INTRON 2 OF THE ALCOHOL DEHYDROGENASE GENE (Adh1-12): A NUCLEAR DNA MARKER FOR MAMMALIAN SYSTEMATICS

BRIAN R. AMMAN, J. DELTON HANSON, LISA K. LONGHOFER,
STEVEN R. HOOFER, AND ROBERT D. BRADLEY

ABSTRACT

DNA sequences from a novel nuclear marker (2nd intron of the vertebrate alcohol dehydrogenase gene, Adh1-12) and the mitochondrial cytochrome-b gene (Cyt-b) were examined in 41 species and six genera of rodents (one Holochilus, 13 Neotoma, six Oryzomys, one Ototylomys, 19 Peromyscus, and one Tylomys). The Adh1-12 dataset was characterized by having a significant level of phylogenetic signal (P < 0.01) and a low level of homoplasy (consistency index = 0.78, retention index = 0.89). Based on parsimony, maximum likelihood, and Bayesian analyses, the Adh1-12 results were congruent with results from Cyt-b and other types of data, and generally were accompanied by high statistical support. All phylogenetic relationships at and above the genus level were strongly supported. Furthermore, compared with Cyt-b, the Adh1-12 dataset provided supported resolution to more species-level relationships within Neotoma, Peromyscus, and Oryzomys, despite fewer overall characters (489 Adh1-12; 1,143 Cyt-b) and fewer parsimony informative characters (126 Adh1-12; 464 Cyt-b). Overall, results of this study indicated that Adh1-12 DNA sequences are useful for addressing phylogenetic relationships within and among the Sigmodontinae and Neotominae. The slower rate of molecular evolution observed in Adh1-12 sequences, coupled with low levels of homoplasy and strength of resolution for supraspecific relationships, indicated that this marker may be useful for resolving phylogenetic relationships at low to intermediate taxonomic levels in mammals.

Key words: alcohol dehydrogenase gene, cytochrome-b gene, molecular systematics, rodents
**Introduction**

The use of nucleotide sequence data has greatly enhanced our knowledge of mammalian systematics. In mammals, the majority of molecular-based phylogenies have been estimated by analysis of maternally inherited, mitochondrial DNA (mtDNA) sequences, especially protein coding genes. Advantages in using mtDNA sequences in phylogeny reconstruction include: technical ease in DNA isolation and polymerase chain reaction amplification, manageable sequence lengths (generally <1,200 base pairs, bp), ease of alignment, availability of "universal" primers, differential rates of molecular evolution in specific regions of the molecule, significant phylogenetic signal, and substantial databases (GenBank, EMBL, etc.) of comparative material. However, there are drawbacks to mtDNA that can result in a misrepresentation of the true phylogeny, including retention of ancestral polymorphisms, incomplete lineage sorting, differential rates of evolution among lineages, and hybridization (Avise 1994; Hillis et al. 1996; Prychitko and Moore 2000). Likewise, base composition of mtDNA protein coding genes (excess of C and A at the 3rd positions -Anderson et al. 1981; Roe et al. 1985; Desjardins and Morais 1990; Prychitko and Moore 2000; Ballard et al. 2002) may result in codon bias and homoplasy (Prychitko and Moore 2000) leading to a biased reconstruction of ancestral sequences (Collins et al. 1994; Lockhart et al. 1994; Perna and Kocher 1995). Although some of these drawbacks can be addressed, nuclear DNA sequences are highly desirable as an alternative dataset to independently test mtDNA hypotheses and to essentially counteract criticisms of mtDNA.

Historically, nuclear DNA sequences have been used to address phylogenetic questions pertaining to relationships above the genus level (e.g., rRNA genes - Gouy and Li 1989; Perasso et al. 1989), principally because nuclear sequence markers with levels of variation suitable for resolving relationships among closely related species or genera are scarcely known and have been difficult to discover. Intron markers are an obvious and likely source of nuclear variability suitable for low- or intermediate-level phylogenetics. With the exception of some conserved areas, such as regulatory sites (Jackson and Hoffmann 1994), most nucleotide positions within introns are adaptively neutral and independently distributed and generally are evolving faster than those within exons, thus increasing their potential in population genetics or phylogenetic analysis of closely related species (Ballard et al. 2002; Prychitko and Moore 2000; Slade et al. 1994; Wickliffe et al. 2003). In addition, introns of various sizes are relatively abundant in the nuclear genome, and usually primer design in conserved areas of adjacent exons enable efficient amplification across taxonomic boundaries. In vertebrates, several introns have been shown to be useful in studies at or below the genus level (Prychitko and Moore 1997, 2000; Oakley and Phillips 1999; Lavoué et al. 2003; Reeder and Bradley 2003; Johanson and Ericson 2004; Helbig et al. 2005), and others have been shown to be relatively invariant at or below the species level (Wickliffe et al. 2003; Carroll and Bradley 2005). Alternatively, Fonseca et al. (in press) and Porter et al. (in review) found intron sequences useful for resolving relationships among closely related species of bats.

