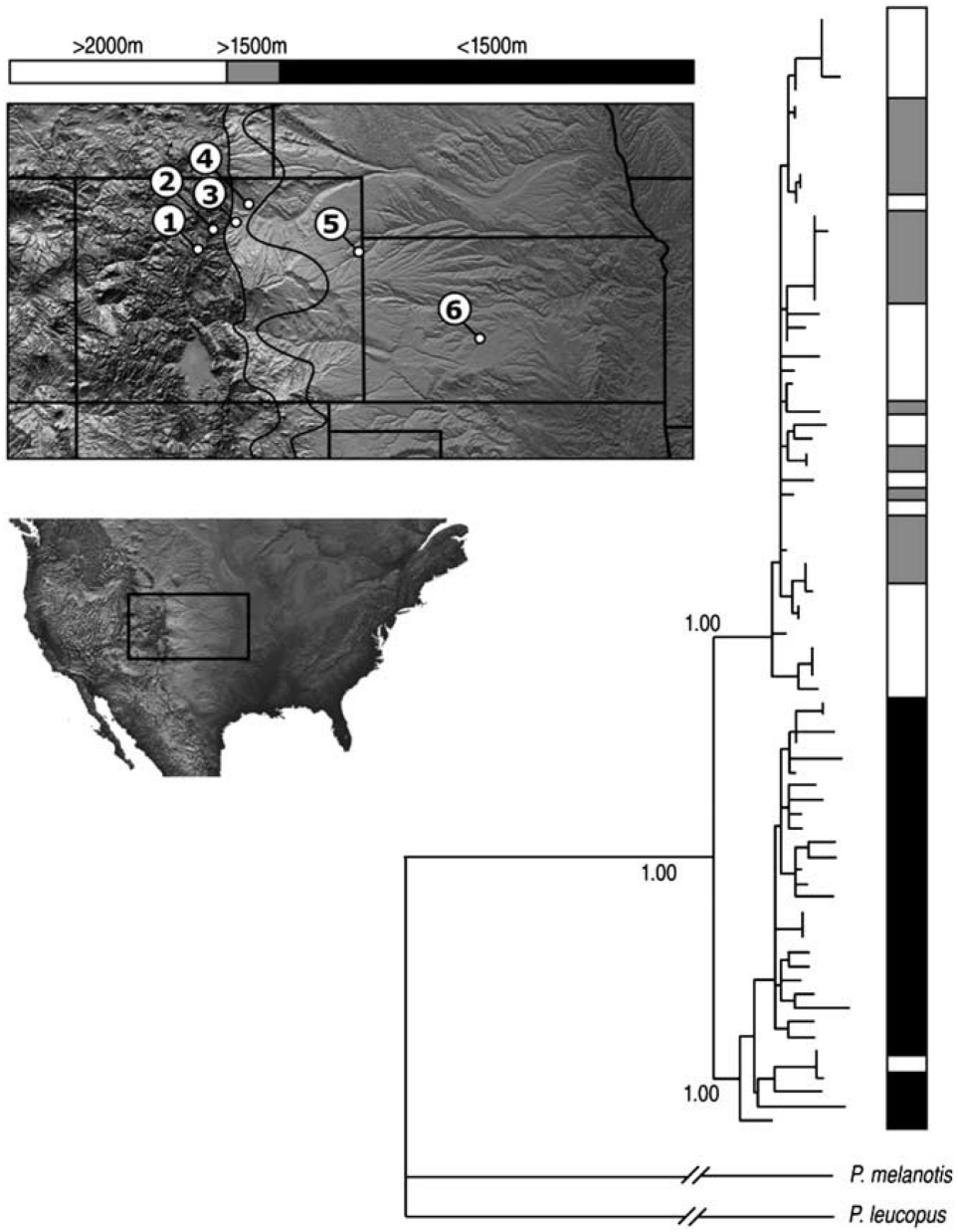


**Figure 2.** Bayesian phylogeny of cytochrome *b* sequences from deer mice that were collected along an altitudinal gradient. The gradient spanned the interface between the southern Rocky Mountains and the Great Plains. The inset map shows the location of six collecting localities in Colorado and Kansas, USA: (1) Summit of Mt. Evans, Clear Creek Co., Colorado (4,347 m); (2) Niwot Ridge, Boulder Co., CO (2,,900 m); (3) Mesa Reservoir, Boulder Co., Colorado (1,672 m); (4) Nunn, Weld Co., Colorado (1,578 m); (5) Bonny Reservoir, Yuma Co., Colorado (1,158 m); and (6) Fort Larned National Monument, Pawnee Co., Kansas (620 m). The tree was based on a GTR+I+G model of nucleotide substitution. Bayesian posterior probabilities are given above each of the main nodes. The haplogroup containing specimens from montane localities (1–4) is nested within clade '1,' and the haplogroup that primarily consists of specimens from plains localities (5–6) is nested within clade '2' (see Figure 1).

**Table 1.** Pairwise levels of net nucleotide divergence between *Peromyscus maniculatus* clades depicted in Figure 1, obtained using the Kimura 2-parameter substitution model.

| DNA haplogroup | | Clade 1 | Clade 2 | Clade 3 | Clade 4 | Clade 5 |
|---|---|---|---|---|---|---|
| Clade 1 | Central and western USA | | | | | |
| Clade 2 | Central USA | 0.017 (0.003) | | | | |
| Clade 3 | Western USA | 0.034 (0.006) | 0.035 (0.006) | | | |
| Clade 4 | Southwestern USA | 0.044 (0.006) | 0.043 (0.006) | 0.029 (0.005) | | |
| Clade 5 | Northeastern USA and eastern Canada | 0.043 (0.007) | 0.040 (0.007) | 0.040 (0.006) | 0.037 (0.005) | |
| Clade 6 | Northcentral USA and central Canada | 0.038 (0.006) | 0.034 (0.006) | 0.041 (0.006) | 0.032 (0.006) | 0.015 (0.004) |

Values in parentheses represent standard errors computed from 10,000 bootstrap replicates.

Of the 311 cytochrome *b* sequences we analyzed from the *P. maniculatus* species complex (defined here as *P. maniculatus*+*P. keeni* to the exclusion of *P. gossypinus*, *P. leucopus* and *P. melanotis*), we recovered 236 unique haplotypes at the nucleotide level and 95 unique haplotypes at the amino-acid level. Twenty-two of the amino-acid haplotypes were recovered more than once in the sample. These nonunique haplotypes were distinguished by a total of 37 amino-acid substitutions at 22 sites.

