J Mol Evol (1999) 48:369–379

JOURNAL OF MOLECULAR EVOLUTION

Multiple Substitutions Affect the Phylogenetic Utility of Cytochrome *b* and 12S rDNA Data: Examining a Rapid Radiation in Leporid (Lagomorpha) Evolution

Kenneth M. Halanych,* Terence J. Robinson

Department of Zoology and Entomology, University of Pretoria, Pretoria, South Africa

Received: 23 April 1998 / Accepted: 14 May 1998

Abstract. Partial sequences of two mitochondrial genes, the 12S ribosomal gene (739 bp) and the cytochrome *b* gene (672 bp), were analyzed in hopes of reconstructing the evolutionary relationships of 11 leporid species, representative of seven genera. However, partial cytochrome *b* sequences were of little phylogenetic value in this study. A suite of pairwise comparisons between taxa revealed that at the intergeneric level, the cytochrome *b* gene is saturated at synonymous coding positions due to multiple substitution events. Furthermore, variation at the nonsynonymous positions is limited, rendering the cytochrome *b* gene of little phylogenetic value for assessing the relationships between leporid genera. If the cytochrome *b* data are analyzed without accounting for these two classes of nucleotides (i.e., synonymous and nonsynonymous sites), one may incorrectly conclude that signal exists in the cytochrome *b* data. The mitochondrial 12S rRNA gene, on the other hand, has not experienced excessive saturation at either stem or loop positions. Phylogenies reconstructed from the 12S rDNA data support hypotheses based on fossil evidence that African rock rabbits (*Pronolagus*) are outside of the main leporid stock and that leporids experienced a rapid radiation. However, the molecular data suggest that this radiation event occurred in the mid-Miocene several millions of years earlier than the Pleistocene dates suggested by paleontological evidence.

Key words: Leporidae — Rabbits — Cytochrome *b* — 12S rRNA — Lagomorph evolution — Saturation — Phylogenetics

Introduction

The Leporidae (Lagomorpha) has been extensively studied from an ecological, taxonomic, and, to a lesser extent, paleontological perspective, yet relatively little is known about its evolutionary history (Dawson 1958, 1981; Hibbard 1963; Myers and MacInnes 1981). The Lagomorpha comprises two recognized families, the Ochotonidae and the Leporidae. The ochotonids (pikas) consist of one genus, *Ochotona*, with 25 species, whereas the Leporidae (hares and rabbits) includes 11 recognized extant genera (approximately 54 species), 8 of which are monotypic (Hoffman 1993; Chapman and Flux 1990). *Lepus* (jackrabbits and hares), comprising ~30 species, and *Sylvilagus* (cottontails), with ~13 species, are the only two species-rich genera within the clade (*Pronolagus* contains 3 species). This lack of diversity provides sharp contrast to the evolutionary radiations within the Rodentia, with which the lagomorphs have been traditionally allied as members of the Glires cohort (Li et al. 1987; Novacek et al. 1988; Novacek 1990, 1993; Meng et al. 1994; for an opposing view see Easteal 1990). Graur and co-workers' (1996) recent suggestion that

* *Present address:* Biology Department, MS 33, Woods Hole Oceanographic Institution, Woods Hole, MA 02543-1049, USA
Correspondence to: K.M. Halanych; *e-mail:* khalanych@whoi.edu

Table 1. The taxa employed in the study and their GenBank accession numbers[a]

Scientific name	Common name	GenBank Accession No.	
		12S rDNA	Cyt *b* GenBank
Brachylagus idahoensis	Pygmy rabbit	U58921	U58930
Bunolagus monticularis	Riverine rabbit	U58922	U58931
Lepus americanus	Snowshoe hare	U58923	U58932
Lepus californicus	Black-tailed jackrabbit	U58924	U58933
Lepus capensis	Cape hare	U58925	U58934
Oryctolagus cuniculus	European rabbit	X54172[b] & U59264	U07566[c]
Pronolagus crassicaudatus	Greater red rabbit	U31044[d]	U58935
Romerolagus diazi	Volcano rabbit	U58926	U58936
Sylvilagus aquaticus	Swamp rabbit	U58927[e]	U58937
Sylvilagus audubonii	Audubon's cottontail	U58928[e]	U58938
Sylvilagus floridanus	Eastern cottontail	U58929[e]	U58939
Ochotona princeps	American pika	U31043[d]	U58940

[a] If no reference is given, then the sequence was determined in this study.
[b] From Mignotte et al. (1990).
[c] From Irwin and Árnason (1994).
[d] From Whiteford (1995).
[e] From Halanych and Robinson (1997).

lagomorphs are allied to primates has been refuted (Halanych 1998). Studying the phylogenetic relationships within the Leporidae will allow future workers to examine the evolutionary mechanisms affecting lagomorph diversification.

Most modern generic treatments of the Leporidae (Corbet 1983; Chapman and Flux 1990) rely heavily on Dawson's (1958, 1981) and Hibbard's (1963) interpretation of the limited paleontological evidence. Existing cytogenetic data, which are available for eight extant genera (Robinson 1980; Robinson et al. 1981, 1983a, b, 1984; van der Loo et al. 1981; Robinson and Skinner 1983), lack informative characters with which to infer phylogeny. The absence of synapomorphies in the paleontological and the cytogenetic data suggests that many leporid genera (including *Lepus* and *Sylvilagus*) originated nearly instantaneously during a single brief radiation.

The purpose of this study was to analyze the generic relationships among leporids by reconstructing their evolutionary history and to shed light on factors which may have influenced their diversification. To this end, sequence data from two mitochondrial genes, cytochrome *b* and the 12S ribosomal subunit genes, were employed. Although the cytochrome *b* data were ineffective at this taxonomic level of phylogenetic analysis, the 12S rDNAs contain some phylogenetic signal which helps elucidate leporid evolutionary history. Parsimony, neighbor-joining, and maximum-likelihood methods were used to clarify the phylogenetic relationships among 11 species, representing 7 of the 11 recognized leporid genera. Like the paleontological and cytogenetic data, the mitochondrial sequence data suggest that these mammals experienced a rapid diversification event.