The purpose of this study was to examine the phylogenetic utility of DNA sequence data from the 2nd intron of the alcohol dehydrogenase gene 1 (Adh1) obtained for 41 species in six genera of rodents. Adh genes, in mammalian systems, code for dimeric zinc-metallo enzymes that catalyze oxidization of alcohols to aldehydes or ketones (Szalai et al. 2002). Seven distinct classes are recognized in vertebrates (Duester et al. 1999; Dolney et al. 2001; Szalai et al. 2002), six of which are present in mammalian species (Duester et al. 1999; Szalai et al. 2002). In rodents, four Adh classes (1, II, III, and IV) have been reported (Crabb and Edenberg 1986; Zhang et al. 1987; Park and Plapp 1991; Bradley et al. 1993, 1998; Duester et al. 1999; Szalai et al. 2002) and are controlled by six genes, Adh1, Adh2, Adh3, Adh4, Adh5a, and Adh5b (Duester et al. 1999; Dolney et al. 2001; Szalai et al. 2002). Adh1 is approximately 13 kilobase pairs in length (Crabb et al. 1989), contains nine exons (coding between 6 and 87 amino acids) and eight introns (ranging from 91 bp to 3.6 kilobase pairs; Zhang et al. 1987). Hereafter, in referring to the 2nd intron, we follow the recommended nomenclature (Adh1-12) of vertebrate alcohol dehydrogenases outlined in Duester et al. (1999).
MATERIALS AND METHODS

Specimens examined.—Specimens examined, GenBank accession numbers, and information associated with museum vouchers are listed in Appendix I. Complete sequences for Adh1-12 and cytochrome-b (Cyt-b) were obtained for 41 individuals and 10 individuals, respectively, and were deposited in GenBank. In addition, 31 Cyt-b sequences were retrieved from GenBank. Holochilus chacarius was designated as the outgroup taxon in phylogenetic analyses of Adh1-12 and Cyt-b data, as previous morphological and molecular studies agree that Holochilus may be outside the remainder of taxa in this study (reviewed in Musser and Carleton 2005). Sequence data were evaluated and relationships were inferred among the 40 ingroup species (13 Neotoma, six Oryzomys, one Ototylomys, 19 Peromyscus, and one Tylomys) representing the subfamilies Neotominae and Sigmodontinae (Musser and Carleton 2005).

Adh1-12 data.—Genomic DNA was isolated from approximately 0.1 g frozen muscle or liver tissues by the method of Smith and Patton (1999). Polymerase chain reaction (PCR; Saiki et al. 1988) primers for Adh1-12 (Table I) were designed from conserved regions in exons 2 and 3 based on sequences from Homo (Ikuta et al. 1986), Geomys (Bradley et al. 1993, 1998), Mus (Zhang et al. 1987), Peromyscus (Zheng et al. 1993), and Rattus (Crabb et al. 1989). A nuclear fragment approximately 650 bp long, encompassing the entire Adh1-12, was amplified in all six genera using two forward primers (2340-1 or EXON II-F) and two reverse primers (2340-11 or EXON III-R). PCR thermal profiles varied only slightly among genera examined: initial denaturation at 95°C for 2-10 min, followed by 25-30 cycles of denaturation at 95°C for 30-60 sec, annealing ramped from 52-53°C down to 46-48°C and back up to 52-53°C at a rate of 0.6°C/sec, extension at 73°C for 1.5 min, and a final extension cycle of 73°C for 4 min. Ramping speed between all three phases of PCR was set at a rate of 1°C/sec. Reaction concentrations (35 µl volume) included approximately 300 ng genomic DNA, 0.07 mM dNTPs, 2.86 mM MgCl₂, 3.5 µl 10X buffer, 0.286 µM primer, and 1.25-1.5 U enzyme (FailSafe PCR Enzyme Mix, Epicentre, Valencia, California). In some reactions, 1.5 U AmpliTaq Gold (PE Applied Biosystems, Foster City, California) was used with an initial denaturation time of 10 min.