**Table 2.** Levels and patterns of nucleotide variation in six cytochrome *b* clades of *Peromyscus maniculatus* depicted in Figure 1.

| Cytochrome b haplogroup | Class | S | H | HD | $\pi$ | $\theta$ | Tajima's $D$ | Fu and Li's $D^*$ |
|---|---|---|---|---|---|---|---|---|
| Clade 1 | Central and western USA (*N*=123) | Synonymous | 121 | 88 | 0.975 | 0.024 | 0.074 | –2.153** | –4.210** |
| | | Nonsynonymous | 53 | 44 | 0.801 | 0.019 | 0.077 | –2.331** | –4.765** |
| | | Total | 133 | 91 | 0.977 | 0.008 | 0.024 | –2.177** | –4.495** |
| Clade 2 | Central USA (*N*=47) | Synonymous | 64 | 35 | 0.984 | 0.018 | 0.046 | –2.156* | –3.006* |
| | | Nonsynonymous | 27 | 20 | 0.798 | 0.012 | 0.042 | –2.401** | –3.931** |
| | | Total | 74 | 35 | 0.984 | 0.006 | 0.016 | –2.246** | –3.305* |
| Clade 3 | Western USA (*N*=35) | Synonymous | 84 | 37 | 0.986 | 0.027 | 0.065 | –2.100* | –3.597* |
| | | Nonsynonymous | 48 | 29 | 0.962 | 0.031 | 0.084 | –2.222** | –4.311** |
| | | Total | 103 | 40 | 0.995 | 0.010 | 0.024 | –2.129* | –3.856** |
| Clade 4 | Southwestern USA (*N*=6) | Synonymous | 10 | 5 | 0.933 | 0.380 | 0.438 | –0.795 | –0.851 |
| | | Nonsynonymous | 2 | 2 | 0.333 | 0.333 | 0.438 | –1.132 | –1.155 |
| | | Total | 12 | 5 | 0.933 | 0.004 | 0.005 | –0.913 | –0.967 |
| Clade 5 | Northeast USA and eastern Canada (*N*=61) | Synonymous | 51 | 31 | 0.953 | 0.015 | 0.034 | –1.855* | –1.409 |
| | | Nonsynonymous | 15 | 13 | 0.569 | 0.006 | 0.022 | –2.157* | –2.250 |
| | | Total | 53 | 32 | 0.954 | 0.005 | 0.011 | –1.886* | –1.461 |
| Clade 6 | Northcentral USA and central Canada (*N*=39) | Synonymous | 61 | 29 | 0.979 | 0.192 | 0.226 | –0.528 | –0.694 |
| | | Nonsynonymous | 12 | 13 | 0.587 | 0.071 | 0.226 | –2.047* | –2.779* |
| | | Total | 73 | 32 | 0.982 | 0.012 | 0.016 | –0.999 | –1.220 |
| Total (clades 1–6) (*N*=311) | Synonymous | 245 | 226 | 0.993 | 0.099 | 0.141 | –0.914 | –3.448** | |
| | Nonsynonymous | 119 | 126 | 0.947 | 0.065 | 0.157 | –1.767* | –6.966** |
| | Total | 286 | 235 | 0.994 | 0.031 | 0.049 | –1.115 | –4.762** |

Abbreviations: S = number of polymorphic sites; H = number of haplotypes; HD = haplotype diversity; $\pi$ = nucleotide diversity; $\theta$ = proportion of segregating sites.
Tajima's $D$ and Fu and Li's $D^*$ compare observed allele frequency distributions with expectations based on neutral theory. * $P < 0.05$; ** $P < 0.01$.

We obtained conflicting results concerning the monophyly of *P. maniculatus* in species-level and genus-level phylogenies. Phylogenetic reconstructions of the species-level data set indicated that *P. keeni* and a subset of haplotypes recovered from *P. maniculatus gambelli* form a monophyletic clade that is sister to the remaining *P. maniculatus* haplotypes. Reconstructions of the genus-level data set supported monophyly of *P. maniculatus* with a Bayesian posterior probability of 0.72 and placed *P. keeni* as the most closely related congener (Figure 3). Overall, species relationships within our genus-level phylogeny were consistent with a recent, independent phylogeny of *Peromyscus* cytochrome *b* sequences (Bradley *et al.*, 2007).

The three main structural domains of the cytochrome *b* protein were characterized by pronounced differences in levels of amino-acid variation (Table 3). Within *P. maniculatus*, the matrix domain was the most variable and the intermembrane domain was the least variable. This suggests that the intermembrane domain is characterized by more stringent functional constraints, which is consistent with the findings of previous studies (Degliesposti *et al.*, 1993; McClellan *et al.*, 2005).

### *Polymorphism and divergence*

McDonald–Kreitman tests (McDonald and Kreitman, 1991) comparing polymorphism within *P. maniculatus* with divergence from *P. melanotis* revealed an excess of replacement polymorphism within the low-altitude haplogroup, within the high-altitude haplogroup, within *keeni*-like and *maniculatus*-like clades and within the pooled sample of all *P. maniculatus* cytochrome *b* sequences (Table 4). The excess of replacement polymorphism was statistically significant in the low-altitude haplogroup and in clade 3. Repeating McDonald–Kreitman tests with cytochrome *b* sequence from a more distantly related outgroup, *P. leucopus* (Figures 1 and 3), yielded similar outcomes for each clade (results not shown). This excess of low-frequency amino-acid polymorphism (reflected by neutrality indices > 1) is a pattern that is commonly observed in mitochondrial genes of humans and other animals (Nachman *et al.*, 1996; Hasegawa *et al.*, 1998; Nachman, 1998; Rand and Kann, 1996, 1998; Fry, 1999) and is typically interpreted as evidence for purifying selection against mildly deleterious mutations.

### *Molecular evolution of cytochrome b within P. maniculatus*

The free-ratio model of codon substitution provided a slightly better fit to the *P. maniculatus* data than the one-ratio model, although the difference in log-likelihood scores was not statistically significant ($2\Delta\ell = 107.98$, $df = 87$, $P < 0.063$). As this result suggests the possibility of heterogeneity in the $\omega$ ratio among the different *P. maniculatus* haplogroups, we proceeded with additional tests to determine if the $\omega$ ratio differed between high- and low-altitude haplogroups. A model that allowed two $\omega$ ratios, one for the pooled sample of high- and low-altitude haplogroups and one for the remaining background branches, provided a better fit to the data than the one-ratio model ($2\Delta\ell = 4.09$, $df = 1$, $P = 0.04$).