Materials and Methods

Data Collection. Partial sequences for the 12S rRNA and cytochrome *b* mitochondrial genes were determined for most of the taxa incorporated in this study. Table 1 gives the original reference for each sequence and the corresponding GenBank accession number. The *Oryctolagus cuniculus* sequence is available for the entire cytochrome *b* gene (Irwin and Árnason 1994), but only part of the 12S rRNA gene (Mignotte et al. 1990) is published. Therefore, the *Oryctolagus* 12S rDNA sequence employed in the present investigation is a composite derived from two individuals—that published by Mignotte et al. (1990) and a new sequence presented here (see Table I). The aligned data sets, which can be obtained from the authors via e-mail on request, have been deposited in TreeBase (http://phylogeny.harvard.edu/treebase).

The availability of tissue from endangered species has limited the scope of our study to leporid taxa from Africa, Europe, and the New World. The three Asian genera (*Caprolagus*, *Pentalagus*, and *Nesolagus*) which are absent from this investigation are all endangered; the only other missing genus, *Poelagus*, is restricted to central Africa, including Uganda and Rwanda (Hoffman 1993), and was not available.

Total genomic DNA was extracted from cultured fibroblast cells using standard protocols (Maniatis et al. 1982). The oligonucleotide primers developed by Pääbo and Wilson (1988) and Kocher et al. (1989) were used to amplify an approximately 760-bp region of the cytochrome *b* gene (primer pair L14724, H15494), and those of Allard and Honeycutt (1992) were used to amplify approximately 820 bp of the 12S ribosomal gene (primer pair A and D). Standard PCR techniques were employed (Hillis et al. 1990). The cycling parameters were 3 min at 95°C for initial denaturation (1 cycle), 45 s at 94°C for denaturation, 45 s at 50°C for annealing of oligonucleotides, 60 s at 72°C for extension (35 cycles), and 5 min at 72°C for the final extension.

Amplified fragments were purified by agarose gel electrophoresis. The use of biotinylated primers during PCR allowed for the capture of single-stranded DNA for sequencing (Sequenase v2.0; US Biochemical). Sequencing was done using the PCR oligonucleotide primers and primers L14841, L15162, and H15149 [for cytochrome *b* (Pääbo and Wilson 1988; Kocher et al. 1989)] and primers B and C [for the 12S rDNA (Allard and Honeycutt 1992)]. Fragments were visualized using 6% acrylamide gel electrophoresis and standard audioradiographic techniques (Hillis et al. 1990).

Table 2. The phylogenetic information in the cytochrome *b* and 12S rDNA data

	Base pairs	Variable characters	% variable characters	Informative characters	% informative characters	G_1 statistic[a]
Cytochrome *b*						
All positions	672	262	38.9	184	27.4	−0.619
1st & 2nd positions	448	59	13.2	28	6.3	−0.426
1st position	224	42	18.8	23	10.2	−0.348
2nd position	224	17	7.6	5	2.2	−0.563
3rd position	224	203	90.6	156	69.6	−0.621
12S rDNA						
All positions	714[b]	217	30.4	85	11.9	−0.510
Stems	339	83	24.5	26	7.7	−0.614
Loops	375	134	35.7	59	15.7	−0.339

[a] All values are significant; see Hillis and Huelsenbeck (1992).
[b] There were 739 total nucleotide positions for the 12S rDNA data, but only 714 of these could be unambiguously aligned.

Phylogenetic Analyses. The alignments for both the 12S rDNA and the cytochrome *b* data were produced with the multiple alignment program Clustal W (Thompson et al. 1994) and corrected by hand for obvious alignment errors. Additionally, the protein coding sequence was used to verify the cytochrome *b* alignment and the secondary structure model of Springer and Douzery (1996) was used to check the 12S rDNA alignment. Regions that could not be unambiguously aligned were excluded from the analyses (in this case, only 25 positions in the 12S rDNA sequence were excluded). The boundaries of the excluded regions were trimmed back to the last uninformative character as in Halanych (1996).

The PAUP software package, Version 3.1.1d25 (Swofford 1993), was used for parsimony analyses, and the PHYLIP software package, Version 3.5 (Felsenstein 1993), was used for neighbor-joining and maximum-likelihood estimates. MacClade Version 3.0 (Maddison and Maddison 1992) was employed to determine various character statistics and tree lengths. Because we can never know with certainty under which transition (T_i)–transversion (T_v) ratio the leporid sequences evolved, we chose an approach which builds a consensus from trees reconstructed under a variety of T_i/T_v weighting schemes. This methodology is admittedly conservative but has the advantage of providing a qualitative estimate of how robust topological features are under different assumptions of nucleotide evolution. For the analyses, the ratios of 1:1, 2:1, 3:1, and 10:1 were used to weight transversions over transitions. Although these values were chosen arbitrarily, the empirical ratios for both the cytochrome *b* data and the 12S rDNA data were determined with the ''state changes and stasis'' option of MacClade by counting the average number of transition and transversion events on 100 random trees (Halanych 1996; see also Halanych and Robinson 1997). Although this method provided an empirical estimate, as sequences become saturated with substitutions, the T_i/T_v value will approach 1.

The Kimura two-parameter model with a Γ correction was chosen for neighbor-joining and likelihood analyses because it represents the simplest model which accounts for the observed nucleotide parameters. For example, below we show that the T_i/T_v ratio is not equal to 1:1 (thus eliminating a Jukes–Cantor model) but also show that the nucleotide frequencies for the 12S data are close to equal (rendering the HKY model not needed). The Γ distribution was used to accurately model for the among-site rate heterogeneity known to exist in mammalian 12S rRNA (e.g., Sullivan et al. 1995).

All nonparametric (traditional) bootstrap values reported here are based on 200 iterations, and in the case of parsimony bootstraps, the general heuristic search algorithm with TBR branch swapping was employed. No gaps were present in the cytochrome *b* data, but the 12S data contained nine gaps at 11 characters, which were treated as missing data in the analyses. All other gaps within the ingroup were contained in the regions which could not be unambiguously aligned and were therefore excluded from the analyses.

Simulations for the parametric bootstrap used the Siminator program (Huelsenbeck et al. 1996). This program evolves sequences under an HKY model for which the T_i/T_v model (κ), nucleotide frequencies, and among-site rate heterogeneity (Γ) can be modeled. For a further description of parametric bootstrapping see Huelsenbeck et al. (1996).

