TOWARD A MOLECULAR PHYLOGENY FOR *PEROMYSCUS*: EVIDENCE FROM MITOCHONDRIAL CYTOCHROME-\(b\) SEQUENCES

ROBERT D. BRADLEY,* NEVIN D. DURISH, DUKE S. ROGERS, JACQUELINE R. MILLER, MARK D. ENGSTROM, AND C. WILLIAM KILPATRICK

Department of Biological Sciences, Texas Tech University, Lubbock, TX 79409-3131, USA (RDB, NDD)
Museum of Texas Tech University, Lubbock, TX 79409-3191, USA (RDB)
Department of Integrative Biology and Monte L. Bean Life Science Museum, Brigham Young University, Provo, UT 84602, USA (DSR)
Department of Zoology, University of Toronto, Toronto, Ontario M5S 3G5, Canada (JRM, MDE)
Department of Biology, University of Vermont, Burlington, VT 05405-0086, USA (CWK)

One hundred DNA sequences from the mitochondrial cytochrome-\(b\) gene of 44 species of deer mice (*Peromyscus* (sensu stricto), 1 of *Habromys*, 1 of *Isthmomys*, 2 of *Megadontomys*, and the monotypic genera *Neotomodon*, *Osgoodomys*, and *Podomys* were used to develop a molecular phylogeny for *Peromyscus*. Phylogenetic analyses (maximum parsimony, maximum likelihood, and Bayesian inference) were conducted to evaluate alternative hypotheses concerning taxonomic arrangements (sensu stricto versus sensu lato) of the genus. In all analyses, monophyletic clades were obtained that corresponded to species groups proposed by previous authors; however, relationships among species groups generally were poorly resolved. The concept of the genus *Peromyscus* based on molecular data differed significantly from the most current taxonomic arrangement. Maximum-likelihood and Bayesian trees depicted strong support for a clade placing *Habromys*, *Megadontomys*, *Neotomodon*, *Osgoodomys*, and *Podomys* within *Peromyscus*. If *Habromys*, *Megadontomys*, *Neotomodon*, *Osgoodomys*, and *Podomys* are regarded as genera, then several species groups within *Peromyscus* (sensu stricto) should be elevated to generic rank. *Isthmomys* was associated with the genus *Reithrodontomys*; in turn this clade was sister to *Baiomys*, indicating a distant relationship of *Isthmomys* to *Peromyscus*. A formal taxonomic revision awaits synthesis of additional sequence data from nuclear markers together with inclusion of available allozymic and karyotypic data.

Key words: cytochrome-\(b\) gene, *Peromyscus*, phylogeny, species group, systematics

et al. 2002). In support of the importance of *Peromyscus* to a variety of scientific disciplines, Dewey and Dawson (2001) referred to *Peromyscus* as “the *Drosophila* of North American mammalogy.” Despite, this significance and associated abundance of information, surprisingly little is known about the phylogenetic relationships within this genus. Given the importance of members of the genus *Peromyscus* as model organisms and status of the genus as one of North America’s most common and diverse groups of rodents, recovering the phylogenetic history of *Peromyscus* is crucial for many scientific disciplines.

In developing a phylogenetic hypothesis for *Peromyscus*, 2 issues must be considered and reviewed concerning taxonomic history and organization. First, the most basic question is, what is *Peromyscus*? Through several revisions and classifications (Carleton 1980, 1989; Hooper 1968; Hooper and Musser 1964; Musser and Carleton 2005; Osgood 1909), recognition of genera and subgenera has influenced the concept of *Peromyscus*. Second, are the species groups established by Osgood (1909) valid taxonomic entities?

Osgood’s (1909) revision of *Peromyscus* included 6 subgenera (*Baioniomys*, *Haplomylops*, *Megalontomys*, *Ochrotomys*, *Peromyscus*, and *Podomys*) in *Peromyscus*. Subsequently, Miller (1912) and Hooper (1958) assigned 2 of these (*Baioniomys* and *Ochrotomys*) to generic status. Three additional subgenera (*Habromys*, *Isthmomys*, and *Osgoodomys*) were proposed by Hooper and Musser (1964). Hooper (1968) therefore included 7 subgenera in his classification. However, Carleton (1980) contended that 5 of these subgenera (*Habromys*, *Isthmomys*, *Megalontomys*, *Osgoodomys*, and *Podomys*) were not as closely related to the subgenera *Peromyscus* and *Haplomylops* than were some forms of *Reithrodontomys*.

Consequently, Carleton (1980) elevated these taxa to generic rank leaving *Peromyscus* and *Haplomylops* as the remaining subgenera in *Peromyscus*. Carleton (1989) subsequently reviewed the genus and subsumed *Haplomylops* within *Peromyscus* (hereafter referred to as *Peromyscus* sensu stricto), a treatment that remains the most comprehensive review of *Peromyscus* and allies to date and has served as the basis for the 2 most recent classifications (Musser and Carleton 1993, 2005) of the genus.

Two other genera, *Neotomodon* and *Nelsonia*, historically have been affiliated with *Peromyscus*. Osgood (1909) appears to have accepted Merriam’s (1898) suggestion that *Neotomodon* was aligned with *Neotoma*. However, Goldman (1910), Davis and Follansbee (1945), Hooper and Musser (1964), and Hooper (1968) hypothesized that *Neotomodon* was a close relative of *Peromyscus*, and Yates et al. (1979) and Rogers et al. (2005) presented data that suggested that it may be a congener of *Peromyscus*. Conversely, Carleton (1980, 1989) and Musser and Carleton (2005) hypothesized that *Neotomodon* was aligned closely with *Podomys* and *Habromys*. *Nelsonia* largely has been ignored in most revisions because of its scarcity in museum collections. Carleton (1980) and Musser and Carleton (2005) placed *Nelsonia* as a separate genus closely related to other neotomines (*Hodomys* and *Neotoma*); however, examination of primitive morphological characters (Hooper and Musser 1964) and karyotypic data (Engstrom and Bickham 1983, Engstrom et al. 1992) suggested an affinity to peromyscines.