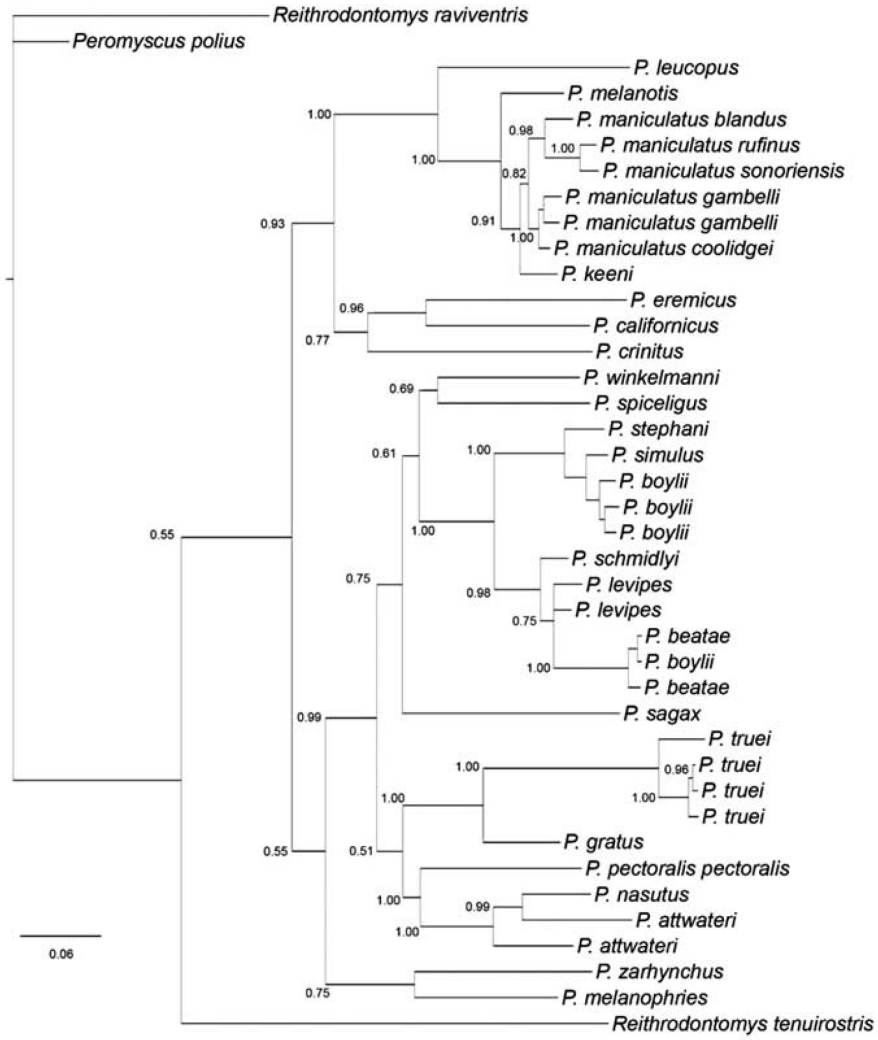


**Figure 3.** Bayesian phylogeny of cytochrome *b* sequences from 26 *Peromyscus* species. Bayesian posterior probabilities are given above each branch. A similar topology was recovered from a maximum likelihood reconstruction using the GTR+I+G model implemented in Treefinder (Jobb *et al.*, 2004); all labeled nodes had bootstrap support values > 50%.

**Table 3.** Amino-acid changes in the three functional domains of *Peromyscus maniculatus* cytochrome *b*.

| Domain | Proportion of polymorphic sites | Proportion of physicochemically radical substitutions | Proportion of replacement sites with multiple substitutions in the *P. maniculatus* phylogeny |
|---|---|---|---|
| Matrix | 0.07 (3/41) | 1.00 (10/10) | 0.90 (37/41) |
| Transmembrane | 0.09 (18/202) | 0.17 (4/24) | 0.29 (59/202) |
| Intermembrane | 0.03 (3/106) | 0.33 (1/3) | 0.00 (0/106) |
| $\chi 2$ (*df* = 2) | 3.387 (P40.05) | 35.777 ($P < 0.001$)* | 27.157 ($P < 0.001$)* |

*Significant after Bonferronni correction ($\alpha = 0.017$; $P < 0.05/3$).

In this two-ratio model, estimates of $\omega$ for each group were <1, consistent with a history of purifying selection within both groups. More complex models that allowed for additional sources of rate variation within the high- and low-altitude haplogroups did not significantly improve likelihood scores (Table 5). Thus, there is no evidence for different rates of nonsynonymous substitution in high- and low-altitude haplogroups.

*Molecular evolution of cytochrome b within the genus Peromyscus*

The free-ratio model provided a better fit to the genus-level data than the one-ratio model ($2\Delta\ell = 138.94$; *df* = 75, $P < 0.001$). As this result indicated rate heterogeneity among the branches of the genus-level *Peromyscus* phylogeny, we tested for evidence of an accelerated rate of nonsynonymous substitution within the clade representing the most broadly distributed *Peromyscus* species, *P. maniculatus*. A model that

**Table 4.** McDonald-Kreitman (1991) tests and neutrality indices (Rand and Kann, 1996) on the basis of cytochrome *b* polymorphism in *Peromyscus maniculatus* relative to divergence from *P. melanotis*.

| Haplogroup | Site class | Polymorphic sites | Divergence to *P. melanotis* | Neutrality index | *P*-value |
|---|---|---|---|---|---|
| All *P. maniculatus* | Nonsynonymous | 47 | 1 | 3.57 | 0.31 |
| | Synonymous | 194 | 13 | | |
| Clades 1–6, excluding clade 3 | Nonsynonymous | 49 | 1 | 3.14 | 0.32 |
| | Synonymous | 157 | 16 | | |
| Clade 3 | Nonsynonymous | 38 | 1 | 13.78 | < 0.01 |
| | Synonymous | 80 | 29 | | |
| High-altitude haplogroup (subset of clade 1) | Nonsynonymous | 11 | 4 | 3.13 | 0.09 |
| | Synonymous | 36 | 41 | | |
| Low-altitude haplogroup (subset of clade 2) | Nonsynonymous | 14 | 2 | 5.22 | 0.03 |
| | Synonymous | 55 | 41 | | |

Phylogenetic relationships and geographic distributions of *P. maniculatus* haplogroups are depicted in Figure 1.
*P*-values are derived from Fisher's exact test.