Results and Discussion

12S rDNA

The final alignment of the 12S rDNA data included 739 positions, of which 714 were unambiguously aligned; of these, 217 were variable and 85 were informative (i.e., parsimony characters; see Table 2). The g_1 statistic for 10^4 randomly generated trees ($g_1 = -0.510$) indicated that the data contained significantly more signal than random (Hillis and Huelsenbeck 1992). However, most of the phylogenetic information was apparently at the tips of the tree since the deletion of taxa to reduce the more recent clades (i.e., the *Sylvilagus* and *Lepus* clades) increased the g_1 to near-zero (−0.037), indicating that little signal was retained. An analysis of nucleotide composition revealed that the data had more adenines (36.6%) and fewer guanines (17.6%) than expected at random. Because the bias was similar across lineages (adenine's range, 34.7–37.7%; cytosine's range, 22.2–25.7%; guanine's range, 16.8–19.0%; thymine's range, 21.5–23.6%), it presumably did not greatly influence the phylogenetic analyses. The basis in base composition, as calculated by Irwin et al. (1991), was 0.1541. Furthermore, these values are comparable to those reported in other mammalian taxa (Springer et al. 1995; Matthee and Robinson 1997).

For the 12S rDNA data, the empirically determined T_i/T_v value of 2.2:1 falls within the weighting ratios employed. To reduce computation time, this value was rounded off to 2:1. The Γ parameter employed (calculated value = 0.531768, rounded to 0.5) is based on an empirical estimation (Sullivan et al. 1995).

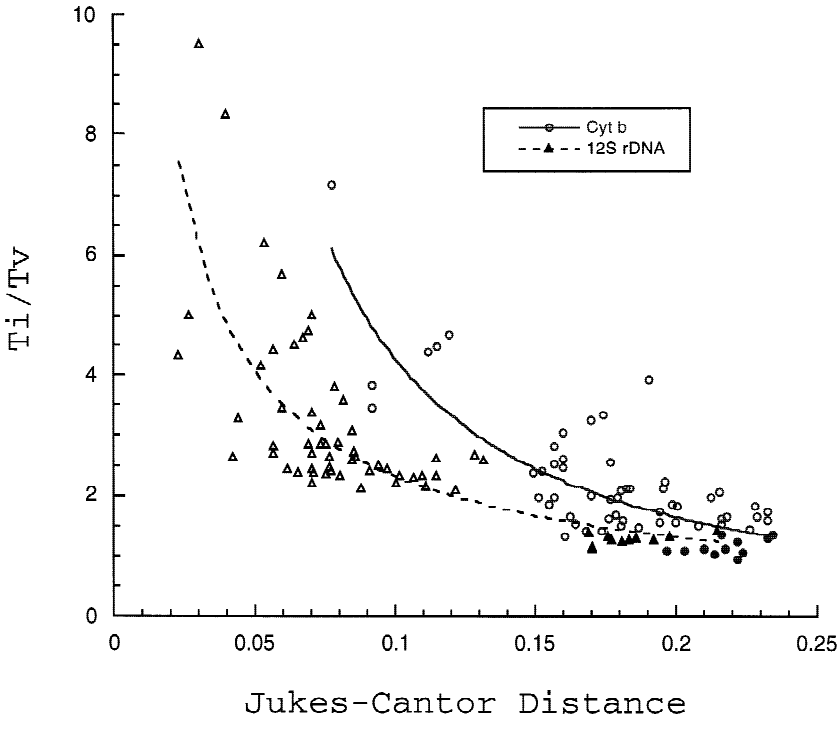

Fig. 1. Saturation plot for each gene. The T_i/T_v ratio was plotted against the Jukes–Cantor distance of all pairwise comparisons of taxa. (○) Cytochrome *b* data; (△) 12S rDNA data. *Open symbols* are pairwise comparisons within the ingroup (leporid/leporid), whereas *filled symbols* are ingroup-to-outgroup comparisons (leporid/ochotonid).

Cytochrome b

The cytochrome *b* data set consisted of 672 nucleotides for which the codon positions were determined. Table 2 shows the distribution of phylogenetic information in the cytochrome *b* gene according to codon position. Clearly, third positions have experienced saturation, given that 90.6% of these sites were variable and 69.6% were ''informative.'' Interestingly, the g_1 was −0.619 when all three positions were included but −0.426 when only the first two were considered. Both of these values are significant as judged by the cutoff values given by Hillis and Huelsenbeck (1992). Although the g_1 indicated that the third position added signal to the data despite its high variability, this increase in signal can be accounted for by the high rate of third-position transitions among the terminal taxa (i.e., the information gained within the *Sylvilagus* or *Lepus* lineages). As with the 12S data, the exclusion of certain taxa showed that most of the phylogenetic signal is at the tips of the tree; the g_1 for all positions with excluded taxa was −0.065.

The analysis of the nucleotide composition of the cytochrome *b* gene revealed fewer guanines than expected at random. The degree of bias was, however, dependent upon the codon position: 24.2% guanine in the first position, 14.8% in the second, and 3.6% in the third. The base composition basis for these positions was 0.0525, 0.1943, and 0.3396, respectively. Similar values have been reported for other mammal groups (Irwin et al. 1991, Matthee and Robinson 1997). The empirical T_i/T_v estimate yielded a value of 1.8:1 when all positions were considered.

Multiple Substitutions and Saturation

Because of the effects of multiple substitutions at the third codon position (Brown et al. 1982; Irwin et al. 1991), we decided to make a closer examination of the data to determine if saturation had destroyed phylogenetic signal. In order to assess saturation, all pairwise comparisons of the T_i/T_v ratio for both cytochrome *b* and 12S rDNA data were plotted against the Jukes–Cantor (JC) distances based on either cytochrome *b* or 12S, respectively (Fig. 1). We used the JC model to allow for easy comparisons to published analyses (e.g., Adkins and Honeycutt 1994; Yoder et al. 1996), but the same conclusions are reached with a more complex correction models.

Perhaps the most noticeable feature is that the cytochrome *b* pairwise comparisons and 12S pairwise comparisons cluster in different regions of the graph space. This finding is expected because, despite occurring on the same genetic locus, these two genes have different modes and tempos of evolution (as demonstrated by Janczewski et al. 1995). The slower substitution rate of the 12S rDNA, and thus smaller genetic distances, is evident in Fig. 1. Several researchers (Graybeal 1993; Adkins and Honeycutt 1994; Meyer 1994) show that a mitochondrial protein-coding gene becomes saturated at the third position around JC distances of 0.15 to 0.20, which is where the cytochrome *b* data cluster in the graph space.

