Genealogical Concordance and the Specific Status of Peromyscus sejugis

Mindy L. Walker, Scott E. Chirhart, Ashli F. Moore, Rodney L. Honeycutt, and Ira F. Greenbaum

From the Department of Biology, Texas A&M University, College Station, TX 77843 (Walker, Moore, and Greenbaum); the Department of Biology, Centenary College of Louisiana, Shreveport, LA 71134 (Chirhart); and the Department of Wildlife and Fisheries Sciences, Texas A&M University, College Station, TX 77843 (Honeycutt).

Address correspondence to I. F. Greenbaum at the address above, or e-mail: i-greenbaum@tamu.edu.

Abstract

Peromyscus sejugis, a peripheral isolate of Peromyscus maniculatus, is a threatened taxon endemic to 2 small islands in the Sea of Cortés. Although its insularity makes the specific recognition of P. sejugis inherently problematic, resolution of this problem has important conservation implications. To evaluate the specific validity and evolutionary history of P. sejugis, we compared sequence variation (ND3/ND4L/ND4) in mtDNA for both island populations of P. sejugis with that for 8 populations of P. maniculatus from mainland Baja California. Each island population of P. sejugis had a single haplotype (0.7% sequence divergence), whereas 11 different haplotypes (mean sequence divergence = 0.68%) were obtained for the populations of P. maniculatus. The mean sequence divergence between the populations of the 2 species was 2.0%. Nested clade analysis supports the conclusion that P. sejugis is an insular isolate of P. maniculatus from mainland Baja California. Although our analysis confirms a low level of mtDNA divergence between P. sejugis and P. maniculatus from Baja California, the genealogical concordance of morphological, chromosomal, microsatellite, and mtDNA haplotype distinctiveness supports the conclusion that the 2 island populations of P. sejugis constitute independent evolutionarily significant units and together represent a phylogenetic species distinct from the P. maniculatus from Baja California.

Due to the peninsula’s dynamic geologic history, heterogeneous habitats, and various areas of endemism, the biogeography and evolutionary history of Baja California, Mexico, are complex. Reconstructions of the faunal history of this region are confounded by the debate over vicariance versus dispersal as the explanation for its modern organismal diversity. The current distributions of vertebrates in this region have historically been attributed to late Pleistocene through Holocene dispersal from continental North America (Orr 1960; Savage 1960). More recently, however, the present-day biotic architecture of mainland Baja California has been hypothesized to have arisen via a series of vicariant events during the late Neogene (5.5–1 mya, Murphy 1983; Grismer 1994; Riddle 1995; Riddle, Hafner, Alexander, and Jaeger 2000; Riddle and others 2000a, 2000b; Carreño and Helenes 2002).

After the purported late Neogene origin of the Sea of Cortés, vicariant events that generated the arid islands in this sea likely led to fragmentation of the mainland arid-adapted Pliocene and Pleistocene faunas of Baja California. Correspondingly, these islands are inhabited primarily by desert-adapted mammals with sister taxa on the proximal mainland (Lawlor and others 2002). Sufficiently isolated island populations are expected to have experienced independent evolutionary trajectories as a result of founder effect with subsequent genetic drift and inbreeding. However, mtDNA sequence data (Hafner and others 2001) suggest that some populations of mammals inhabiting islands in the Sea of Cortés originated by dispersal from nonadjacent mainland sources (restricted gene flow with isolation by distance).

Mice of the genus Peromyscus are a major component of the mammalian fauna of Baja California and the most prevalent mammals on the islands in the Sea of Cortés; 7 peninsular species of Peromyscus (Cricetidae, Neotominae) are variously present on these islands (Hafner and others 2001). Two species in the Peromyscus maniculatus group are currently recognized in the Baja California Peninsular Desert Region. Peromyscus maniculatus is common in Baja California Norte, but its occurrence in Baja California Sur is sparse and poorly documented (Hall 1981). Peromyscus sejugis, the Santa Cruz Island mouse (Dice 1940; Blair 1950; Hooper 1968; Carleton 1980; Hall 1981), is restricted to 2 small islands off the coast of Baja California Sur, Isla Santa Cruz (ISC, 14 km²) and Isla San Diego (ISD, 1.3 km²), in the Sea of Cortés and considered threatened by the Government of Mexico (Álvarez-Castañeda 2001).