Several molecular studies have generated data used to infer phylogenetic relationships among *Peromyscus* and allied genera. Cladistic analyses of protein electrophoretic data (Patton et al. 1981) and chromosomal banding data (Rogers 1983; Rogers et al. 1984; Stangl and Baker 1984) were interpreted as supportive of some aspects of the hypothesis of Hooper and Musser (1964). A DNA–DNA hybridization study by Dickerman (1992) supported the phylogeny of Hooper and Musser (1964) because *Podomys* was contained within the *Peromyscus* cluster, and *Reithrodontomys* was sister to a cluster containing *Podomys*, *Peromyscus*, and *Onychomys*. Similarly, Bradley et al. (2004), Reeder and Bradley (2004, 2007), and Reeder et al. (2006) commented that in some cases *Megalontomys*, *Neotomodon*, and *Osgoodomys* were embedded within *Peromyscus* (sensu stricto). However, Engel et al. (1998) supported the contention of Carleton (1980) that some forms of *Peromyscus* were more closely related to *Reithrodontomys* than they were to *Osgoodomys* and *Isthmomys*. To date, only 1 study (Rogers et al. 2005) was designed specifically to discern among the alternative positions of Hooper (1968) and Carleton (1980). Their phylogenetic analyses of protein electrophoretic data placed all subgenera that Carleton (1980) had elevated to generic rank (*Habromys*, *Megalontomys*, *Osgoodomys*, and *Podomys*) together with *Neotomodon* into a single clade that, with the exception of *Isthmomys*, was identical to the concept of *Peromyscus* as expressed by Hooper (1968).

In addition to providing the 1st revision of *Peromyscus*, Osgood (1909) placed related species into species groups. Osgood’s (1909) recognition of 7 species groups within the subgenus *Peromyscus* (*boylei [= boylii], lepturus, leucopus, maniculatus, megalops, melanophys, and truei*) was conservative because relatively few taxa were described at that time. As more taxa were described, a need arose for additional species groups to partition the discrete morphological variation evident among all taxa and to eliminate paraphyletic assemblages. Hall and Kelson (1959) added the *mexicanus* group; Hooper (1968) recognized the *crinitus* group and placed the *lepturus* group under the subgenus *Habromys*; and Carleton (1989) added the *aztecus, hooperi,* and *furvis* groups and recognized the *californicus* and *eremicus* groups in the subgenus *Haplomylops*. A critical evaluation of phylogenetic relationships within and among species groups is paramount to continued usage of this nomenclatural delineation.

et al. 2004; Engel et al. 1998; Hogan et al. 1997; Sullivan et al. 1995, 1997; Tiemann-Boege et al. 2000), and combinations of these techniques (Bradley et al. 1996b; Hogan et al. 1993). However, most of these studies did not include an adequate sample of taxa to allow construction of a phylogeny of recognized genera and subgenera.

The objective of our study was to develop a molecular phylogenetic hypothesis, based on DNA sequences from the mitochondrial cytochrome-\(b\) (\textit{Cytb}) gene, for species of \textit{Peromyscus}\ (sensu lato—Hooper and Musser 1964) and use that phylogeny to begin to evaluate the monophyly and validity of species groups, subgenera, and genera. We used the concept of Hooper and Musser (1964) of \textit{Peromyscus} as a null hypothesis because it is inclusive and readily refutable based on nonmonophyly. Also, we recognize from the onset that the \textit{Cytb} gene is most useful in recovering phylogenetic relationships among closely related taxa and can lose resolution at deeper nodes. We include a majority of species putatively assigned to this group (51 of 72, or 70%), comprising a sample that represents all genera, subgenera, and species groups, establishing this study as a foundation for a revised and more thoroughly inclusive classification of \textit{Peromyscus}. We propose that our study be viewed as a foundation upon which inclusion of additional taxa and DNA sequences can be added, the ultimate goal being to estimate a robust phylogeny for the genus. In the words of Emmet T. Hooper (1968:33–34), “A taxonomic classification is a tentative thing; it is not sacred. . . . It is a reasonable yet tentative framework against which new systematic data may be tried for fit.” It is in this spirit that we tackle the concept of “what is \textit{Peromyscus}?"

**Materials and Methods**

**Samples.**—Kimura 2-parameter genetic distances (Kimura 1980) and the neighbor-joining method of Saitou and Nei (1987) were used in a preliminary analysis of 189 \textit{Cytb} sequences, generated in our laboratories or obtained from GenBank. In several instances, \(\geq 10\) DNA sequences were available for examination of a single species; consequently 100 individuals representing 44 species of \textit{Peromyscus} (sensu Carleton 1989), 1 species of \textit{Habronomys}, 1 species of \textit{Isthmomys}, 2 species of \textit{Megadontomys}, and the monotypic genera \textit{Neotomodon}, \textit{Osgoodomys}, and \textit{Podomys} were selected for further analysis. The rationale for the reduced data set was to examine the most genetically divergent members within each species (determined from preliminary analyses), include individuals representing geographic extremes within a species distribution, and insure that all individuals of a species were monophyletic, while maintaining the data set at a manageable size. This was accomplished by including a maximum of 3 individuals per species. If possible, individuals with complete DNA sequences were included. Specimen numbers, GenBank accession numbers, collection localities, and specimen voucher numbers are listed in Appendix I.