One would expect that in the absence of substitutional saturation at the nucleotide level, pairwise comparisons between the ochotonid outgroup and the leporid taxa

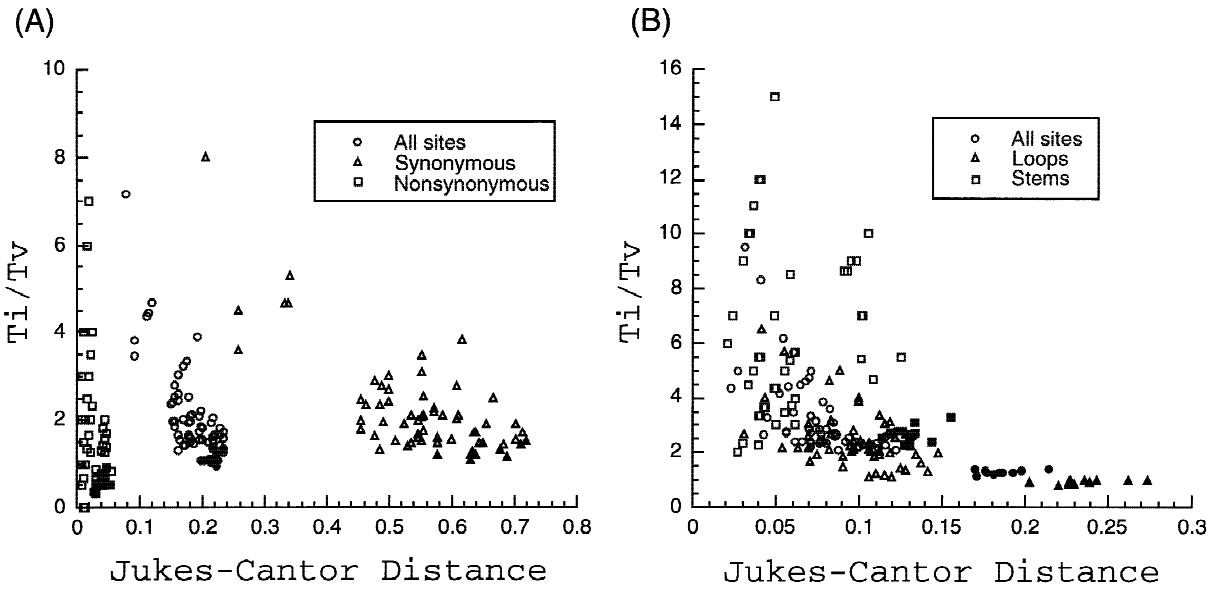

Fig. 2. Saturation plot for nucleotide classes within the genes studied. The T_i/T_v ratio was plotted against the Jukes–Cantor distance of all pairwise comparisons of taxa. **(A)** Plots of the different classes of cytochrome *b* data: (○) all sites; (△) synonymous sites; (□) nonsynonymous sites. **(B)** Plots of the different classes of 12S rDNA data: (○) all sites; (△) sites in loops; (□) sites in stems. *Open symbols* are pairwise comparisons within the ingroup (leporid/leporid), whereas *filled symbols* are ingroup-to-outgroup comparisons (leporid/ochotonid).

should be noticeably greater than comparisons between leporid genera. The separation between leporids and ochotonids [30–40 MYA (Dawson 1981)] is 3–20 times longer than the separation between leporid genera, depending on the estimation employed [conventionally 1–2 MYA (Dawson 1981) or the RFLP estimate of 6–8 MYA (Biju-Duval et al. 1991)]. In Fig. 1, the ochotonid/leporid comparisons (denoted by filled symbols on the graph) for the 12S data group as a discrete cluster around distances of 0.17 to 0.20. However, cytochrome *b* data do not noticeably resolve the ochotonid/leporid comparisons from the leporid/leporid comparisons (open symbols), thus indicating saturation.

Although saturation has occurred in the cytochrome *b* gene, presumably at synonymous sites, this begs the question whether there is any meaningful phylogenetic signal present in the cytochrome *b* data that would allow an assessment of leporid intergeneric relationships. To answer this question, we coded nucleotide sites within the cytochrome *b* fragment as either synonymous (all third-position and all first-position leucine sites) or nonsynonymous (all second-position and all nonleucine first-position sites) and again compared the T_i/T_v ratio to the genetic distance. Figure 2A shows T_i/T_v ratios of the nonsynonymous, the synonymous, and all cytochrome *b* sites plotted against the JC distance values obtained for each of these three categories of cytochrome *b* data. The comparisons based on synonymous positions cluster on the lower far right of the graph, with low T_i/T_v ratios and high divergence values. The only values that have JC distances of <0.4 are the six pairwise comparisons within either the *Lepus* or the *Sylvilagus* genera. In contrast, the nonsynonymous changes have very low divergence values, <0.05, suggesting that there is little phylogenetic information. This suspicion is confirmed if one considers that only 10 parsimony informative characters are observed within the ingroup for 420 nonsynonymous positions. Again, in neither case were the ochotonid/leporid comparisons clearly distinct from the leporid/leporid comparisons.

Clearly, the analysis of all sites together can be misleading; in fact, the problem of saturation is much more severe than is evident when all of the cytochrome *b* data are examined together. When treated as a single class of data, the rates of substitution in the synonymous and nonsynonymous sites yield overall average divergence values that may not accurately reflect the evolution of the gene. The same phenomenon at the amino acid level has been long known (Fitch and Margoliash 1967). This situation may cause workers to infer incorrectly, based on genetic distance values, that significant saturation has not occurred (see Palumbi 1989). For example, synonymous changes may be saturated, but the low variability of nonsynonymous positions may cause the overall distance measurements to average out to less than 0.15–0.20. The degree to which this situation causes error in phylogenetic analyses, for both cytochrome *b* and other protein-coding genes, remains to be seen.