PCR products were purified using the QIAquick PCR purification kit (Qiagen, Valencia, California) and sequenced with Big-Dye version 3.1 chain terminators (Applied Biosystems, Inc., Foster City, California). Appropriate external primers and four internal primers (350F, 350R, 410F, 410R; Table I) were used to sequence each strand entirely. Thermal profile for cycle sequencing was modified from the manufacturer’s recommendation: 25-30 cycles of 95°C for 30 sec denaturation, 50°C for 20 sec annealing, and 60°C for 3 min extension. Sequencing reactions were precipitated and concentrated with standard isopropanol methods, resuspended in 15 ml Hi-Di Formamide, and electrophoresed on an ABI 3100-Avant Genetic Analyzer (Applied Biosystems, Inc., Foster City, California). Sequencher 3.0 software (Gene Codes, Ann Arbor, Michigan) was

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>2340-I (5' PCR primer)</td>
<td>GTAATCAAGTGCAAAGCAGCTG</td>
</tr>
<tr>
<td>2340-II (3' PCR primer)</td>
<td>TAACCCACGTGGTTACTGAGCG</td>
</tr>
<tr>
<td>EXON II-F (5' PCR primer)</td>
<td>GTAATCAAGTGCAAAGCAGCRGCYRTGAGGAG</td>
</tr>
<tr>
<td>EXON III-R (3' PCR primer)</td>
<td>GACCTATCACCTGTGTTTACGCAAGTCCACCC</td>
</tr>
<tr>
<td>350F (internal sequencing primer)</td>
<td>GTGCTAAACATCTCTGATTTTGC</td>
</tr>
<tr>
<td>350R (internal sequencing primer)</td>
<td>GCCTTTTGAATCAAGATGTTTAG</td>
</tr>
<tr>
<td>410F (internal sequencing primer)</td>
<td>CTATAGCACAGCACG</td>
</tr>
<tr>
<td>410R (internal sequencing primer)</td>
<td>TGCTGCTGCTGCTAG</td>
</tr>
</tbody>
</table>
used to assemble and proof resultant fragments. Base calling ambiguities on single strands were resolved by choosing the call on the cleanest strand or by using appropriate IUB ambiguity code if both strands showed the same ambiguity.

Because provisional statements of homology (i.e., sequence alignment) are of special concern for non-coding DNA sequences possibly containing insertion/deletions (Giribet and Wheeler 1999), and because many parameters can affect multiple-sequence alignment and resulting phylogenetic inference (DeSalle et al. 1994; Hickson et al. 2000; Lutzoni et al. 2000; Wheeler 1995), two multiple-sequence alignments of Adh1-12 data were performed independently in Clustal X software (Thompson et al. 1997): one using a 30:4 gap cost-ratio, the other using a 5:4 gap cost-ratio (Van Den Bussche and Hoofer 2001; Van Den Bussche et al. 2002). Alignments subsequently were examined using MacClade software (version 4.05; Maddison and Maddison 2002), ambiguously aligned sites were delimited following methods of Hoofer and Van Den Bussche (2003), and analyses were performed with and without those sites.

Cyt-b data.—Sequences of the entire (1,143 bp) mitochondrial Cyt-b gene were obtained from each of the 41 species to facilitate comparison to the Adh1-12 data. PCR and sequencing primers, conditions, and thermal profiles followed Edwards and Bradley (2002) and Bradley et al. (2004a). Sequence alignments were performed manually and checked in MacClade software (version 4.05; Maddison and Maddison 2002), ambiguously aligned sites were delimited following methods of Hoofer and Van Den Bussche (2003), and analyses were performed with and without those sites.

Data analysis.—Analyses were performed in PAUP* 4.0b10 software (Swofford 2002) or MrBayes 2.01 software (Huelsenbeck and Ronquist 2001). Nucleotide positions were treated as unordered, discrete characters (A, C, G, and T), multiple states as polymorphisms, and gaps as missing. Nucleotide sequences from both Adh1-12 and Cyt-b were evaluated four ways: 1) base frequencies, number of transitions, number of transversions, and number of substitutions per 100 bp were estimated within and compared among the five ingroup genera; 2) levels of phylogenetic signal were estimated using the g-statistic (Hillis and Huelsenbeck 1992) for 100,000 randomly drawn trees; 3) genetic distances (uncorrected “p”) were obtained and compared using pairwise comparisons of taxa; and 4) maximum likelihood, Bayesian likelihood, and parsimony analyses were performed and compared among taxa.

Based on hierarchical likelihood ratio tests (hLRTs) in Modeltest software (Posada and Crandall 1998), the following models of nucleotide substitution and associated parameters best fit the data: Adh1-12—Hasegawa-Kishino-Yano (HKY) model with allowances for gamma distribution of rate variation (G), ti/tv = 2.308, πA = 0.301, πC = 0.191, πG = 0.183, πT = 0.325, α = 1.662; Cyt-b—general time reversible (GTR) with allowances for G and proportion of invariant sites (I), RAC = 1.310, RAG = 10.479, RCT = 2.473, RCG = 0.708, RCT = 32.137, πA = 0.375, πC = 0.343, πG = 0.079, πT = 0.203, α = 0.687, pinv = 0.456. The HKY + G model best fit the Adh1-12 data with and without ambiguous characters, although specific model parameters differed slightly; Adh1-12 values reported were calculated without ambiguous characters.