**Table 5.** Maximum-likelihood analysis of variation in the $d_N/d_S$ ratio (= $\omega$) across branches of species- and genus-level cytochrome *b* phylogenies.

| Model | No. of parameters | Parameter estimates | Likelihood |
|---|---|---|---|
| Lineage model: *Peromyscus maniculatus* haplogroups | | | |
| M0 | 1 | $\omega_b = 0.041$ | –3793.93 |
| M1 (free-ratio) | 88 | — | –3739.94 |
| (Background) + (high and low) | 2 | $\omega_b = 0.034$<br>$\omega_{hl} = 0.063$ | –3791.88 |
| (Background) + (high) + (low) | 3 | $\omega_b = 0.034$<br>$\omega_h = 0.070$; $\omega_l = 0.0058$ | –3791.80 |
| (Background) + (high stem) + (high crown) + (low stem) + (low crown) | 5 | $\omega_b = 0.034$<br>$\omega_{hs} = 0.119$; $\omega_{hc} = 0.064$<br>$\omega_{ls} = 1 \times 10^{-4}$; $\omega_{lc} = 0.062$ | –3790.94 |
| Lineage model: *Peromyscus* species | | | |
| M0 | 1 | $\omega = 0.028$ | –9289.81 |
| M1 (free-ratio) | 76 | — | –9220.34 |
| (Background) + (*P. maniculatus* total group) | 2 | $\omega_b = 0.028$<br>$\omega_{Pm} = 0.031$ | –9289.76 |
| (Background) + (*P. maniculatus* stem) + (*P. maniculatus* crown) | 3 | $\omega_b = 0.028$<br>$\omega_s = 2/0$<br>$\omega_c = 0.025$ | –9286.72 |

'High' and 'low' refer to the high- and low-altitude haplogroups of *P. maniculatus* (see text for details). Subscripts denote the following: b= background branches, h= high-altitude haplogroup, l= low-altitude haplogroup, s= stem, c= crown, and Pm= *P. maniculatus* clade.

allowed three separate $\omega$ ratios (one for the stem of the *P. maniculatus* clade, one for the crown and one for the background branches) provided a significantly better fit to the data than a more simple model that allowed two separate ratios (one for the entire *P. maniculatus* clade (stem+crown) and one for the background branches; $2\Delta\ell = 6.07$; $df = 1$; $P < 0.014$) (Table 5). This heterogeneity in $\omega$ ratios is attributable to the fact that the stem of the *P. maniculatus* clade was characterized by two nonsynonymous substitutions and zero synonymous substitutions. Ancestral sequence reconstruction (using the codeml program) revealed that the two substitutions occurred at residue positions 110 and 345. The substitution at residue position 110 involved replacement of a polar threonine with a hydrophobic methionine in the *bc*-loop of the matrix domain, and the substitution at position 345 involved replacement of an uncharged tyrosine with a positively charged histidine in the *H* helix of the transmembrane domain. Neither of these two substitutions is predicted to have important effects on cytochrome *b* function. Residue positions 110 and 345 are among the most variable sites in cytochrome *b* of mammals (Supplementary Figure S1), which suggests that these two sites are not subject to particularly stringent functional constraints. Sites that were identified in earlier studies as potential targets of positive selection in human cytochrome *b* (Mishmar *et al*., 2003; Ruiz-Pesini *et al*., 2004) were invariant both within *P. maniculatus* and among *Peromyscus* species.

Branch-site models revealed no evidence for positive selection in the *P. maniculatus* lineage. The model allowing $\omega > 1$ for a proportion of sites in the stem of the *P. maniculatus* clade did not identify any sites under positive selection, and had the same log-likelihood score (lnL = –9198.96) as a simplified model in which the entire cytochrome *b* gene was constrained to $\omega = 1$ in foreground and background lineages.

## Conclusion

We detected no evidence of positive selection on cytochrome *b* variation among phylogeographically distinct haplogroups of *P. maniculatus* or among different species of *Peromyscus*. Results of our McDonald–Kreitman tests revealed evidence for purifying selection against mildly deleterious amino-acid mutations in *P. maniculatus*, and results of our phylogenetic

analysis of codon-substitution patterns revealed that the evolution of cytochrome *b* in *Peromyscus* is chiefly governed by purifying selection. Our results stand in contrast to reports of positive selection on cytochrome *b* in recent human history (Mishmar *et al*., 2003; Ruiz-Pesini *et al*., 2004) and during the diversification of anthropoid primates (Andrews *et al*., 1998; Grossman *et al*., 2004). If it is true that human cytochrome *b* has undergone adaptive modifications in response to metabolic challenges associated with the colonization of temperate and sub-Arctic environments, our results demonstrate that such patterns are not universal among broadly distributed, temperate zone mammals. The evolutionary pattern described for humans is clearly not mirrored by species, such as deer mice, that inhabit an even more diverse range of thermal environments.

**Acknowledgments —** We thank S. Sabatino and C. Dingle for help in the lab, F. Hoffmann for help with data analysis, and H. Moriyama for help with protein structure modeling. F. Hoffmann, T. Miller, H. Moriyama, R. Nichols, G. Orti, A. Runk, T. Zera, and two anonymous reviewers for *Heredity* provided helpful comments on the manuscript. This work was funded by grants to JFS from the National Science Foundation (DEB-0614342), the National Institutes of Health (R01 HL087216) and the Nebraska Research council and grants to EJG from the American Society of Mammalogists and the University of Nebraska (LSIGRP, School of Biological Sciences).

## References

Andrews TD, Jermiin LS, Easteal S (1998). Accelerated evolution of cytochrome *b* in simian primates: adaptive evolution in concert with other mitochondrial proteins? J Mol Evol 47: 249–257.

Avise JC, Shapira JF, Daniel SW, Aquadro CF, Lansman RA (1983). Mitochondrial DNA differentiation during the speciation process in *Peromyscus*. Mol Biol Evol 1: 38–56.

Bielawski JP, Yang Z (2004). A maximum likelihood method for detecting functional divergence at individual codon sites, with application to gene family evolution. J Mol Evol 59: 121–132.