Traditionally, the cytochrome *b* gene has been a valuable genetic marker for several levels of phylogenetic analysis and has been used to resolve evolutionary events from <50 MYA (e.g., Meyer and Wilson 1990; Edwards et al. 1991; Irwin et al. 1991; Irwin and Árnason, 1994) to very recent intraspecific evolutionary events (e.g., Moritz et al. 1987; Zang and Ryder 1995). However, the present data suggest that there is a window within the

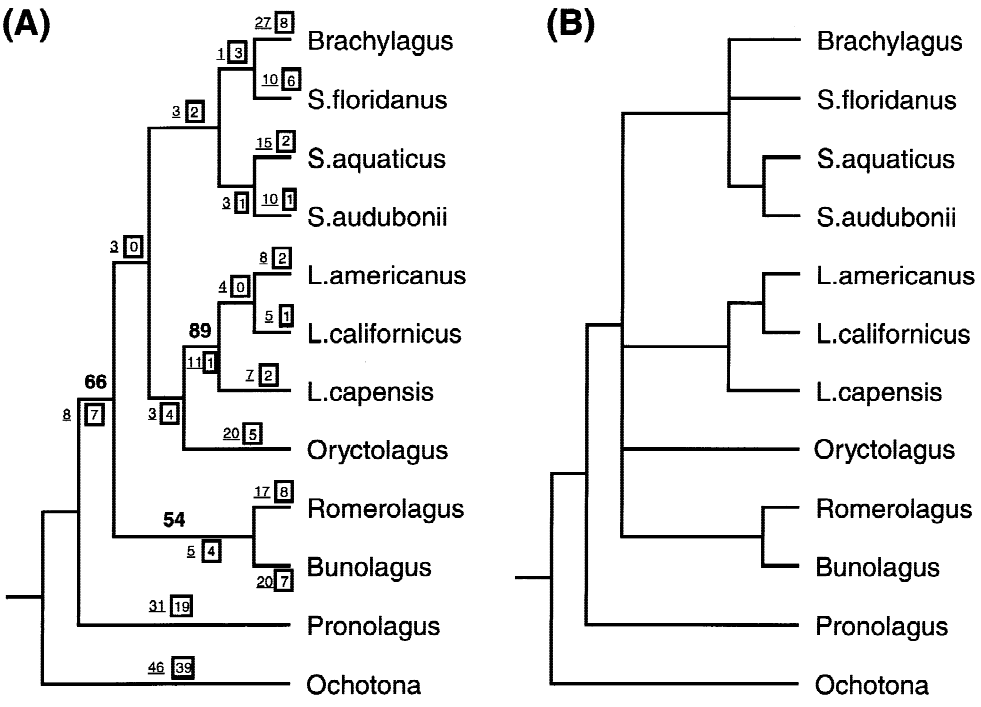

Fig. 3. Results of the parsimony analyses based on the 12S rDNA sequences. **A** The most-parsimonious tree obtained when transversions were weighted twice transitions. The tree was produced with the heuristic search program of PAUP (Swofford 1993) and is 491 steps in length (see text for details). Bootstrap values obtained from 200 iterations using a heuristic search are shown *on the branches* in *larger boldface text.* If no value is shown, the node was supported in <50% of the iterations. The numbers *underlined along the branch* are the number of transitional changes, and the *boxed numbers* are the number of transversional changes along the given branch. **B** The strict consensus of all the most-parsimonious trees obtained under the four transversion weighting schemes (1:1, 2:1, 3:1, and 10:1). This tree is based on 12S rDNA data.

time scale of resolution offered by the cytochrome *b* gene in which the synonymous positions have become saturated with nucleotide substitutions (obliterating phylogenetic signal), and nonsynonymous changes have not accumulated in sufficient numbers for meaningful phylogenetic analyses. This situation, which was also described in bufonid frogs (Graybeal 1993), is likely to occur in similar genes such as the mitochondrial COI or COII genes.

In contrast, the 12S data provide a strikingly different result. The 12S rDNA data was coded as either ''stem'' or ''loop'' (i.e., base-pairing or non-base-pairing, respectively) following the model of Springer and Douzery (1996). The T_i/T_v versus genetic distance plot (Fig. 2B) reveals that both the stem and the loop regions are representative of the 12S data taken as a whole. Moreover, the wide distribution of divergence values, and the clear separation between the ochotonid/leporid and leporid/leporid comparisons, suggests that saturation is not a problem in this particular instance. For some of the pairwise comparisons of the 12S stem data, there were no transversional differences (data not shown on plot). Because of the confounding effects of saturation, we opted to exclude the cytochrome *b* data from further phylogenetic analyses (except for a combined data analysis; see below), as it would only serve as a source of homoplasy. Consequently, subsequent phylogenetic conclusions are based only on the 12S data.

Phylogenetic Analyses

Figure 3A shows the single best tree of 491 steps obtained by parsimony analysis (using the general heuristic search algorithm in PAUP with TBR branch swapping and 100 random addition replicates) for the 12S rDNA data when transversions are weighted twice transitions (i.e., the empirical value). The single tree produced from an equal T_i/T_v weighting had 378 steps and a CI of 0.714 (CI excluding uninformative characters = 0.526). Although the parsimony analyses of the 12S data indicate that the topology was dependent on the T_i/T_v weighting employed, several consistent results were evident across all weighting schemes. The tree shown in Fig. 3B is a strict consensus of all the most parsimonious trees generated under the four weighting schemes employed in our investigation. Although the position of *Oryctolagus* is ambiguous Fig. 3, it clustered with *Lepus* in all 12S rDNA analyses except when the equal weighting option (T_i/T_v = 1:1) was employed. When transversions are weighted 10 times transitions, a *Lepus–Oryctolagus* association is supported in 74% of the bootstrap iterations. In general, the bootstrap support was weak, especially when transversions were not weighted heavily. The node defining *Lepus* monophyly was the only one consistently supported with high bootstrap values (87–94%) across T_i/T_v weighting schemes.