Bayesian analyses were run at least two million generations with four Markov-chains, random starting trees for each chain, and trees sampled every 100th generation. For each data set, two independent analyses were run to assess whether chains converged on the same posterior probability distribution and whether likelihood values became stable (Huelsenbeck et al. 2002). Model parameters were treated as unknown variables (with uniform priors) to be estimated in each Bayesian analysis (Leaché and Reeder 2002). Burn-in values (initial set of unstable generations to be ignored) were based on empirical evaluation of likelihoods converging on stable values. Clade reliabilities were assessed using posterior probabilities (values ≥ 0.95 regarded as significant).

Maximum likelihood analyses were performed with full heuristic searches, neighbor-joining starting trees, and tree-bisection-reconnection branch swapping. Parsimony analyses with all characters and substitution types given equal probabilities (i.e., unweighted), were conducted with full heuristic searches with 10 random additions, starting trees
by simple addition, and tree-bisection-reconnection branch swapping. Clade reliabilities were assessed in parsimony analyses using bootstrapping methods with 250 iterations (Felsenstein 1985). Due to prohibitive computation time under maximum likelihood, bootstrapping was performed for 100 iterations for just the Adh1-12 dataset (100 iterations). Values ≥ 70 were regarded as strong support.

Results

Adh1-12 data.—Complete sequence of Adh1-12 averaged 530 bp for the 41 rodents examined, ranging from 494 (Holochilus) to 576 (Peromyscus californicus). Alignment of sequences resulted in 614 aligned sites (125 ambiguously aligned) with the 5:4 gap cost-ratio, and 607 aligned sites (133 ambiguously aligned) with the 30:5 gap cost-ratio, corresponding to the insertion of more gaps with a lower gap cost ratio. As results from all subsequent analyses, phylogenetic analyses in particular, essentially were identical regardless of alignment and with and without the ambiguously aligned characters excluded, only results based on the 5:4 alignment with 125 ambiguous characters removed are reported and discussed.

After removing 125 ambiguous characters, 489 characters were available for analysis, of which 260 were constant and 126 were parsimony informative. Overall nucleotide frequencies varied slightly among the three genera (Table 2), averaging 30.47% (A), 17.51% (C), 17.94% (G), and 34.08% (T). The transition to transversion ratio was approximately 2.28 to 1. The number of heterozygous sites ranged from zero (37 taxa) to three (N. stephensi, N. mexicana, and O. perenensis), with a mean heterozygosity of 0.34 per taxon. The g1 statistic was skewed significantly left (-0.65; P < 0.01), indicating strong phylogenetic signal (Hillis and Huelsenbeck 1992).

Cyt-b data.—Complete sequences of Cyt-b (1,143 bp) were included for the 41 rodents examined herein. Sequence alignment was unequivocal and contained no stop codons. Of the 1,143 characters, 589 were constant and 475 parsimony informative. Nucleotide variation was distributed across codon positions as expected for protein-coding genes (1st position, 111; 2nd position, 37; 3rd position, 327). Overall nucleotide frequencies varied slightly among the three genera (Table 2), averaging 31.67% (A), 28.15% (C), 12.45% (G), and 27.74% (T), and transition to transversion ratio was approximately 2.68 to 1. The g1 statistic was skewed significantly left (-0.49; P < 0.01), indicating strong phylogenetic signal (Hillis and Huelsenbeck 1992).

Phylogenetic analyses.—For Adh1-12 and Cyt-b, Bayesian likelihoods reached stability before 100,000 generations (i.e., burn-in = 1,000), thinning the data to 19,000 sample points. Topology and posterior probabilities for nodes and model parameters for all sets of runs agreed. Maximum likelihood analysis resulted in a single best tree for both Adh1-12 (LnL = -2,718.63) and Cyt-b (LnL = -14,379.92) data sets (Figs. 1 and 2).

Table 2. Nucleotide base composition (A, C, G, and T), number of transition substitutions per 100 bp (#Ti/100 bp), number of transversion substitutions per 100 bp (#Tv/100 bp), and the uncorrected p distance (p Dist) for the 3 genera possessing multiple taxa. All values were averaged across taxa within each of the 3 genera for intron 2 of the alcohol dehydrogenase locus (Adh1-12) and the mitochondrial cytochrome-b gene (Cyt-b).