**Sequence data.**—Mitochondrial DNA was extracted from liver samples (0.1 g), either frozen or preserved in 95% ethanol and purified using the Wizard Miniprep kit (Promega, Madison, Wisconsin) or the DNeasy Tissue Kit (Qiagen, Valencia, California). For some specimens, genomic DNA was isolated following methods of Smith and Patton (1999). Polymerase chain reaction (Saiki et al. 1988) parameters and primers, summarized in Durish et al. (2004) and Rogers et al. (2007), were used to amplify the complete \textit{Cytb} gene (1,143 base pairs [bp]). The resulting polymerase chain reaction product was purified using the QIAquick polymerase chain reaction purification kit (Qiagen) and cycle sequencing was conducted using the ABI Big Dye version 3.0 ready reaction mix (PE Applied Biosystems, Foster City, California). Excess dye terminator was removed using Millipore Multiscreen Filter Plates for High Throughput Separations (Millipore, Billerica, Maryland) and samples were analyzed on an ABI 3100 Avant automated sequencer (PE Applied Biosystems). Sequencher software (versions 3.1 and 4.1.1; Gene Codes Corporation, Ann Arbor, Michigan) and Vector NTI 7.0 software (Informax, Inc., Bethesda, Maryland) were used to align and proof sequences. A small number of samples were amplified and sequenced using modified protocols. Standard phenol–chloroform extractions of frozen liver were subsequently reduced and purified by alcohol precipitation (Sambrook et al. 1989). A 2-step touchdown polymerase chain reaction thermal profile was used (an initial 2–3 min of denaturing at 94°C; 40 s at 94°C, 40 s at 56°C, 1 min 45 s at 72°C, for 15–17 cycles; 40 s at 94°C, 40 s at 50°C, 1 min 45 s at 72°C for 21–26 cycles; 5-min final extension at 72°C). Where polymerase chain reaction amplification was difficult, downward adjustments of annealing and extension temperatures were made. Amplicons were electrophoresed on a 1.5% agarose gel. Resulting bands were extracted for cycle sequencing and purified by spin column reduction. The DYNAMIC Direct Cycle Sequencing kit (GE Healthcare Bio-Sciences Inc., Baie D’Urfé, Quebec, Canada) was used for cycle sequencing, and samples were analyzed on a LICOR LongReader 4200 analyzer (LI-COR Biosciences, Lincoln, Nebraska). All sequences were edited and aligned using Sequencher version 4.1.1 (Gene Codes Corporation). Sequences generated in this study were deposited into GenBank and accession numbers are reported in Appendix I.

**Data analyses.**—Based on phylogenetic relationships presented in Carleton (1980), Bradley et al. (2004), and Reeder and Bradley (2004), \textit{Oryzomys palustris, Sigmodon hispidus, Nyctomys sumichrasti, Tylomys nudicaudus}, and \textit{Ototylomys phyllotis} were used as outgroup taxa. Because of the uncertain phylogenetic relationships among the tribes \textit{Baiomiini}, \textit{Ochrotomyini}, and \textit{Reithrodontomyini}, \textit{Baiomys taylori}, \textit{Neotoma mexicana}, \textit{Onychomys arenicola}, \textit{Ochrotomys nutalli}, \textit{Reithrodontomys mexicanus}, and \textit{Reithrodontomys megalotis} all were included as reference taxa. By including these taxa, all genera previously assigned to \textit{Peromyscus} (e.g., \textit{Baiomys} and \textit{Ochrotomys}) were represented. Given spurious nucleotides (unreliable in that they did not match well with other \textit{Peromyscus} sequences) in the 3′ region of the \textit{P. polius} sequence, this sequence was truncated to include only the first 700 bp. All GenBank accession numbers for outgroup and reference samples are listed in Appendix I. Likelihood, Bayesian, and parsimony models (PAUP* [Swofford 2002]
and MRBAYES 3.1.1.1 (Huelsenbeck and Ronquist 2001) were used to generate hypotheses concerning phylogenetic relationships of taxa. Variable nucleotide positions were treated as unordered, discrete characters with 5 possible states: A, C, G, T, or missing.

Some authorities advocate removal of 3rd-position information in sequence data when saturation is suspected (e.g., Ericson and Johansson 2003; Huchon et al. 2002). However, synonymous substitutions at 3rd positions can be phylogenetically informative, even in sequences that are moderately saturated (Poux and Douzery 2004; Yoder et al. 1996; Yoder and Yang 2000). The value of this information also has been demonstrated in cases where taxa are closely related (Björklund 1999; Hästad and Björklund 1998). All models employed acknowledge base position characteristics; therefore, all codon positions were included in analyses.

Unconstrained parsimony analysis (PAUP*—Swofford 2002) was conducted, using unweighted characters and excluding uninformative characters. All parsimony analyses employed a heuristic search using tree-bisection-reconnection, with random stepwise addition of taxa and 100 repetitions to obtain the most-parsimonious tree-set. Bootstrap analysis (Felsenstein 1985) with 1,000 iterations and 50 random sequence addition replicates was used to evaluate internal stability and establish nodal support.

Transition–transversion plots (subroutines in DAMBE—Xia 2000; Xia and Xie 2001) regressed against sequence divergence values were used to assess relative degrees of saturation at each codon position. These plots suggested a moderate degree of saturation with increasing distance among taxa at 3rd positions (not shown). Bias in transitions and transversions was assessed by means of index values (transition/transversion [Ts/Tv]) calculated for the overall data matrix and at each codon position using the program MEGA (Kumar et al. 1993). Data were reanalyzed under the same search options as the unconstrained parsimony, but with application of either a priori constraints (transition–transversion weighting, codon position downweighting), or a posteriori constraints (successive re-weighting by rescaled consistency index [RC]) following recommendations in the literature (e.g., Barker and Lanyon 2000; Farias et al. 2001; Farris 1969; Honeycutt et al. 1995; Horovitz and Meyer 1995; Huchon et al. 2002; Irwin et al. 1991).

MODELTEST software (Posada and Crandall 1998) was used to determine the model of DNA evolution best fitting the data. The GTR+I+G model generated significantly better likelihood scores than all other models and included the following parameters: base frequencies (A = 0.3920, C = 0.3297, G = 0.0517, and T = 0.2266), rates of substitution (A–C = 0.4007, A–G = 6.7581, A–T = 0.6293, C–G = 0.4783, C–T = 7.5063, G–T = 1.00), proportion of invariant sites (I = 0.4008), and gamma distribution (G = 0.6155). Using the above parameters and maximum likelihood methods, an optimal tree(s) was generated using the heuristic search option in PAUP* (Swofford 2002).

A Bayesian model (MRBAYES—Huelsenbeck and Ronquist 2001) was used for comparison to the likelihood method and to generate support values (clade probabilities). A GTR+I+G model with a site-specific gamma distribution and sites partitioned by codon was used with the following options: 4 Markov chains, 20 million generations, and sample frequency of every 1,000th generation. O. palustris was designated as the outgroup and 2 runs were conducted simultaneously. After a visual inspection of likelihood scores, convergence statistics, and potential scale-reduction factors, the first 1,000 trees were discarded and a consensus tree (50% majority rule) was constructed from remaining trees. Additionally, 4 independent runs using 2 million generations and above conditions were conducted to investigate stabilization of likelihood scores.