Bradley RD, Durish ND, Rogers DS, Miller JR, Engstrom MD, Kilpatrick CW (2007). Towards a molecular phylogeny for *Peromyscus*: evidence from mitochondrial cytochrome-*b* sequences. J Mammal 88: 1146–1159.

Carleton MD (1989). Systematics and evolution. In: Kirkland GL, Layne JN (eds). Advances in the Study of Peromyscus (Rodentia). Texas Tech University Press: Lubbock, Texas. pp 7–141.

Chappell MA, Snyder LR (1984). Biochemical and physiological correlates of deer mouse alpha-chain hemoglobin polymorphisms. Proc Natl Acad Sci USA 81: 5484–5488.

Cook SF, Hannon JP (1954). Metabolic differences between three strains of *Peromyscus maniculatus*. J Mammal 35: 553–560.

Degliesposti M, Devries S, Crimi M, Ghelli A, Patarnello T, Meyer A (1993). Mitochondrial cytochrome-*b*—evolution and structure of the protein. Biochim Biophys Acta 1143: 243–271.

Dragoo JW, Lackey JA, Moore KE, Lessa EP, Cook JA, Yates TL (2006). Phylogeography of the deer mouse (*Peromyscus maniculatus*) provides a predictive framework for research on hantaviruses. J Gen Virol 87: 1997–2003.

Elson JL, Turnbull DM, Howell N (2004). Comparative genomics and the evolution of human mitochondrial DNA: assessing the effects of selection. Am J Hum Genet 74: 229–238.

Felsenstein J (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. J Mol Evol 17: 368–376.

Fontanillas P, Depraz A, Giorgi MS, Perrin N (2005). Nonshivering thermogenesis capacity associated to mitochondrial DNA haplotypes and gender in the greater white-toothed shrew, *Crocidura russula*. Mol Ecol 14: 661–670.

Fry AJ (1999). Mildly deleterious mutations in avian mitochondrial DNA: evidence from neutrality tests. Evolution 53: 1617–1620.

Goldman N, Yang Z (1994). A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol Biol Evol 5: 725–726.

Goodman M (1982). Positive selection causes purifying selection. Nature 295: 630.

Grossman LI, Willdman DE, Schmidt TR, Goodman M (2004). Accelerated evolution of the electron transport chain in anthropoid primates. Trends Genet 20: 579–584.

Hall ER (1981). The Mammals of North America, Vols. I & II. John Wiley & Sons: New York.

Hasegawa M, Cao Y, Yang Z (1998). Preponderance of slightly deleterious polymorphism in mitochondrial DNA: nonsynonymous/synonymous rate ratio is much higher within species than between species. Mol Biol Evol 15: 1499–1505.

Hogan KM, Hedin MC, Koh HS, Davis SK, Greenbaum AF (1993). Systematic and taxonomic implications of karyotypic, electrophoretic, and mitochondrial-DNA variation in Peromyscus from the Pacific Northwest. J Mammal 74: 819–831.

Grossman LI, Wildman DE, Schmidt TR, Goodman M (2004). Accelerated evolution of the electron transport chain in anthropoid primates. Trends Genet 20: 579–584.

Jobb G, von Haeseler A, Strimmer K (2004). TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. BMC Evol Biol 4: 18.

Kivisild T, Shen P, Wall DP, Do B, Sung R, Davis K *et al*. (2006). The role of selection in the evolution of human mitochondrial genomes. Genetics 172: 373–387.

Kumar S, Tamura K, Nei M (2004). MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform 5: 150–163.

Lansman RA, Avise JC, Aquadro CF (1983). Extensive genetic variation in mitochondrial DNA's among geographic populations of the deer mouse, *Peromyscus maniculatus*. Evolution 37: 1–16.

Lemmon AR, Moriarty EC (2004). The importance of proper model assumption in Bayesian phylogenetics. Syst Biol 53: 265–277.

MacMillen RE, Garland TJ (1989). Adaptive physiology. In: Kirkland GL, Layne JN (eds). Advances in the Study of Peromyscus (Rodentia). Texas Tech University Press: Lubbock, Texas. pp 143–168.

Maddison WP, Maddison DR (1989). Interactive analysis of phylogeny and character evolution using the computer program MacClade. Folia Primatol (Basel) 53: 190–202.

McClellan DA, Palfreyman EJ, Smith MJ, Moss JL, Christensen RG, Sailsbery JK (2005). Physicochemical evolution and molecular adaptation of the cetacean and artiodactyl cytochrome *b* proteins. Mol Biol Evol 22: 437–455.

McDonald JH, Kreitman M (1991). Adaptive protein evolution at the Adh locus in *Drosophila*. Nature 351: 652–654.

Mishmar D, Ruiz-Pesini E, Golik P, Macaulay V, Clark AG, Hosseini S *et al*. (2003). Natural selection shaped regional mtDNA variation in humans. Proc Natl Acad Sci USA 100: 171–176.

Moilanen JS, Majamaa K (2003). Phylogenetic network and physicochemical properties of nonsynonymous mutations in the protein-coding genes of human mitochondrial DNA. Mol Biol Evol 20: 1195–1210.

Nachman MW (1998). Deleterious mutations in animal mitochondrial DNA. Genetica 102–103: 61–69.

Nachman MW, Brown WM, Stoneking M, Aquadro CF (1996). Nonneutral mitochondrial DNA variation in humans and chimpanzees. Genetics 142: 953–963.

Nielsen R, Yang Z (1998). Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics 148: 929–936.

Posada D, Crandall KA (1998). Modeltest: testing the model of DNA substitution. Bioinformatics 14: 817–818.

Rand DM, Kann LM (1996). Excess amino acid polymorphism in mitochondrial DNA: contrasts among genes from *Drosophila*, mice, and humans. Mol Biol Evol 13: 735–748.

Rand DM, Kann LM (1998). Mutation and selection at silent and replacement sites in the evolution of animal mitochondrial DNA. Genetica 102–103: 393–407.

Rolfe DF, Brown GC (1997). Cellular energy utilization and molecular origin of standard metabolic rate in mammals. Physiol Rev 77: 731–758.

Ronquist F, Huelsenbeck JP (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572–1574.

Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19: 2496–2497.

Ruiz-Pesini E, Mishmar D, Brandon M, Procaccio V, Wallace DC (2004). Effects of purifying and adaptive selection on regional variation in human mtDNA. Science 303: 223–226.