<table>
<thead>
<tr>
<th>Taxon/Sequence</th>
<th>% A</th>
<th>% C</th>
<th>% G</th>
<th>% T</th>
<th>#Ti/100 bp</th>
<th>#Tv/100 bp</th>
<th>p Dist</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peromyscus Adh1-12</td>
<td>30.16</td>
<td>17.48</td>
<td>18.58</td>
<td>33.77</td>
<td>2.31</td>
<td>1.09</td>
<td>3.41%</td>
</tr>
<tr>
<td>Peromyscus Cyt-b</td>
<td>31.72</td>
<td>27.45</td>
<td>12.62</td>
<td>28.21</td>
<td>8.38</td>
<td>2.79</td>
<td>12.29%</td>
</tr>
<tr>
<td>Neotoma Adh1-12</td>
<td>30.35</td>
<td>17.73</td>
<td>17.59</td>
<td>34.33</td>
<td>1.67</td>
<td>0.50</td>
<td>2.22%</td>
</tr>
<tr>
<td>Neotoma Cyt-b</td>
<td>32.24</td>
<td>28.93</td>
<td>12.43</td>
<td>26.40</td>
<td>8.73</td>
<td>2.90</td>
<td>11.61%</td>
</tr>
<tr>
<td>Oryzomys Adh1-12</td>
<td>31.34</td>
<td>17.33</td>
<td>16.97</td>
<td>33.37</td>
<td>4.48</td>
<td>2.10</td>
<td>6.70%</td>
</tr>
<tr>
<td>Oryzomys Cyt-b</td>
<td>31.16</td>
<td>28.43</td>
<td>11.98</td>
<td>28.43</td>
<td>7.73</td>
<td>5.75</td>
<td>14.00%</td>
</tr>
</tbody>
</table>
Parsimony analysis resulted in 62 most-parsimonious trees (length = 363, CI = 0.78, RI = 0.89) and a single most-parsimonious tree (length = 3,321, CI = 0.28, RI = 0.48) for the most-parsimonious tree (length = 3,321, CI = 0.28, RI = 0.48) for Neotominae. Based on the g statistic and consistency and retention indices, the Adh1-12 dataset had a greater overall phylogenetic signal and less homoplasy than the Cyt-b dataset. Adh1-12 sequences also provided statistically supported resolution to slightly more relationships than Cyt-b (Fig. 3). For example, Adh1-12 strongly supported all relationships examined at and above the genus level, including monophyly of Neotoma and Peromyscus and a sister relationship between Ototylomys and Tylomys, relationships that have been documented repeatedly by analysis of morphological and DNA sequence data (Carleton 1980; Bradley et al. 2004b; Edwards and Bradley 2002; Reeder and Bradley 2004, in press a and b). Adh1-12 supported a diphyletic origin for Oryzomys, a relationship also supported by Cyt-b and recent studies of nuclear markers (Myers and Tucker 1995; Smith and Patton 1993; Weksler 2003).

It is particularly noteworthy that, compared with Cyt-b, the Adh1-12 dataset provided supported resolution to more species-level relationships within Neotoma, Peromyscus, and Oryzomys (Fig. 3), despite fewer overall characters (489 Adh1-12, 1,143 Cyt-b) and fewer parsimony informative characters (126 Adh1-12, 464 Cyt-b). Furthermore, all but one of the species-level relationships were congruent with results from Cyt-b. The exception involved the sister species relationships of N. albigula, N. leucodon, and N. micropus. Adh1-12 analysis provided strong support for a sister relationship between N. albigula and N. leucodon (parsimony, 98%; maximum likelihood, 94%; Bayesian, 100%). Cyt-b analysis strongly supported a sister relationship between N. leucodon and N. micropus (parsimony, 86%; Bayesian, 100%). The difference, however, is that Adh1-12 support was based on three characters, whereas Cyt-b support was based on 22 characters. We view these particular Adh1-12 results with caution, and suggest additional study of nuclear markers will be necessary to help determine whether Adh1-12 results accurately reflect a conflict between the mitochondrial and nuclear genomes or represents the case of misleading inference of relationship due to inadequate variability.

Results of this study indicated that Adh1-12 DNA sequences are useful for addressing phylogenetic relationships within and among the Sigmodontinae and Neotominae. Based on the g statistic and consistency and retention indices, the Adh1-12 dataset had a greater overall phylogenetic signal and less homoplasy than the Cyt-b dataset. Adh1-12 sequences also provided statistically supported resolution to slightly more relationships than Cyt-b (Fig. 3). For example, Adh1-12 strongly supported all relationships examined at and above the genus level, including monophyly of Neotoma and Peromyscus and a sister relationship between Ototylomys and Tylomys, relationships that have been documented repeatedly by analysis of morphological and DNA sequence data (Carleton 1980; Bradley et al. 2004b; Edwards and Bradley 2002; Reeder and Bradley 2004, in press a and b). Adh1-12 supported a diphyletic origin for Oryzomys, a relationship also supported by Cyt-b and recent studies of nuclear markers (Myers and Tucker 1995; Smith and Patton 1993; Weksler 2003).