From data for a single individual of P. sejugis and one of P. maniculatus from Baja California Norte, Hogan and others (1997) reported an mtDNA (ND3/ND4L/ND4 region)
sequence divergence of 2% (mistakenly listed as 0.02%). Similarly, low sequence divergence at the COIII gene for 3 individuals of P. sejugis from each island and 3 specimens of P. maniculatus from 1 locality in Baja California Sur led Hafner and others (2001) to suggest that a more thorough sampling of peninsular P. maniculatus would likely indicate that P. sejugis should be included as a subspecies of P. maniculatus. The specific distinction of P. sejugis and peninsular P. maniculatus is based on general morphological characteristics and appears to be supported by karyotypic data. Relative to P. maniculatus from peninsular Baja California, Burt’s (1932) recognition of P. sejugis considered that the latter is larger in size and has a duller pelage, a lighter lateral line, a longer rostrum, and noninflated frontals (Alvarez-Castañeda 2001). These taxa also differ for independent pericentric inversions of the pleiomorphic (for Peromyscus) acrocentric condition of chromosome 13 (Smith and others 2000). Character state differences for the presence of distal heterochromatin on the short arm of chromosome 13 suggest a lack of recent gene flow between the 2 island populations of P. sejugis (Smith and others 2000).

Herein we present an analysis and a comparison of sequence variation in the mitochondrial ND3/ND4L/ND4 (NADH dehydrogenase) genes for both populations of P. sejugis (n = 20) and 8 populations of P. maniculatus from mainland Baja California (n = 96). The objectives of this study were to further assess the taxonomic status of P. sejugis and to address the question of whether the populations of P. sejugis are the result of fragmentation of, or dispersal from, an ancestral mainland population.

Materials and Methods

Qiagen purification kits and procedures (Qiagen, Inc., Valencia, CA) were used to isolate DNA from frozen (~80 °C) liver or spleen samples of specimens of P. sejugis (Mexico—Baja California del Sur: ISD [n = 13] and ISC [n = 7]) and P. maniculatus (Mexico—Baja California del Norte: Vallecitos [VLL, n = 26]; Laguna Hanson, Sierra Juarez [LH, n = 29]; 16’ S, 5’ E, or 8’ S, 9’ E Valle de Trinidad [VDT, n = 31]; 3 km SW Colonio Vicente Guerrero [CVG, n = 3]; Mision San Fernando [MSF, n = 1]; 27 km S Punta Prieta [PP, n = 4]; Baja California del Sur: 25° SE Guerrero Negro [GN, n = 1] and 11 km S Todos Santos [TS, n = 1], Figure 1). The animal use in this research was conducted in accordance with the Guide for Care and Use of Laboratory Animals and approved by the Texas A&M University Laboratory Animal Care and Use Committee.

Amplification with the polymerase chain reaction (PCR) and sequencing of the 1439 bp fragment of the mitochondrial ND3/ND4L/ND4 genes as well as tRNAArg and the 3′-end of tRNAGly generally followed the techniques of Arevalo and others (1994). The primers used for PCR amplification and sequencing included: PI’, Marg, ND4L, and Nap2. Amplification reactions (Perkin Elmer/Cetus DNA Thermal Cycler) were conducted with the following reagents and concentrations: 1 µl DNA (ca. 100 ng), 12.3 µl H2O, 2.5 µl of 10× PCR Buffer II (PE Applied Biosystems, Foster City, CA), 2.5 µl of 25 mM MgCl2, 0.5 µl bovine serum albumin, 4 µl of 8 mM deoxynucleoside triphosphates (Amersham Pharmacia Biotech, Piscataway, NJ), 1.0 µl of forward and reverse primers, and 0.2 µl Taq (TaKaRa). Amplifications proceeded in 3 stages, including an initial denaturation cycle at 95 °C for 5 min; followed by 35 cycles of 1 min each at 95 °C, 50 °C, and 72 °C; and concluded with an extension cycle of 10 min at 72 °C. Amplified products were purified using exonuclease I in combination with shrimp alkaline phosphatase (ExoSAP, USB, Cleveland, OH).

Each sequencing reaction was performed with a Big Dye sequencing kit (PE Applied Biosystems) in a Perkin Elmer/ Cetus DNA Thermal Cycler. Amplifications followed the protocol outlined in the sequencing kit, and sequences were obtained on an Applied Biosystems 377 automated sequencer. Fragments were sequenced in both directions; sequence alignments and the formation of contigs were conducted using the program Sequencher 4.1.1 (Gene Codes Corporation, Ann Arbor, MI). Each unique sequence was scored as an individual haplotype. GenBank Accession numbers for each of the haplotypes are presented in the Results.