**RESULTS**

The DNA sequences obtained in this study (n = 40) were combined with published sequences from GenBank (n = 60). Seventy-eight of 100 sequences were complete (1,143 bp). Of the remaining 22 sequences, 8 were missing fewer than 100 bp, 12 were >50–75% complete, and 2 (P. keeni and *P. polionotus*) were approximately 30% complete. Across all taxa, nucleotide frequencies estimated from PAUP* (Swofford 2002) for the complete data set were: A = 31.7%, C = 27.4%, G = 12.7%, and T = 28.2%; and for phylogenetically informative sites only: A = 36.3%, C = 35.9%, G = 6.0%, and T = 21.8%. Nucleotide change at each codon position agreed with observations for mammalian Cytb in general (Honeycutt et al. 1995; Irwin et al. 1991). Ratios for transitions to transversions were: 3.5, 2.9, and 1.9 at 1st, 2nd, and 3rd positions, respectively, and 2.1 overall. For the entire gene sequence, 514 nucleotides were conserved, 629 were variable, and 509 were phylogenetically informative.

The majority of substitutions occurred at 3rd positions (58%), with 66% being transitions. Nucleotide substitutions at 1st positions accounted for the majority of remaining replacements (26%), of which most were transitions (79%). In addition, there was a slight bias toward an A–T-rich nucleotide composition (28.2% and 31.7%, respectively), accounting for approximately 60% of all substitutions. A large proportion of substitutions represented synonymous changes reflecting a modest amino acid proportion of leucine (12–15% depending on taxon). Nonsynonymous substitutions occurred relatively frequently (187 of 381 variable amino acids, of which 108 were phylogenetically informative).

Unconstrained parsimony analysis of 509 unweighted characters (informative characters) generated 4 equally parsimonious trees of 5,132 steps, with a consistency index (C.I) of 0.1775, homoplasy index (HI) of 0.8225, and retention index (RI) of 0.5705 (Table 1). A bootstrap (majority-rule) consensus tree generated from the 4 most-parsimonious trees is depicted in Fig. 1. All species represented with multiple samples formed monophyletic clades, with 1 exception; 1 sample of *P. mexicanus* was sister to *P. gymnotis*. Bootstrap values showed support for a large clade uniting members of the tribes Baiomyini (*Baiomys*), Neotomini (*Neotoma*), and Peromyscini (*Peromyscus* sensu lato), *Onychomys*, *Ochrotomys*, and *Reithrodontomys*; however, no support was present for a group
that would represent a monophyletic *Peromyscus* (either sensu lato or sensu stricto). Six smaller clades united members that generally represented the *boyllii*, *aztecus*, *mexicanus*, *melanophrys*, *leucopus*, and *maniculatus* species groups. Support for relationships between species group was present only for the *leucopus* and *maniculatus* species groups. Relationships among peromyscine–neotomine taxa were unresolved. In addition, a dynamically weighted analysis using the rescaled consistency index (RC) was conducted (Farris 1969) and the unconstrained, most-parsimonious tree was set as the starting iteration. Iterations were repeated until tree topology and tree statistics stabilized (3 iterations). A strict consensus of the 2 equally parsimonious trees maintained the integrity of the larger species groups, established under the unconstrained model (not shown). Although tree lengths were not readily comparable, overall tree statistics indicated a more robust evolutionary hypothesis.

Examination of transitions and transversions by codon position, both empirically and by bivariate regression against branch length under one of the models of evolution with fewest specified parameters (Jukes and Cantor 1969), indicated moderate saturation in transitions at the 3rd position. Therefore, 5 additional analyses were conducted to downweight or correct for transitional biases. First, transitions were downweighted by a factor of 2 relative to transversions (Ts/Tv = 2) to account for the overall 2:1 ratio of transitions to transversions. Second, each codon position was assessed for independent transition–transversion biasing and dynamic weighting according to RC values converged upon the general phylogenetic hypothesis recovered with likelihood and Bayesian analyses (Fig. 2).

The likelihood analysis (GTR+I+G model) generated 2 trees (not shown) that were identical in topology except for placement of *P. ochraventer*. In both trees, *P. ochraventer* was affiliated with members of the *truei* species group but differed with respect to sister relationships. All taxa of *Peromyscus* (sensu lato) were contained within a single clade; however, the 2 samples of *Reithrodontomys* were included in this clade as a sister taxon to *P. polius*. As in the parsimony analysis, all conspecific taxa formed monophyletic clades. In addition, composition of species groups was similar to that obtained in parsimony and Bayesian analyses (below), although branching patterns within clades occasionally differed. No support values were calculated because of computational constraints.

Two simultaneous Bayesian analyses generated identical trees (Fig. 2). Clade probability values differed slightly between analyses; therefore if 2 support values differed at a node, both were reported. All taxa of *Peromyscus* (sensu lato) were contained within a single well-supported clade (I, value = 98) except for samples of *I. pirreensis* and *P. polius*. *I. pirreensis* was sister to *R. megalotis* but not supported by significant clade probability values (79 and 81). The sample of *P. polius* was sister to *Peromyscus* (sensu stricto) but not supported by a significant clade probability value (89). Taxa representing species groups formed well-supported monophyletic clades except for members of the *truei* and *mexicanus* groups. *P. sagax* was strongly supported as sister to samples of *H. ixtilani*. Three genera (*Hubromys* plus *P. sagax*, *Neotomodon*, and *Podomys*) recognized by Carleton (1980) and Musser and Carleton (2005) were contained within a well-supported clade along with members of the *aztecus*, *boyllii*, *difficilis*, and *furvus* species groups. Similarly, *Osgoodomys* was placed within a well-supported clade containing members of the *californicus*, *eremicus*, *leucopus*, *maniculatus*, *crinitus*, and *hooperi* species.
groups. *Megadontomys* was placed in the large *Peromyscus* (sensu lato) clade but its position was unresolved.