It is particularly noteworthy that, compared with Cyt-b, the Adh1-12 dataset provided supported resolution to more species-level relationships within Neotoma, Peromyscus, and Oryzomys (Fig. 3), despite fewer overall characters (489 Adh1-12, 1,143 Cyt-b) and fewer parsimony informative characters (126 Adh1-12, 464 Cyt-b). Furthermore, all but one of the species-level relationships were congruent with results from Cyt-b. The exception involved the sister species relationships of N. albigula, N. leucodon, and N. micropus, and perhaps represented a situation of misleading phylogenetic inference. Whereas Adh1-12 analysis provided strong support for a sister relationship between N. albigula and N. leucodon (parsimony, 98%; maximum likelihood, 94%; Bayesian, 100%), Cyt-b analysis strongly supported a sister relationship between N. leucodon and N. micropus (parsimony, 86%; Bayesian, 100%). The difference, however, is that Adh1-12 support was based on three characters, whereas Cyt-b support was based on 22 characters. We view these particular Adh1-12 results with caution, and suggest additional study of nuclear markers will be necessary to help determine whether Adh1-12 results accurately reflect a conflict between the mitochondrial and nuclear genomes or represents the case of misleading inference of relationship due to inadequate variability.

Overall, the Adh1-12 results are congruent with results from Cyt-b and other types of data, and generally are accompanied by high statistical support. Furthermore, multiple sequence alignment, which can be problematic with non-coding DNA sequences (Giribet and Wheeler 1999; Hoofer and Van Den Bussche 2003; Van Den Bussche et al. 2002), was not of particular concern in this study. Although we identified 125 ambiguously aligned characters in the 5:4 alignment and 133 in the 30:5 alignment, results from analysis of both alignments with and without the ambiguous characters were identical in topology and statistical support. Thus, most of the phylogenetic signal in the Adh1-12 dataset was associated with polymorphisms (nucleotide substitutions) rather than insertions/deletions. Yet, some insertion/deletions were informative. For example, a 5 bp (nucleotide positions 335-339) and a 4 bp (nucleotide positions 487-490) insertion/deletion were present in species of Oryzomys and Holochilus but absent in Neotoma and Peromyscus. Similarly, a 17
Figure 1. Maximum likelihood phylogram (LnL -2,718.627) from analysis of 5·4 gap cost-ratio alignment of complete Adh1-12 sequences (489 base pairs; 614 aligned sites minus 125 ambiguously aligned sites) using best-fit model (HKY + G, ti/tv = 2.308, πA = 0.301, πC = 0.191, πG = 0.183, πT = 0.325, α = 1.662). *Holochilus* was the designated outgroup. *N.* = *Neotoma*. *O.* = *Oryzomys*. *Ot.* = *Ototylomys*. *P.* = *Peromyscus*. *T.* = *Tylomys*. 
Figure 2. Maximum likelihood phylogram (LnL -14,379.922) from analysis of complete Cyt-b sequences (1,143 base pairs) using best-fit model (GTR + G + I; $R_{AC} = 1.310, R_{AG} = 10.479, R_{AT} = 2.473, R_{CG} = 0.708, R_{CT} = 32.137, \pi A = 0.375, \pi C = 0.343, \pi G = 0.079, \pi T = 0.203, \alpha = 0.687, \rho_{AC} = 0.456$). *Holochilus* was the designated outgroup. *N.* = *Neotoma*. *O.* = *Oryzomys*. *Ot.* = *Ototylomys*. *P.* = *Peromyscus*. *T.* = *Tylomys.*
Figure 3. Cladograms from phylogenetic analyses of complete Adh1-I2 (left) and Cyt-b (right) DNA sequences. *Holochilus* was the designated outgroup. In order, numbers above and below nodes are bootstrap proportions (250 iterations) from parsimony analysis, bootstrap proportions (100 iterations) from maximum likelihood analysis, and posterior probability proportions from Bayesian analysis. Maximum likelihood bootstrapping was feasible only for Adh1-I2 data. Only nodes with ≥ 70% bootstrap support or ≥ 0.95 posterior probabilities, or both, are shown. Vertical dotted line indicates sister relationship between *N. micropus* and clade of *N. floridana + N. magister*. *N.* = *Neotoma*. *O.* = *Oryzomys*. *Ot.* = *Ototylomys*. *P.* = *Peromyscus*. *T.* = *Tylomys*.
bp insertion/deletion (nucleotide positions 502-518) was present in *Neotoma* and *Peromyscus* but was absent in *Oryzomys* and *Holochilus*. These results are encouraging toward the goal of recovering reliable phylogenetic relationships at low and intermediate taxonomic levels from a nuclear intron. Further study of this intron, along with other nuclear and mitochondrial markers, should aid our understanding of the phyletic limitations of *Adh1-12* as well as organismal genealogy.