Sequences of haplotypes were compared, and their within- and among-locality variation was analyzed. Phenetic and phylogenetic analyses included neighbor-joining (NJ), maximum parsimony (MP), and maximum likelihood (ML)
approaches in PAUP*4b10 (Swofford 2002). MP analyses of the sequence data consisted of heuristic searches using tree bissection–reconnection (TBR) branch swapping with equal weights and random additions. Similar heuristic searches were performed with unequal weights for transitions and transversions (following Hogan and others 1997). Uncorrected (p) distances were computed according to Swofford (2002). Modeltest 3.06 (Posada and Crandall 1998) identified HKY + I + G (Hasegawa, Kishino, and Yano + invariant sites + gamma distribution for variable sites, Hasegawa and others 1985) as the most appropriate model of nucleotide evolution for the data. Both NJ and ML analyses employed this model; ML analysis (a = 0.7648; p$_{str}$ = 0.6163; transition:transversion ratio = 6.6503) was conducted with heuristic searches using TBR branch swapping with equal weights and random additions. All analyses included reference sequences (Hogan and others 1997; Chirhart and others 2001) for single individuals of Peromyscus maniculatus rufinus (GenBank Accession U40250) and Peromyscus maniculatus austerus (GenBank Accession U40249); sequence data for Peromyscus melanotis (Hogan and others 1997; Chirhart and others 2001) from Hidalgo, Mexico (GenBank Accession U40247) were used as the out-group for all phylogenetic analyses. Bootstrap estimates (Felsenstein 1985) based on 1000 replications were obtained for MP and ML analyses.

Haplotypes were also subjected to nested clade analysis (NCA) using TCS (Clement and others 2000) and GeoDis (Posada and others 2000) and to subsequent phylogeographic inference. The inference key (Templeton and others 1995; Appendix I, Templeton 1998) was applied to the GeoDis output in order to attempt to differentiate between alternative biogeographic hypotheses. Although simulation studies (Irwin 2002; Knowles and Maddison 2002) have criticized the NCA inference key’s capacity to evaluate alternative phylogeographic hypotheses, the method does provide a statistical framework for characterizing genetically distinct populations a posteriori through examination of their geographic distribution and frequency of haplotypes. Templeton (2004) maintained that NCA is complementary to an a priori procedure described by Knowles and Maddison (2002).

Results
Each of the island populations of P. sejugis exhibited a single unique haplotype: Haplotype 12, ISD (GenBank Accession U40255), and Haplotype 13, ISC (GenBank Accession U40253), differing by a p distance of 0.7%. The 8 population samples of P. maniculatus from peninsular Baja California exhibited a total of 11 different haplotypes with a mean p distance of 0.68%. Haplotype 1 (GenBank Accession DQ077697) was the most frequent and geographically widespread (and thus presumably ancestral) haplotype, characterizing 38 of the 96 individuals of P. maniculatus (14 VLL, 12 LH, 10 VDT, 1 CVG, and 1 PP). Other relatively frequent haplotypes included the following: 17 individuals with Haplotype 2 (6 VLL, 5 LH, and 6 VDT; GenBank Accession DQ077693), 19 with Haplotype 3 (6 VLL, 7 LH, and 6 VDT; GenBank Accession DQ077696), and 8 with Haplotype 4 (5 LH and 3 VDT; GenBank Accession DQ077694). The haplotype exhibited by the single individual from GN (Haplotype 5, GenBank Accession DQ077698) was shared with 1 individual from CVG. Twelve individuals were characterized by haplotypes that were not observed at any other locality: Haplotype 6 (GenBank Accession DQ077695) and 7 (GenBank Accession DQ077703; 3 individuals each from VDT), Haplotype 8 (GenBank Accession DQ077699; 1 CVG), Haplotype 9 (GenBank Accession DQ077702; the single individual from MSF), Haplotype 10 (GenBank Accession DQ077700; 3 PP), and Haplotype 11 (GenBank Accession DQ077701; the single individual from TS).

The mean uncorrected (p) distance between the sequences of P. sejugis and those from the 8 populations of P. maniculatus was 2%. All analyses (NJ, MP, and ML) recovered nearly identical topologies (Figure 2), with the NJ tree providing more resolution among the haplotypes of P. maniculatus. All trees grouped the 2 populations of P. sejugis as sister to the populations from peninsular Baja California and these as being distinct from, and sister to, the reference samples of P. m. austerus and P. m. rufinus. The MP and NJ analyses placed the TS haplotype as basal among the populations from peninsular Baja California; the ML analysis did not resolve the TS haplotype relative to the other peninsular samples (Figure 2).