**DISCUSSION**

Bootstrap support and clade probability values indicated strong support (in most cases) for clades located terminally, minimal support for clades in the midregion of topologies, and strong support (for a few nodes) at the base of topologies. Therefore, resolution and phylogenetic inference varied across the 3 analyses. The pattern of support values is not surprising given caveats and limitations associated with using a rapidly evolving gene such as *Cytb* to address relationships among genera and subgenera. Although the Cytb gene has proven useful in recovering phylogenetically useful information at a variety of taxonomic levels, strength of its utility can be lineage-dependent and declines with evolutionary depth (Gissi et al. 2000; Springer et al. 2001; Yoder et al. 1996; Zardoya
At the outset we assumed that support at basal nodes would be ambiguous. This was the case for the parsimony analysis (Fig. 1), where only a single basal node possessed substantial bootstrap support. However, several basal nodes of the Bayesian tree (Fig. 2) had significant clade probability values. Below we summarize phylogenetic relationships that were consistent across analyses, identify groups of taxa that received support (bootstrap or clade probability values), and comment on status and composition of species groups and genera recognized by previous authors.

**Fig. 2.**—Maximum-likelihood tree calculated using Bayesian inference methods. Roman numerals refer to major clades as identified in the text. Clade probability values (depicted as percentages) are shown above branches (values to the left of slash are from the 1st run, those to the right are from the 2nd run). Asterisks indicate significant nodal support (95 or greater). Names to the right of dotted lines indicate species group or generic affiliation. Indeterminate species groups (ISG) as defined by Musser and Carleton (2005).
In all analyses (maximum parsimony, maximum likelihood, and Bayesian inference), multiple samples representing conspecific taxa formed monophyletic clades with the exception of *P. mexicanus*. One sample of *P. mexicanus* was sister to the clade containing the 2 samples of *P. gymnотis*, whereas the 2nd sample of *P. mexicanus* was basal to that clade. Even if *P. gymnотis* is not a valid taxon, the amount of genetic divergence among these samples (8.3% Kimura 2-paramenter) warrants additional investigation (Baker and Bradley 2006).

**Phylogenetic relationships among species.**—Although we were unable to include all taxa in *Peromyscus* (sensu lato), approximately 70% of the species were examined. The majority of unsampled taxa belong to the *eremicus* species group and the genus *Habromys*, where we examined only 1 of 10 species and 2 of 6 species, respectively. In addition, we were unable to include *P. sejegis* (maniculatus group), *P. mекiистrus* (melanophrys group), *P. granдis* and *P. yucatanicus* (mexicanus group), *P. bullatus* (truei group), *P. slevинi* (indeterminate group), *P. t. flavидas*, and *M. nelsoni*. Given that most taxonomic groups (genera, subgenera, and species groups) were well sampled, the phylogeny depicted in Fig. 2 can be used as a reasonable hypothesis for estimating phylogenetic relationships. To avoid a detailed description of phylogenetic relationships, we refer to Fig. 2 and clades labeled with Roman numerals.

**Species groups.**—Osgood’s (1909) original concept of species groups within the subgenus *Peromyscus* formed the basis and provided a model for subsequent classifications as reflected in the addition of 7 species groups (aztecus, californicus, crinitus, eremicus, furvus, hooperi, and mexicanus) by Hall and Kelso (1959), Hooper (1968), and Carleton (1989). After the elevation of *Habromys* to the generic level (Carleton 1989) and the concomitant removal of the *lepturus* species group, Musser and Carleton (2005) recognized 13 species groups within *Peromyscus* (sensu stricto). In addition, Musser and Carleton (2005) placed 6 taxa (*melanocarpus*, *mayensis*, *ochraventer*, *perotaralis*, *sagax*, and *slevini*) as incertae sedis.

Given that some species groups are monotypic (californicus, crinitus, and hooperi) and others are somewhat underrepresented in our data set (eremicus and megalops), it is relatively difficult to assess their validity. However, if one compares the overall topology, branching patterns, and equivalency of clades, a pattern emerges (Fig. 2) that offers support for recognition of the 13 species groups identified by Musser and Carleton (2005). With the exception of the furvus, mexicanus, and truei species groups, all other species groups with 2 or more taxa are monophyletic and are well supported (clade probability and bootstrap values). The truei species group was split into 2 clades, the difficilis assemblage represented by attwateri, difficilis, nasutus, and ochraventer and the truei assemblage including gratus and truei. Durish et al. (2004) noted this dichotomy and suggested that the 2 assemblages may represent separate species groups.

Four of the 5 taxa (*melanocarpus*, *ochraventer*, *perotaralis*, *polius*, and *slevini*) placed as incertae sedis by Musser and Carleton (2005) also were problematic in our analyses. Only the placement of *P. ochraventer*, which received support for being included in the difficilis assemblage (attwateri, difficilis, and nasutus), was resolved. Placement of *P. melanocarpus* was unresolved; however, its association with a clade containing members of the furvus, megalops, melanophrys, and mexicanus species groups was well supported. Although *P. perotaralis* showed some affinity with the truei group, this relationship was not strongly supported; consequently, *perotaralis* was placed (unresolved) in a clade with members of several species groups. *P. polius* was placed as the basal member of the *Peromyscus* (sensu lato) clade; however, the clade probability value of 89 indicated weak support for this arrangement. Placement of *P. polius* warrants some caution because it is represented by a single sample amplified from a skin clip (Tiemann-Boege et al. 2000).

Relationships among species groups were not well supported, in most cases, probably as an artifact of lack of resolution among clades in the middle regions of the tree. Exceptions included recovery of the aztecus and boylii species groups and the leucopus and maniculatus species groups as sister clades. A clade containing californicus and eremicus group members as well as Osgoodomys was well supported, indicating a strong affiliation among these taxa. Although there was little support for relationships among species groups, there were 3 clades that were well supported and provided phylogenetic information. The 1st clade (II) contained samples representing Osgoodomys and the californicus, eremicus, leucopus, maniculatus, crinitus, and hooperi species groups. The 2nd clade (IV) contained samples representing Habromys, Neotomodon, and Podomys; the truei, boylii, aztecus, and furvus species groups; and ochraventer and perotaralis. The 3rd clade (V) contained samples representing the mexicanus, melanophrys, furvus, and megalops species groups and melanocarpus.