Storz JF 2007. Hemoglobin function and physiological adaptation to hypoxia in high-altitude mammals. J Mammal 88: 24–31.

Storz JF, Kelly JK 2008. Effects of spatially varying selection on nucleotide diversity and linkage disequilibrium: Insights from deer mouse globin genes. Genetics 180: 367–379.

Storz JF, Sabatino SJ, Hoffmann FG, Gering EJ, Moriyama H, Ferrand N *et al*. 2007. The molecular basis of high-altitude adaptation in deer mice. PLoS Genet 3: e45, 0448–0459.

Sun C, Kong Q, Zhang Y (2007). The role of climate in human mitochondrial DNA evolution: a reappraisal. Genomics 89: 338–342.

Yang Z (1997). PAML: a program package for phylogenetic analysis by maximum likelihood. CABIOS 13: 555–556.

Yang Z, Nielsen R (2000). Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. Mol Biol Evol 17: 32–43.

Yang Z, Nielsen R (2002). Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. Mol Biol Evol 19: 908–917.

Zhang J, Nielsen R, Yang Z (2005). Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Mol Biol Evol 22: 2472–2479.

Supplementary Information follows.

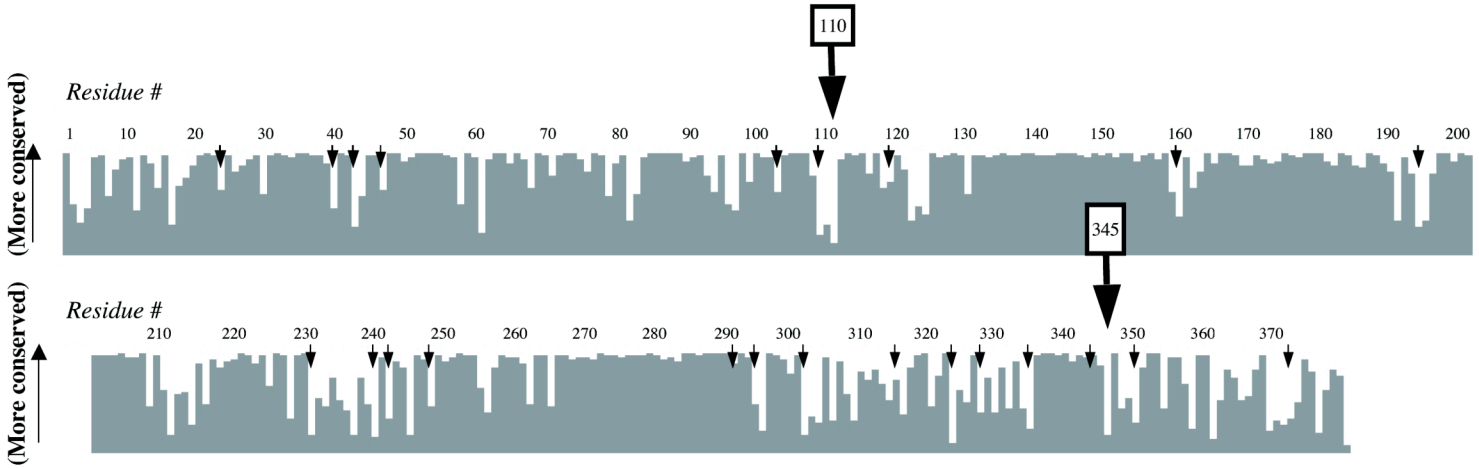


**Fig. S1**. Normalized residue conservation scores for full-length cytochrome *b* sequences from 632 mammalian species (Castresana 2001). Sequences were aligned using Clustal X (Thompson et al. 1994). Higher scores (grey bars) are indicative of a higher degree of conservation. Locations of the substitutions occurring at the stem of the *P. maniculatus* species complex (indicated by large arrows) and within *P. maniculatus* (indicated by smaller arrows) are among the most variable sites in the alignment of mammalian cyt*b* sequences.

References

Castresana J (2001). Cytochrome b phylogeny and the taxonomy of great apes and mammals. *Mol Biol Evol* 18: 465–471.

Thompson JD, Higgins DG, Gibson TJ (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignments through sequence weighting, position specific gap penalties and weight matrix choice. *Nucl Acids Res* 22: 4673–4680.

Table S1. Collection localities and Genbank accession numbers for cytochrome *b* sequences used for species-level analyses of *Peromyscus maniculatus*. Phylogenetic relationships and geographic distributions of clades are shown in Figure 1.

| LOCALITY | CLADE(S) | GENBANK ACCESSION # |
|---|---|---|
| Marin Co., CA | 1 | EU006766, EU006767, EU006768, EU006769, EU006770, EU006771, EU00672 |
| Monterey Co., CA | 3 | EF666142 |
| Merced Co., CA | 1,3 | EF666144, EF666145, EF666146, EF666147, EF666148, EF666149, EF666150, EF666151, EF666152, EF666153 |
| Tuolumne Co., CA | 1,3 | EF666154, EF666155, EF666156, EF666157, EF666158, EF666159, EF666160, EF666161, EF666162, EF666163, EF666164, EF666165, EF666166, EF666167, EF666168, EF666169, EF666170, EF666171, EF666172, EF666173, EF666174, EF666175, EF666176, EF666177, EF666178 |
| Santa Clara Co., CA | 3 | EF666143 |
| Mono Co., CA | 1,3 | EF666179, EF666180, EF666181, EF666182, EF666183, EF666184 |
| White Pine Co., NV | 1 | EF666185, EF666186, EF666187, EF666188, EF666189, EF666190, EF666191, EF666192, EF666193, EF666194, EF666195, EF666196 |
| Clear Creek Co., CO | 1 | EF666227, EF666228, EF666229, EF666230, EF666231, EF666232, EF666233, EF666234, EF666235, EF666236, EF666237, EF666238, EF666239, EF666240, EF666241 |
| Yuma Co., CO | 1,2 | EF666212, EF666213, EF666214, EF666215, EF666216, EF666217, EF666218, EF666219, EF666220, EF666221, EF666222, EF666223, EF666224, EF666225, EF666226 |
| Boulder Co., CO | 1 | EF666255, EF666256, EF666257, EF666258, EF666270, EF666271, EF666272, EF666273, EF666274, EF666275, EF666276, EF666277 |
| Pawnee Co., KS | 2 | EF666197, EF666198, EF666199, EF666200, EF666201, EF666202, EF666203, EF666203, EF666204, EF666205, EF666206, EF666207, EF666208, EF666209, EF666210, EF666211 |
| Weld Co., CO | 1 | EF666242, EF666243, EF666244, EF666245, EF666246, EF666247, EF666248, EF666249, EF666250, EF666251, EF666252 |
| British Columbia, | 1 | DQ385698, DQ385699, DQ385700, DQ385701, DQ385702 |