NCA (Figure 3) identified phylogroups identical to those in the NJ tree, and the number of intermediate haplotypes identified by TCS in the haplotype network was concordant with the NJ branch lengths (number of steps between phylogroups). The haplotypes (including the presumed ancestral haplotype) from the localities in Baja California Norte and GN clustered with one another with few missing intermediates (Figure 3). This phylogroup linked to the haplotype from southernmost Baja California Sur (TS) and then to the ISD and ISC haplotypes of P. sejugis, respectively. TCS analysis generated no reticulations. GeoDis identified 2 clades that violated the null hypothesis of panmixia (Table 1). These included a first-level clade (geographically) between the 2 populations of P. sejugis (Clade 1-3) and a second-level clade (geographically) between the populations of P. sejugis and the samples from Baja California (Clade 2-2). According to the inference key, the latter significant value led to a result that was inconclusive with regard to distinguishing between fragmentation and isolation by distance as responsible for the origin of P. sejugis.

Discussion
Microevolutionary Factors
Although our sample sizes of P. sejugis were not large, the observation of endemic and invariant haplotypes is supported by previous studies indicating minimal genetic variation in these 2 small-island populations. For the same individuals as examined in this study, banded karyotypes were invariant within each island population (Smith and others 2000), and microsatellite variation (Chirhart and others 2005) was minimal and significantly lower than in reference mainland populations of P. maniculatus (Kansas) and the other species
in the *P. maniculatus* group. Additionally, Avise and others (1974) reported only 1.7% per-individual heterozygosity at 23 allozymic loci for a total of 31 individuals of *P. sejugis* representing both islands.

The apparently fixed and endemic haplotypes of the 2 island populations of *P. sejugis* and the comparatively rich haplotypic diversity of peninsular *P. maniculatus* are consistent with the conclusion (Smith and others 2000) that there has been a lack of recent gene flow among these populations. The paucity of genetic variation and the pattern of divergence observed for the mitochondrial genes in *P. sejugis* suggest that the origin of these populations has been strongly influenced by genetic drift. Similarly, genetic drift has apparently been the primary process accounting for the differences between the 2 populations of *P. sejugis*.

**Phylogeographic Considerations**

The phenetic analyses, phylogenetic analyses, and NCAs of the sequence variation across the *ND3/ND4L/ND4* region of the mtDNA (Figures 2 and 3) were entirely consistent in indicating an overall genetic continuity among the populations of *P. maniculatus* and in clustering the 2 island populations of *P. sejugis* outside of the group of mainland populations and distinct from one another. Phylogeographic analysis (Table 1, Figure 3) suggested that the 2 populations of *P. sejugis* became separated via past fragmentation but did not provide definitive distinction between past fragmentation and isolation by distance as the mechanism responsible for the current pattern of mtDNA variation between *P. sejugis* and *P. maniculatus* from Baja California. However, we favor the fragmentation model for the following reasons: 1) the autapomorphies mentioned above establish the mainland and island populations as distinct evolutionary units with no evidence of the homogenizing effects of gene flow, 2) islands in this same archipelago (Isla Espiritu Santo and Isla San
Table 1. Summary of inferences (Templeton 1998) regarding demographic events deduced from clades with significant nested clade values

<table>
<thead>
<tr>
<th>Clade</th>
<th>( \chi^2 )</th>
<th>Nested clades</th>
<th>( D_c )</th>
<th>( D_n )</th>
<th>Chain of inference</th>
<th>Demographic event</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-3</td>
<td>( P = 0.0 )</td>
<td>VI (interior)</td>
<td>0.0, &lt; ( P = 0.0^b )</td>
<td>3.9573, &gt;, ( P = 0.0 )</td>
<td>1 yes, 2 yes, 3 yes, 5 yes, 15 no</td>
<td>Past fragmentation</td>
</tr>
<tr>
<td>2-2</td>
<td>( P = 0.004 )</td>
<td>VIII (tip)</td>
<td>0.0, &lt; ( P = 0.02 )</td>
<td>3.9571, &lt;, ( P = 0.0 )</td>
<td>1 yes, 2 yes, 3 no, 4 yes, 9 yes, 10 no</td>
<td>Geographic sampling scheme inadequate to discriminate between fragmentation and IBD(^d)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>V (interior)</td>
<td>NS(^c)</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-3</td>
<td></td>
<td></td>
<td>3.9572, &lt;, 0.004</td>
<td>206.3, &lt;, 0.004</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) > or < indicates a \( D_c \) (clade distance) or \( D_n \) (nested clade distance) value that is significantly larger or smaller, respectively, than expected if haplotypes were distributed randomly.