**Genera and subgenera.**—Examination of the topology and support values associated with Fig. 2 provides several alternative interpretations of the status of genera and subgenera as proposed by Osgood (1909) and subsequent reviewers (Carleton 1980, 1989). First, Baiomys and Ochrotomys are basal to *Peromyscus* and are distantly removed from *Peromyscus*, as suggested in several studies (e.g., Engstrom and Bickham 1982; Reeder et al. 2006); in fact Reeder et al. (2006) and Musser and Carleton (2005) recognized Ochrotomys as a new tribe. Second, 5 genera (*Habromys*, *Neotomodon*, *Podomys*, *Megadontomys*, and Osgoodomys) and the subgenus *Haplomyomys* are embedded within *Peromyscus* (sensu stricto) and these arrangements are strongly supported. Therefore, to include all taxa in a monophyletic clade would necessitate recognition of clade 1 as *Peromyscus*. This also suggests a more basal position for Isthmomys than previously reported (Carleton 1980; but see Rogers et al. 2005) but has the advantage of supporting the recognition of Baiomys and Reithrodontomys as genera. Another option would be to recognize strongly supported nodes (clades II–V) as equal taxonomic ranks and invoke existing generic names where appropriate. However, this would result in no fewer than 5 generic names to recognize various clades and branching patterns depicted in the phylogenetic tree. This option does not offer a realistic or desirable solution at this time.
Phylogenetic relationships and status of Peromyscus.—Although our study is the most inclusive to date from a taxon-sampling standpoint (we examined approximately 70% of the species in Peromyscus (sensu lato), addition of remaining species and sequence data will impact overall phylogenetic reconstruction of the group. Moreover, in our discussion of phylogenetic relationships and status of species groups and genera we have not included any of the morphologic, allozymic, or chromosome data that are available for these taxa. Given that these complementary data are informative and necessary for testing the question of what is a Peromyscus, we have postponed a formal synthesis (revised classification) until more taxa can be included into a single analysis and these alternative (and potentially new) data sets are considered. However, several observations are worthy of further discussion.

Although it is premature to abandon the use of generic names, with the exception of Isthmomys, the concept of Peromyscus (sensu lato) was supported in all studies for which all genera were represented, including chromosome banding (Rogers et al. 1984; Stangl and Baker, 1984), allozymes (Rogers et al. 2005), and Cytb (present study). In addition, constraining taxa in our molecular data set to correspond to Peromyscus (sensu stricto) resulted in a significantly worse ($P < 0.02$) tree based on the test of Shimodaira and Hasegawa (1999). Second, the historical concept of species groups received strong support in our study. With a couple of minor discrepancies (placement of mayensis and stirtoni and diphyle of the truei group), the contents and validity of the 13 species groups of Musser and Carleton (2005) were supported by examination of the Cytb data. Levels of congruence among these studies are noteworthy, especially given that many of the species groups were compiled before the formal development of phylogenetic methodologies (Hennig 1966). Additionally, our study illustrates the accuracy and thoroughness of Osgood’s (1909) monograph and Carleton’s (1989) synthesis. Third, we provide the most comprehensive coverage of molecular phylogenetic relationships at the species level presented to date. Although the Cytb data contain considerable phylogenetic signal, definition of content and resolution of the phylogeny of Peromyscus awaits addition of taxa and sequence data, and synthesis of existing morphological, molecular, and behavioral data sets. Clearly, however, the current concept of the genus as comprising only the former subgenera Peromyscus and Haplomyomys is in need of revision.

Acknowledgments

We thank B. Amman for providing some of the DNA sequences. J. W. Bickham and R. L. Honeycutt (Texas A&M University) and R. J. Baker (Museum, Texas Tech University) kindly provided tissue loans. R. R. Chambers, M. L. Haynie, J. D. Hanson, A. O. Stallings, and D. D. Henson provided helpful comments on previous versions of this manuscript. This research was supported by the National Science Foundation (DBI-0139501), Professional Development Grants, Department of Integrative Biology and the M. L. Bean Life Science Museum, Brigham Young University (to DSR), and by a grant from the National Institutes of Health (DHHS A141435-01 to RDB).

Literature Cited


Kumar, S., K. Tamura, and M. Nei. 1993. MEGA: molecular evolutionary genetics analysis, version 1.02. Institute of Molecular Evolutionary Genetics, Pennsylvania State University, University Park.


Submitted 9 October 2006. Accepted 5 February 2007.

Associate Editor was Carey W. Krajewski.

APPENDIX I

Specimens examined.—For each specimen examined, the collection locality is given; specimens are from the United States unless otherwise noted. Some sequences were obtained from GenBank and do not have locality data. Museum acronyms, museum catalogue numbers, and GenBank accession numbers (AF, AY, DQ, EF, U, and X) are provided in parentheses, respectively. Abbreviations for museum acronyms (in parentheses and to the left of the semicolon) follow Háfner et al. (1997): Abilene Christian University Natural History Collection (ACU/NHC); Angelo State University Museum Natural History Collections (ASNHIC); California Academy of Sciences (ASMC); California Academy of Sciences Museum of Vertebrate Zoology (MVZ); The Museum of Vertebrate Zoology (MVZ); The Museum of Vertebrate Zoology (MVZ); The Museum of Vertebrate Zoology (MVZ); Texas Cooperative Wildlife Collection (TCWC); United States National Museum (USNM); University of Alaska Museum (UAM); University of Michigan Museum of Zoology (UMMZ) and Zoological Museum of the University of Michigan (UMMZ); and Zaddok Thompson Natural History Collection (ZTNHC). If museum catalogue numbers were unavailable, specimens were referenced with the corresponding collector’s numbers: CWK (personal collector number, C. William Kilpatrick, FSH (personal collector number, K. Nutt), or TK number (special number of the Museum of Vertebrate Zoology (MVZ)).

Bioaonos taylori.—Texas; Garza County, 22.5 km S, 0.6 km E Post (TTU454633, AF548469).

Habromys ixtlani.—MEXICO: Oaxaca; Distrito de Ixtlan, 28 km SW (by road) La Esperanza, 2,950 m (BYU15271, DQ861395); 2.2 km N Llano de las Flores, UTM 14-69458E-1931126N (TTU82703, DQ973099; UNAM—catalogue number unavailable, TK93160, DQ004822).

Isthmomys pirrensis.—Panama: Darien; Cerro Pirre, summit, 1.5 km (ROM16308, DQ863298; ROM16309, DQ863299).