**ACKNOWLEDGMENTS**

Thanks to B. Dnâté Baxter, Nevin D. Durish, Michelle L. Haynie, Dallas D. Henson, Lindsey D. Porr, Andrew O. Stallings, Ryan R. Chambers, and two anonymous reviewers for providing helpful comments on previous versions of this manuscript. Thanks to Robert J. Baker for use of laboratory equipment and computers during various stages of this study. This research was supported by a grant from the National Institutes of Health (DHHS A141435-01 to RDB).

**LITERATURE CITED**


Specimens examined.—Collection localities, museum acronyms, and GenBank accession numbers are provided for each specimen examined in this study. Specimens are from the United States unless otherwise noted. Abbreviations for museum acronyms (in parentheses and to the left of the semicolon) follow Hafner et al. (1997): Abilene Christian University Natural History Collection (ACUNHC), Angelo State University Museum Natural History Collections (ASNUHC), Monte L. Bean Life Science Museum (BYU), Texas Cooperative Wildlife Collection (TCWC), Museum of Texas Tech University (TTU), The Museum of Southwestern Biology (MSB), and Universidad Nacional Autónoma de México (UNAM). If museum catalogue numbers were unavailable, specimens were referenced with the corresponding TK number (special number of the Museum of Texas Tech University). GenBank accession numbers (AF, AY, U, and DQ) for Adh1-12 and Cyt-b are provided in parentheses to the right of the semicolon, respectively.

Holochilus chacarius.—Paraguay: Departamento Pte. Hayes; Estancia Loma Pora, 23°33’.15’S, 57°34’.3’W (TTU104423; DQ227456, DQ227455).

Neotoma albigula.—New Mexico; Yuma Co., 3.7 km S, 5.6 km W Somerton, UTM 11 708569E 3608362N (TTU76451; AY817648, AF376477).

Neotoma cinerea.—Utah; San Juan Co., Owaehamo Bridge (MSB121427; AY817635, AF186799).

Neotoma floridana.—South Carolina; Richland Co., Congaree Swamp National Monument, 33°49’N, 80°50’W (MSB74955; AY817637, AF294335).

Neotoma fascipes.—California; Riverside Co., Rancho Capistrano (Ortega Mountains) (TTU81391; AY817632, AF376479).

Neotoma goldmani.—Mexico: Nuevo Leon; 1 km S Providencia (TTU45227; AY817656, AF186829).

Neotoma isthminca.—Mexico: Oaxaca; Las Minas, UTM 15 191165E 1824954N (TTU82665; AY817630, AF329079).

Neotoma lepida.—California; Orange Co., Irvine Lake, 1.3 km E Fremont Canyon on Lake View access road (TTU79131; AY817633, AF307835).

Neotoma leucodon.—Texas; Kerr Co., Kerr Wildlife Management Area, UTM 14 452336E 3330772N (TTU71198; AY817643, AF186806).

Neotoma magister.—Virginia; Madison Co., Shenendoah National Park, White Oak Canyon, 38°34’.36”N, 78°22’.30”W (MSB74952; AY817641, AF294336).

Neotoma microps.—Texas; Jeff Davis Co., Mount Livermore Preserve, UTM 13 579953E 3389871N (TTU101643; AY817645, AF294346).

Neotoma picta.—Mexico: Nuevo Leon; 1 km S Providencia (TTU45227; AY817656, AF186829).

Neotoma picta.—California; Orange Co., Irvine Lake, 1.3 km E Fremont Canyon on Lake View access road (TTU79131; AY817633, AF307835).

Oryzomys albigularis.—Ecuador: Pichincha; 10 km NW Quito, Tandayapa Valley, 0°00’.13”N, 78°40’.70”W (ACUNHC917; DQ207945, DQ224407).

Oryzomys albigularis.—Nicaragua: Selva Negra; Altajo Trail (TTU101644; DQ207950, DQ224410).
Appendix I (cont.)

Oryzomys cousei. — HONDURAS: Olancho; 4 km E Catacamas (Escuela de Sembrador), UTM 16 624523E 1637511N (TTU84697; DQ207948, DQ185383).

Oryzomys melanotis. — HONDURAS: Atlantida; Lancetilla Botanical Gardens, UTM 16 451012E 1740282N (TTU84374; DQ207947, DQ224409).

Oryzomys palustris. — Texas; Galveston Co., Texas City, Virginia Point (TTU82920; DQ207949, DQ185382).