| | | |
|---|---|---|
| Canada | | |
| Labrador, Canada | 5 | DQ385809, DQ385810, DQ385811, DQ385812, DQ385813, DQ385814, DQ385815 |
| Manitoba, Canada | 6 | DQ385723, DQ385724, DQ385725, DQ385726, DQ385727, DQ385728, DQ385729, DQ385730, DQ385731, DQ385732, DQ385733 |
| Ontario, Canada | 5 | DQ385734, DQ385735, DQ385770, DQ385771, DQ385772, DQ385773, DQ385774, DQ385775, DQ385776, DQ385777, DQ385778, DQ385779, DQ385780, DQ385781, DQ385782 |
| Quebec, Canada | 5 | DQ385783, DQ385784, DQ385785, DQ385805, DQ385806, DQ385807, DQ385808, DQ385749, DQ385750, DQ385751, DQ385752, DQ385753, DQ385800, DQ385801, DQ385802, DQ385803, DQ385804, DQ385795, DQ385796, DQ385797, DQ385798, DQ385799, DQ385792, DQ385793, DQ385794, DQ385789, DQ385790, DQ385791, DQ385786, DQ385787, DQ385788 |
| Baja CA, Mexico | 3 | DQ385706, DQ385707, DQ385708, DQ385709 |
| Humboldt Co., CA | 1 | DQ385703, DQ385704, DQ385705 |
| Santa Barbara Co., CA | 3 | DQ385711, DQ385712, DQ385710, DQ385713, DQ385714 |
| Ada Co., ID | 1 | DQ385672, DQ385673, DQ385674, DQ385675 |
| Des Moines Co., IA | 2 | DQ385633, DQ385634, DQ385635, DQ385636, DQ385637 |
| Penobscot Co., ME | 6 | DQ385754, DQ385755, DQ385756 |
| Cass Co., MN | 6 | DQ385736, DQ385737, DQ385738, DQ385739 |
| Keweenaw Co., MI | 1 | DQ385643, DQ385644, DQ385645 |
| Mackinac Co., MI | 5 | DQ385821, DQ385822, DQ385823, DQ385824, DQ385825, DQ385826, DQ385827 |
| Gallatin Co., MT | 1 | DQ385676, DQ385677, DQ385678, DQ385679 |

| | | |
|---|---|---|
| Churchill Co., NV | 1 | DQ385715, DQ385680, DQ385681, DQ385682 |
| Hidalgo Co., NM | 2,4 | DQ385718, DQ385719, DQ385720, DQ385721 |
| Bernalillo Co., NM | 2 | DQ385753, DQ385754, DQ358655, DQ385756 |
| San Juan Co., NM | 2 | DQ385667, DQ385668, DQ385669, DQ385670, DQ385671 |
| Catron Co., NM | 2 | DQ385651, DQ385652 |
| Otero Co., NM | 2,4 | DQ385658, DQ385659, DQ385660, DQ385661, DQ385662, DQ385663, DQ385664, DQ385665, DQ385666, DQ385722 |
| Roosevelt Co., NM | 2,4 | DQ385628, DQ385629, DQ385630, DQ385631, DQ385632 |
| Sullivan Co., NY | 6 | DQ385740, DQ385741, DQ385742 |
| Clinton Co., NY | 6 | DQ385746, DQ385747, DQ385748 |
| Somerset Co., PA | 6 | DQ385762, DQ385763, DQ385764, DQ385765, DQ385766 |
| Sevier Co., TN | 5 | DQ385816, DQ385817, DQ385818 |
| Unicoi Co., TN | 6 | DQ385767, DQ385768, DQ385769, DQ385819, DQ385820 |
| Salt Lake Co., UT | 1 | DQ385684, DQ385685, DQ385686, DQ385687, DQ385688 |
| Madison Co., VA | 6 | DQ385757, DQ385758, DQ385759, DQ385760, DQ385761 |
| King Co., WA | 1 | DQ385696 |
| Kittitas Co., WA | 1 | DQ385689 |
| Pierce Co., WA | 1 | DQ385697, DQ385716 |
| Snohomish Co., WA | 1 | DQ385695 |
| Kitsap Co., WA | 1 | DQ385694 |
| Cochise Co., AZ | *Peromyscus melanotis* | EU574689-EU574701 |
| Leon Co., FL | *Peromyscus gossypinus* | DQ385624, DQ385225 |

Table S2. Species identities and Genbank accession numbers for cytochrome *b* sequences used in genus-level analyses of *Peromyscus*, along with information about altitudinal distributions. "Low elevation" refers to species generally found below 1000m, "High elevation" refers to species generally found above 2500m, "Intermediate elevation" refers to species found between 1000m and 2500m.

| Species ID | Genbank ID | Distribution* | References |
|---|---|---|---|
| *P. beatae* | AF131914<br>AF131921 | **Intermediate to high elevation**<br>Humid forests of Mexico and Central America | 1 |
| *P. levipes* | AF155396<br>AF131928 | **Intermediate to high elevation**<br>Central America and the Sierra Madre of Mexico | 1,2 |
| *P. schmidlyi* | AY322522 | **High elevation**<br>Pine-oak forests of Durango, Mexico | 3 |
| *P. boylii* | AF131915<br>AF155388<br>AF155387<br>AF155392 | **High elevation**<br>Chaparral, pinyon-juniper woodlands and pine-oak forests of the western US and Sierra Madre Occidental, Mexico | 4,5 |
| *P. simulus* | AAG25892 | **Low elevation**<br>Coastal lowlands, plains, and river valleys of Sinaloa and Nayarit, Mexico | 1,6 |
| *P. stephani* | AAD43126 | **Low elevation**<br>Desert islands in the sea of Cortez, Mexico | 1,2 |
| *P. spiceligus* | AAQ83477 | **Low to intermediate elevation**<br>Sierra Madre Occidental, West-central Mexico | 1,7 |
| *P. winkelmanni* | AAG25895 | **High elevation**<br>Montane oak and pine forests of west-central Mexico (Guererro and Michoacán) | 1,8 |
| *P. sagax* | AAD43119 | **Intermediate elevation** | 2 |
| *P. truei (comanche, montipinoris, truei)* | AF155412<br>AY376434<br>AY376431<br>AY376432 | **Low to high elevation**<br>Central Oregon to southern Mexico, California to western Texas. (*P.t. comanche:* Texas, *montipinoris:* southcentral California, *truei:* southwestern US) | 1,5,9 |