The neighbor-joining analyses were run with the

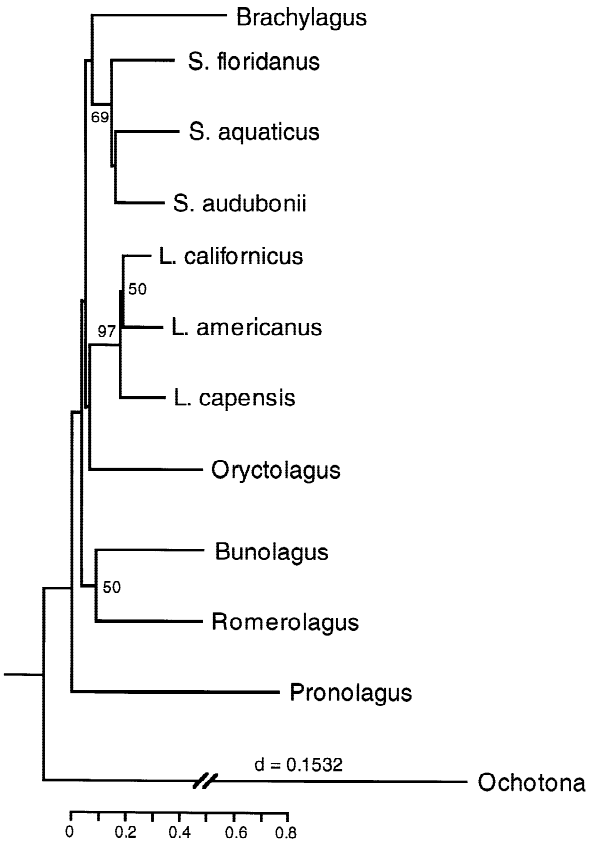

Fig. 4. The neighbor-joining topology produced using a Kimura two-parameter model with an empirically estimated Γ correction ($\alpha = 0.5$) to account for among-site rate variation and weighting transversions twice transitions. The tree obtained is based on the 12S rDNA sequence data. The scale at the bottom shows divergence along the branches as estimated with a Kimura two-parameter model with a Γ correction; the tree is zeroed at the most basal leporid node. Neighbor-joining bootstrap values (200 iterations) are shown next to the relevant node unless the node was supported in <50% of the bootstrap iterations.

DNAdist and Neighbor programs of PHYLIP and employed a Kimura two-parameter model with a correction for among-site nucleotide rate variation (Jin and Nei 1991) based on the empirically estimated Γ distribution ($\alpha = 0.5$). The tree produced when a T_i/T_v ratio of 2:1 was used is shown in Fig. 4. An analysis using a Kimura two-parameter model without the Γ correction and the maximum-likelihood analysis based on equal weighting produced the same topology. For maximum likelihood (using the DNAml program, PHYLIP, with 10 iterations of the jumble option), all other T_i/T_v ratios caused *Brachylagus* to cluster with *S. floridanus,* and not outside of the *Sylvilagus* clade, thus producing the same tree as the parsimony analysis when transversions are weighted twice transitions (Fig. 3A).

However, maximum-likelihood analyses employing two rate categories (means determined empirically = 0.14 and 1.88) to more accurately model the different rates of substitutions in stems and loops, found unorthodox trees. For all T_i/T_v ratios, the three *Sylvilagus* species are paraphyletic at the base of the reconstructed topologies (trees not shown; similar results were found when four rate categories were employed). Because abundant morphological (Dawson 1958, 1981; Hibbard 1963; Corbet 1983; Chapman and Flux 1990; Chapman and Ceballos 1990) and cytogenetic (Robinson et al. 1983b, 1984) data argue against this unconventional result, these topologies were not considered to be accurate. The discussion of genetic distances (see below) focuses on those values obtained for the 12S rDNA sequence data when a Kimura two-parameter model with a Γ correction is employed (Table 3). For the purpose of comparison, the cytochrome *b* values are shown but not discussed.

Although we have shown that the cytochrome *b* data apparently contain little phylogenetic signal, there was still the possibility that including it might increase the topological resolution. To test for this possibility, we conducted combined data analyses with both the 12S and the cytochrome *b* data. According to the incongruence length difference test [ILD; employed as described by Cunningham (1997)], our data sets do not show significant heterogeneity and can be combined. The parsimony bootstrap tree based on the combined data showed almost no additional resolution for intergeneric relationships. These results (Fig. 5A) were based on a reconstruction that considered only T_v events at third coding positions (all events for 12S and cytochrome *b* first position and second position were used, and search parameters as above), and similar results were obtained if all third-position changes were considered. Although the combined data greatly increase the support for the node separating *Pronolagus* from the other leporids (bootstrap value = 99%), it decreases support for the *Bunolagus–Romerolagus* node (bootstrap value < 50%). As expected, the addition of cytochrome *b* data also adds to the support and resolution within the *Lepus* and the *Sylvilagus* clades; the relationships within these clades are more thoroughly examined elsewhere (Halanych and Robinson 1997; Halanych et al. in press). We believe that the combined data analyses further support our claim that the cytochrome *b* data lack phylogenetic signal at the level of divergence examined here.

Parametric Bootstrap/Simulations

The topology in Fig. 3A can be judged to be poorly resolved based on the low (nonparametric) bootstrap values obtained. In this particular case, the low bootstrap values and short internal branch lengths were due to limited phylogenetic information along intergeneric branches. The limited signal could be the result of a rapid radiation or due to an unknown mechanism which affected the gene's evolution. It should also be pointed out that the 12S data resulted in greater resolution at both shallower (Halanych and Robinson 1997) and deeper nodes in the tree. Thus, lack of molecular change and substitutional saturation are not in themselves causes of

Table 3. Genetic distances based on a Kimura two-parameter model with an empirically estimated Γ correction, α = 0.5; Cytochrome *b* data are above the diagonal and 12S rDNA data are below the diagonal

	Bid	*Bmo*	*Lam*	*Lcl*	*Lcp*	*Ocu*	*Pcr*	*Rdi*	*Saq*	*Sau*	*Sfl*	*Opr*
Bid	—	0.1955	0.1650	0.1574	0.1657	0.1909	0.2421	0.2061	0.1795	0.1650	0.1617	0.2392
Bmo	0.0932	—	0.1619	0.1633	0.1659	0.1837	0.2305	0.2055	0.1573	0.1544	0.1748	0.2434
Lam	0.0877	0.0814	—	0.0933	0.1139	0.1613	0.2471	0.1980	0.1856	0.1883	0.1726	0.2318
Lcl	0.0771	0.0626	0.0235	—	0.0932	0.1694	0.2430	0.1911	0.1907	0.1834	0.1776	0.2526
Lcp	0.0821	0.0723	0.0309	0.0272	—	0.1852	0.2483	0.2113	0.1922	0.1699	0.1772	0.2508
Ocu	0.0961	0.0865	0.0712	0.0603	0.0685	—	0.2319	0.2052	0.2048	0.1889	0.1875	0.2356
Pcr	0.1356	0.1324	0.1030	0.1043	0.1095	0.1184	—	0.2437	0.2120	0.2218	0.2302	0.2316
Rdi	0.0997	0.0784	0.0719	0.0705	0.0787	0.0899	0.1259	—	0.2270	0.2090	0.2240	0.2423
Saq	0.0860	0.0829	0.0650	0.0604	0.0700	0.0748	0.1181	0.0795	—	0.0777	0.1166	0.2118
Sau	0.0716	0.0750	0.0526	0.0540	0.0572	0.0714	0.1131	0.0718	0.0405	—	0.1211	0.2195
Sfl	0.0784	0.0869	0.0573	0.0608	0.0574	0.0670	0.1146	0.0768	0.0448	0.0432	—	0.2265
Opr	0.2057	0.1981	0.1819	0.1815	0.1884	0.2116	0.2301	0.1959	0.1788	0.1868	0.1928	—