Oryzomys perenensis. — PERU: Loreto; Maynas, 25 km S Iquitos (Estacion Biologia Allpahuayo) (TTU98606; DQ207946, DQ224408).

Ototylomys phyllotis. — HONDURAS: Atlantida; Lancetilla Botanical Gardens, UTM 16 451012E 1740282N (TTU84374; AY817624, no Cyt-b data); MEXICO: Quintana Roo, 1 km N Noh-bee (ASNH7254; no Adh1-I2 data, AY009788).

Peromyscus attwateri. — Oklahoma; McIntosh Co., 5 km E Dustin (TTU55688; AY817626, AF155384).

Peromyscus beateae. — MEXICO: Chiapas, Yalentay, UTM 15 524171E 1852486N (UNAM catalogue number unavailable, TK93279; AY994223, no Cyt-b data); Veracruz; Xometla (TCWC48060; no Adh1-I2 data, AF131921).

Peromyscus boylii. — California; San Diego Co., Heise County Park (TTU83102; AY994225, no Cyt-b data); MEXICO: Aguascalientes, Rincon de Romos (TCWC48438; no Adh1-I2 data, AF131924).

Peromyscus californicus. — California; Los Angeles Co., Chatsworth Reservoir Park (TTU83292; AY994211, AF155393).

Peromyscus crinitus. — Utah; Emery Co., Cottonwood Canyon, 39°16′51.8″N, 111°10′31.9″W (BYU18639; AY994213, AY376413).

Peromyscus difficilis. — MEXICO: Tlaxcala; 2 km NE Tepeitila (TTU82690; AY994219, AY376416).

Peromyscus eremicus. — California; Los Angeles Co., Kanan Dome Road (TTU81850; AY994212, no Cyt-b data); California; Los Angeles Co., Calabasas Creekside Park (TTU83249; no Adh1-I2, AY322503).

Peromyscus gratiss. — MEXICO: Michoacán; 4 km E Costazos (UNAM catalogue number unavailable, TK46354; AY994218, AY376421).

Peromyscus hylocetes. — MEXICO: Michoacán; Estacion Cerra Burror, Microadas; 3,270 m (UNAM catalogue number unavailable, TK45309; AY994235, DQ000481).

Peromyscus leucopus. — Texas; Presidio Co., Las Palomas Wildlife Management Area, UTM 13 529481E 3321292N (TTU75694; AY994240, no Cyt-b data); Dickens Co., 0.9 km E Afton (TTU catalogue number unavailable, TK47506; no Adh1-I2 data, AF131926).

Peromyscus levipes. — MEXICO: Michoacán, Las Minas, 3 km SW Tuxpan (UNAM catalogue number unavailable, TK47819; AY994224, DQ000477).

Peromyscus manicolatus. — Arkansas; Mississippi Co., Dillahunty Pecan Orchard (TTU17830; AY994242, no Cyt-b data); Alaska, Alexander Archipelago, Bushy Island, Petersburg Quad, 56°15′45″N, 132°58′52″W (UAM50770; no Adh1-I2 data, AF119261).

Peromyscus megalops. — MEXICO: Guerrero; 6.4 km SSW Filo de Caballo (TTU82712; AY994217, DQ000475).

Peromyscus melanophrys. — MEXICO: Durango; 2.2 km S, 2.5 km E Vicente Guerrero (TTU75509; AY994216, AY322510).

Peromyscus mexicanus. — MEXICO: Chiapas; 14.4 km N Ocozocoautla (TTU82759; AY994236, AY376425).
Appendix I (cont.)

*Peromyscus pectoralis.*—MEXICO: Jalisco; 30 km W Huejuquilla del Alto (UNAM catalogue number unavailable, TK48645; AY994221, no Cyt-b data; TTU75575; no Adh1-12 data, DQ000476).

*Peromyscus perfulvus.*—MEXICO: Michoacán; Tunel de Riego, 2 km E Cerro Colorado, 1290 M, 19°19'220"N, 100°28'308"W (UNAM catalogue number unavailable, TK47926; AY994215, DQ000474).

*Peromyscus schmidlyi.*—MEXICO: Durango; 6.1 km W Coyotes, Hacienda Coyotes, UTM 13 465908E 2634281N (TTU81617; AY994228, AY322524).

*Peromyscus spicilegus.*—MEXICO: Michoacán; KM 81 carr., Arío de Rosales-La Huacana, 1602 m, 19°10'59"N, 101°43'42"W (UNAM catalogue number unavailable, TK 47888; AY994232, DQ000480).

*Tylomys nudicaudatus.*—GUATEMALA: Izabal, Cerro San Gil (TTU62082; AY817625, AF307839).