Megadontomys erythrophys.—MEXICO: Oaxaca Municipio Teotitlan, 1.5 km S Puerto de la Soledad, 18°9.742′N, 96°59.852′W, 2,600 m (BYU16076, DQ861373).

Megadontomys thomasi.—MEXICO: Guerrer; 6.4 km SSW Filo de Caballo (UNAM—catalogue number unavailable, TK93388, AY195795).

Neotoma mexicana.—New Mexico; Los Alamos County, Los Alamos (TTU79129, AF294345).

Neotomodon alstoni.—MEXICO: Michoacán; Ladera, Cerro del Buurro, 3 km W Opopo-Tacamborro (UNAM—catalogue number unavailable, TK45302, AY195796); Mexico; Nevado de Toluca (UNAM—catalogue number unavailable, TK93093, AY195797); Morelos; Lagunas de Zempoala, 3 km W Huiztilac, 19°02.020′N, 99°17.627′W, 3,018 m (BYU15513, DQ861374).

Nyctomys sumichrasti.—HONDURAS: Atlantida; Lancellata Botanical Garden (TTU84848, AY195801).

Ochrotomys nuttalli.—Texas; Wood County, 5.6 km SE Quitman (TCWC31929, AY195798).

Onychomys arenicola.—Texas; Presidio County, Big Bend Ranch State Natural Area (TTU67559, AY195793).

Oryzomys palustris.—Texas; Galveston County, Texas City, Virginia Point (TTU82920, DQ185382).

Osogoodymys banderanus.—MEXICO: Jalisco; 6 km SE Chamela (TTU377754, AF155383); Michoacán; Coalcoman, Tehuantepec, 17.5 km WSW Coalcoman, 1,350 m, 18°42.95′N, 103°18.19′W (UNAM—catalogue number unavailable, TK45952, DQ000473).

Ototylomys phillois.—MEXICO: Campeche; El Remate 14 km W Tuncuche (ASNHC7236, AY009798).

Peromyscus attwateri.—Oklahoma; McIntosh County, 5 km E Dustin (TTU55688, AF155384); Texas; Knox County, 4.8 km E Benjamin (ZTNHC—catalogue number unavailable, DQ27, AF155385).

Peromyscus aztecus.—MEXICO: Veracruz; 8.8 km N Huatusco (TCWC47976, U89968).

Peromyscus beatae.—MEXICO: Veracruz; Xometa (TCWC48060, AF131921); Guerrero: 6.4 km SW Filo de Caballo (TCWC45222, AF131922); HONDURAS: Francisco Morazan; 3.2 km NE El Hatillo (TCWC52288, AF131914).

Peromyscus boylit.—California; Monterey County, Hastings Natural History Reservation (MVZ—catalogue number unavailable, KN120, AF155386); Utah; Garfield County, Henry Mountains, Mt. Pennell, Sidewell Spring, 2,652 m (MSB123149, AF155392); MEXICO: Jalisco; 30 km W Huejuquilla del Alto (TTU81702, AF155388).

Peromyscus californicus.—California; San Diego County, San Onofre State Beach, 5.6 km NNE on Christmas Road (TTU81275, AF155393).

Peromyscus crinitus.—Utah; Emery County, Cottonwood Canyon, 39°16′51.8′′N, 111°10′31.9′′W (BYU18639, AY376413); Kane County, 59 km E, 25 km N Kanab, UTM 12-41446E-4121440N, 1,450 m (BYU18065, EF028168); 59 km E, 25 km N Kanab, UTM 12-415040E-4121500N, 1,450 m (BYU18029, DQ681378).

Peromyscus difficilis.—MEXICO: Hidalgo; McIntosh County, 5 km E Dustin (TTU55688, AF155384); Texas; Knox County, 4.8 km E Benjamin (ZTNHC—catalogue number unavailable, DQ27, AF155385).

Peromyscus eremicus.—Arizona; Yavapai County, Sycamore Station, 34°23′28.2′′N, 112°3′1.3′′W (TTU797750, AY195799); California; Los Angeles County, Calabasas Creekside Park (TTU83249, AY322503); MEXICO: Zacatecas; 16.8 mi S Concepcion del Oro (TTU45197, DQ973100).

Peromyscus evideo.—MEXICO: Oaxaca; 5.6 km S Suchtepec (TCWC41855, DQ728970).

Peromyscus furvus.—MEXICO: Veracruz; 5 km W Naolinco de Victoria, Municipio Naolinco, 19°39′12′′N, 96°55′26′′W, 1,650 m (CNMA32298, AF271032); Oaxaca: 1.5 km S Puerto de la Soledad, 18°09.74′2′′N, 96°59.852′W, 2,690 m, Municipio Teotitlan (BYU15389, AF271012); Hidalgo, 21.8 km NE Metepec (ZTNHC—catalogue number unavailable, CWK913, AF271005).

Peromyscus gossypinus.—Oklahoma; Seminole County, 5.6 km E Seminole (TTU55019, DQ973101); Texas; Cass County, White Oak
Creek Wildlife Management Area, UT 15-350030E-3678790N (TTU80682, DQ973102).

**Peromyscus gratus**.—MEXICO: Durango; 6.1 km W Coyotes, Hacienda Coyotes, UT 13-465908E-2634281N (TTU81621, AY322509); Mexican; 4 km E Costzeo (UNAM—catalogue number unavailable, TK45309, DQ000481).

**Peromyscus hylotis**.—MEXICO: Michoacán; Puerto Garnica (ZTNHC—catalogue number unavailable, CWK2032, AF155403).

**Peromyscus keeners**.—Washington; Gray Harbor County, 0.8 km NE Grisdale (ZTNHC—catalogue number unavailable, CWK2235, DQ973108); Alaska; Alexander Archipelago, Bushy Island, Petersburg Quad (UAM50770, AF119261).

**Peromyscus leptopus**.—Texas; Dickens County, 0.9 km E Afton (TTU—catalogue number unavailable, TK47506, AF131926); La Salle County, Chaparral Wildlife Management Area, UT 14-454285E-3136441N (TTU104424, DQ973104; TTU101645, DQ000483).