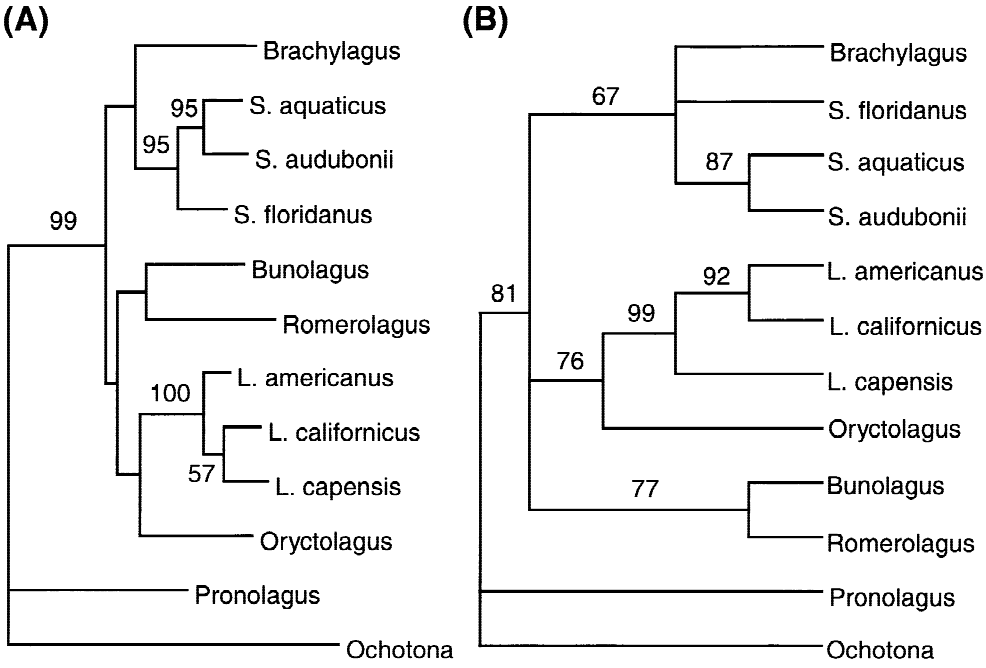

Fig. 5. **A** Results of the combined 12S and cytochrome *b* analysis. The search parameters used to recover this tree are the same as in Fig. 3A. The analysis employed the empirical T_i/T_v ratio for the 12S data but considered only T_v events at third-codon positions (and all changes at first and second sites) of the cytochrome *b* data. The bootstrap values were calculated from 200 iterations with search criteria as in Fig. 3A. Although relative branch lengths are shown, because of the weighting scheme employed, branch lengths are not directly proportionally to the number of changes. **B** Results of the parametric bootstrap/simulation study. Branch lengths for the model topology were calculated via likelihood (HKY model with empirical estimates of κ, Γ, and nucleotide frequencies). Each simulated data set was analyzed as in Fig. 3A (except that single point estimates were obtained as opposed to a full heuristic search). The parametric bootstrap values shown are of 1000 simulations.

the poor resolution. The latter point, along with congruence with other data (see below), supports a rapid radiation hypothesis. In order to explore this, we used a parametric bootstrap approach to assess the probability of recovering the modeled topology [i.e., the short internal branches (for a similar example see Hillis 1996)].

The topology presented in Fig. 3A was used as the model for simulation analyses. Branch lengths were calculated via likelihood (using PAUP*4.0d60) under an HKY model with κ (3.631435), Γ (0.584274), and the nucleotide frequencies estimated from the 12S data. One thousand simulated data sets were created. Reconstructions were performed identically to the analyses based on the observed 12S rDNA data (heuristic search with TBR and T_i/T_v weighting of 2:1) except that single-point estimates of the topology were used rather than several random additions of taxa. (This is not problematic given the number of replications employed; moreover, it facilitates the compilation the resultant trees.) Parametric bootstrap values obtained from the majority-rule consensus of the 1000 point estimates are shown in Fig. 5B.

The parametric bootstrap groups a ''*Bunolagus/Romerolagus*'' clade, an ''*Oryctolagus/Lepus*'' clade, and a ''*Brachylagus/Sylvilagus*'' clade with mod-

erate values. More interestingly, however, this analysis illustrates that if the intergeneric branches are short (an assumption of the model based on observed data), we do not expect to recover these internal branches with a high degree of confidence. Therefore, the simulations support the notion that the poor bootstrap support in the observed data are due to short internal branches, a finding that is consistent with a rapid radiation hypothesis.

Leporid Evolution

The reconstructions based on the 12S rDNA data offer two important insights to leporid evolution: (1) the majority of leporid genera arose from a single rapid diversification event, and (2) distance data (Table 3) and tree topology based on the 12S rDNA data suggest that the rock rabbits, *Pronolagus*, are distinct from other leporids.

The internal branches of the leporid phylogeny are remarkably short (see Figs. 3A and 4) compared to the terminal branches. Such short branches and low bootstrap values (both parametric and nonparametric) observed within a clade are suggestive of a very rapid diversification event (see Lara et al. 1996). Thus, it is no surprise that the exact branching pattern of the trees is dependent upon which model of evolution (e.g., T_i/T_v ratio or method of reconstruction) was employed during the phylogenetic analyses. The fossil evidence (Dawson 1958, 1981; Hibbard 1963) and the absence of shared derived cytogenetic characters (Robinson et al. 1984, and references therein) are consistent with the interpretation of a rapid radiation.