| | | | |
|---|---|---|---|
| *P. gratus* | AAS45263 | **High elevation**<br>Mountains of southwestern New Mexico and central Mexico | 5 |
| *P. difficilus* | AAS45259 | **High elevation**<br>Rocky forested areas, Colorado, US to Oaxaca, Mexico | 1, 2 |
| *P. nasutus* | AAS45265 | **Intermediate elevation**<br>Rocky outcrops and among boulders in pinyon-juniper-oak woodland in New Mexico, eastern Arizona, western Texas, central Colorado and Northern Mexico | 1,5 |
| *P. attwateri* | AAD43100 | **Low elevation**<br>Rocky habitats in central Texas, Oklahoma, and Arkansas, US | 1,5,10 |
| *P. pectoralis* | AAS45266 | **Intermediate elevation**<br>Rocky terrain in the Central Plateau and Sierra Madre of Mexico, comparable areas in Texas, New Mexico, and Oklahoma | 1,11 |
| *P. melanophries* | AAQ83475 | **Intermediate elevation**<br>Dry rocky habitats within the Mexican Plateau, southern Durango and Coahuila through Oaxaca | 1,12 |
| *P. zarhynchus* | AAP34316 | **Intermediate to high elevation**<br>cloud forests of Chiapas, Mexico | 13 |
| *P. maniculatus (gambelli, coolidgei, rufinus, sonoriensis, blandus)* | EF666175<br>DQ385707<br>DQ385714<br>EF666232<br>DQ785703<br>DQ385721 | **Low to high elevation**<br>Across most of North America into central Mexico; see text of this manuscript for further discussion and references | 1,4,5 |
| *P. keeni* | DQ385716 | **Low to high elevation**<br>Coastal lowlands to sub-alpine forests, Pacific Northwestern US and British Columbia | 5 |
| *P. melanotis* | AF155398 | **High elevation**<br>Pine–fir forest and grasslands, central Mexico and southeastern Arizona | 1 |
| *P. leucopus* | AF131926 | **Low to high elevation**<br>Mature forests, and marshes, eastern US and eastern Mexico | 1, 14 |
| *P. californicus* | AAD43108 | **Low to intermediate elevations**<br>Foothills of coastal Californian ranges and the Sierra Nevada | 1,5,15 |

| | | | |
|---|---|---|---|
| *P. eremicus* | AAQ83468 | **Low to intermediate elevations**<br>Desert shrub and riparian habitats in Northern Mexico, Baja and Southwestern US | 1,5,16 |
| *P. crinitus* | AAS45253 | **Low to high elevation**<br>Canyons in western US and Northern Baja, Mexico | 1,17 |
| *P. polius* | AF155403 | **High elevation**<br>West-central Chihuahua, México | 2 |
| *Reithrodontomys raviventris* | AY859470 | **Low elevation**<br>Salt Marshes of the San Francisco Bay Area, US | 18 |
| *R. tenuirostris* | AY859463 | **High elevation**<br>Evergreen Cloud forest in Chiapas, Mexico and Guatemala | 19 |

References


[1] Carleton MD (1989). Systematics and evolution. In: Kirkland GL and Layne JN (eds). *Advances in the study of* Peromyscus *(Rodentia)*, Texas Tech University Press: Lubbock, Texas. pp 7–141.

[2] Wilson DE, and Reeder DM (eds). (2005). *Mammal Species of the World. A Taxonomic and Geographic Reference* (3rd ed). Smithsonian Institution Press, Washington, D.C. 2,142 pp.

[3] Bradley RD, Carroll DS, Haynie ML, Muñiz Martínez R, Hamilton MJ, and Kilpatrick CW (2004). A new species of *Peromyscus* from western Mexico. *J Mammal* **85**: 1184–1193.

[4] Hall, ER (1981) *Mammals of North America* (vol. II), John Wiley & Sons, New York. pp 601-1181.

[5] Wilson DE and Ruff S (1999). *The Smithsonian book of North American mammals*. Smithsonian Institution Press, Washington DC. 776 pp.

[6] Roberts JR, Schmidly DJ, and Bradley RD (2001). *Peromyscus simulus*. *Mammalian Species* **669**:1-3.

[7] Roberts HR, Schmidly DJ, and Bradley RD (1998). *Peromyscus spicilegus*. *Mammalian Species* **596**:1-4.

[8] Álvarez-Castañeda ST (2005). *Peromyscus winkelmanni*. *Mammalian Species* ***765***:1-3.

[9] Hoffmeister DF (1981). *Peromyscus truei*. *Mammalian Species* **161**:1-5.

[10] Schmidly DJ (1974). *Peromyscus attwateri*. *Mammalian Species* **48**:1-3.

[11] Schmidly DJ (1974). *Peromyscus pectoralis*. *Mammalian Species* **49**:1-3.

[12] Carleton MD and Lawlor TE (2005). *Peromyscus* from Santa Catalina Island, Sea of Cortez, Mexico: taxonomic identities and biogeographic implications. *J Mammal* **86**: 814-825.

[13] McClellan DA and Rogers DS (1997). *Peromyscus zarhynchus*. *Mammalian Species* **562**:1-3.

[14] Lackey JA, Huckaby DG, and Ormiston BG (1985). *Peromyscus leucopus*. *Mammalian Species* **247**:1-10.

[15] Merritt JF (1978). *Peromyscus californicus*. *Mammalian Species* **85**:1-6.

[16] Veal R and Caire W (1979). *Peromyscus eremicus*. *Mammalian Species* **118**:1-6.

[17] Johnson DW and Armstrong DM (1987). *Peromyscus crinitus*. *Mammalian Species* **287**:1-8.

[18] Shellhammer, HS (1982). *Reithrodontomys raviventris*. *Mammalian Species*. 169:1-3

[19] Arellano, E. & Rogers, DS (1994). *Reithrodontomys tenuirostris*. *Mammalian Species*. 477:1-3.