**Peromyscus levipes**.—MEXICO: Nuevo Leon; Cola de Caballo (ZTNHC—catalogue number unavailable, CWK2436, AY376423).

**Peromyscus madrensis**.—MEXICO: Nayarit; Isla María Madre (USNM512599, AF155397).

**Peromyscus maniculatus**.—Texas; Wichita County, 6.5 km S Wichita Falls (TTU38739, DQ000484); Washington; Pierce County, Ft. Lewis, Reganberg Mount (TTU104426, DQ973111); MEXICO: Durango; 6.1 km W Coyotes Hacienda Coyotes, UT 13-465908E-2634281N (TTU81622, AY322508).

**Peromyscus mayensis**.—GUATEMALA: Huetengano; Santa Eulalia, 16 km NW Santa (ROM98339, DQ836360; ROM98360, DQ836301).

**Peromyscus megalops**.—MEXICO: Guerrero; 6.4 km SSW Filo de Caballo (TTU82712, DQ000475).

**Peromyscus melanocarpos**.—MEXICO: Oaxaca; Municipio Santiago Comaltepec, La Esperanza, 11 km SW Hacienda San Isidro (CMMC29192, EF028173); Municipio Tlahuitoltepec, vicinity Santa María Yacoch, 17°15’S, 96°00’45”W, 2400 m (CMMC136, EF028175).

**Peromyscus melanophilus**.—MEXICO: Durango; 2.2 km S, 2.5 km E Vicente Guerrero (TTU75509, AY322510); Jalisco; 30 km W Huejuquilla del Alto (UNAM—catalogue number unavailable, TK48638, AY376424); Zacatecas; 16.8 km S Concepcion del Oro (TTU49351, DQ973105).

**Peromyscus melanotis**.—MEXICO: Durango; 12 km E Ojitos, UT 13-385011E-2775718N (UNAM—catalogue number unavailable, TK70997, AF155398).

**Peromyscus mexicanus**.—MEXICO: Chiapas; 14.4 km N Ocozocauac, 15-451772E-1864243N (TTU82759, AY376425); Veracruz; 10 km SE (por carretera) Zongolica, 1,850 m, (CNMA34309, EF028174).

**Peromyscus nasutus**.—New Mexico; Lincoln County, 6.4 km S Carrizo So (TTU78401, AF155399); Texas; Jeff Davis County, Mt. Livermore Preserve, UT 13-579953E-3389871N (TTU78316, AY376426).

**Peromyscus nuditus**.—COSTA RICA: Heredia; 2 km NE Getzemani, 10.04612°N, 84.10377°W (MSB61868, AY041200).

**Peromyscus oaxacensis**.—MEXICO: Oaxaca; 1.4 km N Llano de las Flores (TCWC45192, U89987).

**Peromyscus ochraventer**.—MEXICO: San Luis Potosí; 0.8 km E Las Abritas (TCWC48405, DQ973106).

**Peromyscus pectoralis**.—Texas; Killeen County, Walter Buck Wildlife Management Area, UT 14-423359E-3366337N (TTU71252; AF155400); MEXICO: Durango; 2.4 km SE Los Herreras (UNAM—catalogue number unavailable, TK48519, AY322511); Jalisco; 30 km W Huejiquilla del Alto (UNAM—catalogue number unavailable, TK48642, AY376427).

**Peromyscus perfurians**.—MEXICO: Michoacán; Tunel de Riego, 2 km E Cerro Colorado, 1,290 m, 19°19’220”N, 100°28’308”W (UNAM—catalogue number unavailable, TK47926, DQ000474).

**Peromyscus polionotus**.—Captive reared (Peromyscus Stock Center, Columbia, South Carolina; X89972).

**Peromyscus sagas**.—MEXICO: Chiuhuahua; 4.8 km SW Santa Barbara (TCWC47255, AF155403).

**Peromyscus schmidlyi**.—MEXICO: Durango; 30 km SW Ojitos (TTU81635, AY322520); 12 km E Ojitos, UT 13-385011E-2775718N (TTU81602, AY355405); 6.1 km W Coyotes, UT 13-465908E-2634281N (TTU81617, AY370610).

**Peromyscus simulans**.—MEXICO: Sinaloa; 6.4 km E Concordia, Highway 40 (TCWC45592, AF131927).

**Peromyscus spicilegus**.—MEXICO: Durango; San Juan de Camarones, UT 13-356906E-2757418N (TTU81640, AY322512; UNAM—catalogue number unavailable, TK70919, DQ973107); Michoacán; km 81 carr. Ario de Rosales–La Huacana, 1,602 m, 19°10’59”N, 101°43’42”W (UNAM—catalogue number unavailable, TK47888, DQ000480).

**Peromyscus stephani**.—MEXICO: Sonora; Isla San Esteban (UMMZ117385, AF155411).

**Peromyscus stirtioni**.—NICARAGUA: Parque Nacional, Volcan Masaya (MCZ61922, DQ973108).

**Peromyscus truei**.—Arizona; Apache County, 35°34’51”N, 109°34’33”W (TTU104427, AY376433); California; Alameda County, Strawberry Canyon (MYV157329, AF108703); Texas; Armstrong County, 1.2 km N, 10 km E Wayside (TTU61543, AY376428).

**Peromyscus winkelmanni**.—MEXICO: Michoacán; 11 km WSW Dos Aguas (TCWC45621, AF131930); Guerrero; Filo de Caballo (TCWC45175, U89983).

**Peromyscus zurhyncha**.—MEXICO: Chiapas; Yalentay, UT 15-385117E-1852456N (UNAM—catalogue number unavailable, TK46354, AY376421); Puebla; 5 km SE San Antonio, UT 14-664598E-2030332N (TTU82701, AY376422).

**Peromyscus oaxacensis**.—MEXICO: Oaxaca; 1.4 km N Llano de las Flores (TCWC45192, U89987).

**Peromyscus oaxacensis**.—MEXICO: Oaxaca; 1.4 km N Llano de las Flores (TCWC45192, U89987).

**Peromyscus oaxacensis**.—MEXICO: Oaxaca; 1.4 km N Llano de las Flores (TCWC45192, U89987).

**Peromyscus oaxacensis**.—MEXICO: Oaxaca; 1.4 km N Llano de las Flores (TCWC45192, U89987).

**Peromyscus oaxacensis**.—MEXICO: Oaxaca; 1.4 km N Llano de las Flores (TCWC45192, U89987).