\(^b\) \( P \) values indicate the probability that the \( D_c \) or \( D_n \) values estimated from the data were observed by chance.

\(^c\) NS = not significant at \( P < 0.05 \).

\(^d\) IBD = isolation by distance.

Jose), and thus presumably ISD and ISC, were severed from the present-day La Paz Peninsula during a Pleistocene glacial advance 25 000–17 000 years ago (Beal 1948), thus potentially fragmenting a formerly contiguous ancestral population, and 3) the geologic transgressions that resulted in the present sea level 6000 years ago (Avise and others 1974) would likely have precluded recent waif dispersal and gene flow. Although a more intensive sampling of southern peninsular deer mice might reveal intermediate sequences that would support an isolation by distance-based origin of \( P. sejugis \), the scarcity of \( P. maniculatus \) in Baja California Sur will likely preclude obtaining the samples necessary to test this hypothesis.

**Taxonomy**

Due to the complications of insular allopatri, any decision pertaining to the specific distinction of \( P. sejugis \) is necessarily subjective. Given the minimal distribution and threatened status of \( P. sejugis \), however, the specific recognition of \( P. sejugis \) has important conservation implications and should be carefully considered. Although our analysis of sequence variation across the \( ND3/ND4L/ND4 \) region confirms initial observations (Hogan and others 1997; Hafner and others 2001) of a low level of mtDNA divergence between \( P. sejugis \) and \( P. maniculatus \) from Baja California, this alone does not warrant the conclusion that \( P. sejugis \) should be considered as conspecific with \( P. maniculatus \). Moreover, the pattern of genealogic concordance of the morphological (Burt 1932), chromosomal (Smith and others 2000), microsatellite (Chirhart and others 2005), and mtDNA variation supports the conclusion that \( P. sejugis \) and \( P. maniculatus \) from Baja California are independent evolutionary lineages and represent separate phylogenetic species (Cracraft 1983; Nixon and Wheeler 1990) as modified by inferences from coalescent theory (Avise and Ball 1990; Avise and Wollenberg 1997).

A similar, although less compelling, argument could be applied to hypothesize that the 2 island populations of \( P. sejugis \) are separate phylogenetic species. Although mtDNA, karyotypic, microsatellite, and morphological data are concordant in indicating a lack of recent gene flow between the 2 island populations of \( P. sejugis \), the differences in these characteristics reflect less time since divergence than do those that distinguish the 2 populations of \( P. sejugis \) from peninsular \( P. maniculatus \). Whereas the difference between the alternate \( ND3/ND4L/ND4 \) haplotypes of the 2 populations of \( P. sejugis \) was 0.7%, the mean sequence divergence between the haplotypes of \( P. sejugis \) and those of peninsular \( P. maniculatus \) was 2.0%. The NCA (Table 1) primary-level nesting of the populations of \( P. sejugis \) and secondary-level nesting for \( P. sejugis \) versus the peninsular deer mice is consistent with the hypothesis that there has been less time since the divergence of the 2 island populations of \( P. sejugis \). Karyotypically, the \( P. sejugis \) from ISC differ from those on ISD by the presence of distal heterochromatin on the short arm of chromosome 13, whereas both island populations of \( P. sejugis \) differ from peninsular \( P. maniculatus \) by a unique inversion of this same chromosome (Smith and others 2000). Although the 2 island populations of \( P. sejugis \) were either monomorphic for the same allele or shared the same alleles at similar frequencies at 8 microsatellite loci, these populations exhibited a fixed difference at 1 microsatellite locus and unique low frequency alleles at 2 other loci (Chirhart and others 2005). Alvarez-Castañeda (2001) reported minor qualitative differences in the shape of the nasals and angle of the sutures between the frontals and parietals between specimens of \( P. sejugis \) from the 2 islands but noted that external and cranial measurements did not differ. Given the circumstances and available data, we conclude that \( P. sejugis \) should retain its specific distinction and that the 2 island populations of \( P. sejugis \) at least constitute evolutionarily significant units (for review see Hey and others, 2003, and references therein), which, for conservation purposes, should be considered as separate entities and managed independently.

**Acknowledgments**

This research was partially supported by National Institutes of General Medical Sciences grant GM 27014 (to I.F.G.), National Science Foundation grant
References


Hogan KM, Davis SK, Greenbaum IF. 1997. Mitochondrial-DNA analysis of the systematic relationships within the *Peromyscus maniculatus* species group. J Mammal 78:733–43.


Received November 28, 2005
Accepted March 16, 2006

Corresponding Editor: William Modi