Previous studies have suggested that this radiation occurred in the middle to late Pleistocene (Dawson 1981; Hibbard 1963; Chapman and Flux 1990). However, based on mtDNA restriction site analysis, Biju-Duval et al. (1991) suggest that *Sylvilagus*, *Lepus*, and *Oryctolagus* last had a common ancestor 6–8 MYA, and the 12S rDNA data in this study suggest an even older date. By assuming that the ochotonids and leporids separated in the Oligocene or late Eocene [i.e., 30–40 MYA (Dawson 1981)], we can roughly estimate the generic diversification event. Using the average sequence divergence of 0.1978 between *Ochotona* (family Ochotonidae) and the Leporidae representatives (note the averages across the three *Lepus* and three *Sylvilagus* species were used in the calculations), we obtain a range of 0.002473–0.003297 sequence divergence per million years from the ancestor (0.1978 divergence ÷ 2 lineages ÷ 30 or 40 MY) for the 12S rDNA data. Given that the average distance between the leporid genera is 0.0803, our estimate of this radiation event is approximately 12.2 to 16.3 MYA.

This estimation of a divergence date assumes equal rates of nucleotide change across lineages (i.e., a molecular clock). In order to test this assumption, we compared the maximum-likelihood trees constructed with (DNAmlk program, PHYLIP) and without a molecular clock. A Kishino–Hagesawa (1989) test revealed that the scores of the two trees were not significantly different; the trees had the same topology (Fig. 4) but different branch lengths. Furthermore, if one calculates the divergence estimates using the greatest distances to calibrate the ochotonid–leporid split and the shortest intergeneric distances to date the radiation event (i.e., those values which give an absolute minimum estimate), the diversification event is still occurred at least 6.2 MYA. However, these divergence times are not intended to be a firm estimate but illustrate that, based on our 12S data, the radiation that produced the extant leporid genera may have occurred much earlier than the Pleistocene time scale (up to 2 MYA) previously suggested by Dawson (1958, 1981) and Hibbard (1963). Our estimate should be interpreted in a qualitative nature rather than a quantitative one.

Genetic distances between *Pronolagus*, the rock rabbits, and other genera range from 0.1047 to 0.1373, whereas the other intergeneric distances range from 0.0526 (*L. americanus–S. audubonii*) to 0.0899 (*Oryctolagus–Romerolagus*). By comparison, *Ochotona* is 0.1807 to 0.2258 different from the leporids. This finding, coupled with the observation that *Pronolagus* is consistently placed basal to the other leporid lineages in the phylogenetic reconstructions, suggests that the *Pronolagus* lineage is quite different from other leporids and is the only valid extant member of the Palaeologinae examined herein [the status of *Nesolagus* and *Pentalagus* still needs to be assessed (see Dice 1929; Corbet 1983)]. This finding supports previous work which suggests that *Pronolagus* is distinct morphologically from most leporids (Dice 1929; Corbert 1983). However, given the low bootstrap values supporting the node which excludes *Pronolagus* from the other leporids, this evolutionary uniqueness of *Pronolagus* remains equivocal.

Acknowledgments. Our research was supported by an Open Research Programme grant from the South African Foundation for Research Development to T.J.R. We thank C.A. Matthee, P. Bloomer, and anonymous reviewers for criticism of the manuscript and H. Erasmus (Northern Cape Nature Conservation, South Africa) and Mary M. Peacock (University of Arizona, Tucson) for tissue/DNA from *Lepus capensis* and *Ochotona princeps*, respectively. The work was conducted while K.M.H. was a postdoctoral fellow at the University of Pretoria, and his support from the Foundation for Research Development and the University is gratefully acknowledged.

References

Adkins RM, Honeycutt RL (1994) Evolution of the primate cytochrome c oxidase subunit II gene. J Mol Evol 38:215–231

Allard MW, Honeycutt RL (1992) Nucleotide sequence variation in the mitochondrial 12S rRNA gene and the phylogeny of African molerats (Rodentia: Bathyergidae). Mol Biol Evol 9:27–40

Biju-Duval C, Ennafaa H, Dennebouy N, Monnerot M, Mignotte F, Soriguer RC, Gaaïed AE, Hili AE, Mounolou JC (1991) Mitochon-

drial DNA evolution in lagomorphs: Origin of systematic heteroplasmy and organization of diversity in European rabbits. J Mol Evol 33:92–102

Brown WM, Prager EM, Wang A, Wilson AC (1982) Mitochondrial DNA sequences of primates: Tempo and mode of evolution. J Mol Evol 18:225–239

Chapman JA, Ceballos G (1990) The cottontails. In: Chapman JA, Flux JEC (eds) Rabbits, hares and pikas, status survey and conservation action plan. IUCN, Gland, Switzerland, pp 95–110

Chapman JA, Flux JEC (1990) Introduction and overview of the Lagomorphs. In: Chapman JA, Flux JEC (eds) Rabbits, hares and pikas, status survey and conservation action plan. IUCN, Gland, Switzerland, pp 1–6

Corbert GB (1983) A review of the classification in the family Leporidae. Acta Zool Fennica 174:11–15

Cunningham CW (1997) Can three incongruence tests predict when data should be combined? Mol Biol Evol 14:733–740

Dawson MR (1958) Later Tertiary Leporiade of North America. Univ Kans Paleo Contrib Vertebrata (Art 6):1–75

Dawson MR (1981) Evolution of the modern lagomorphs. In: Myers K, MacInnes CD (eds) Proceedings of the World Lagomorph Conference. University of Guelph Press, Guelph, Ontario, pp 1–8

Dice LR (1929) The phylogeny of the Leporidae, with the description of a new genus. J Mammal 10:340–344

Easteal S (1990) The pattern of mammalian evolution and the relative rate of molecular evolution. Genetics 124:165–173

Edwards SV, Arctander P, Wilson AC (1991) Mitochondrial resolution of a deep branch in the genealogical tree for perching birds. Proc R Soc Lond B 234:99–107

Felsenstein J (1993) PHYLIP (phylogeny inference package), Version 3.5c. Department of Genetics, University of Washington, Seattle

Fitch WM, Margoliash E (1967) A method for estimating the number of invariant amino acid coding positions in a gene using cytochrome c as a model case. Biochem Genet 4:579–593

Graur D, Duret L, Gouy M (1996) Phylogenetic position of the order Lagomorpha (rabbits, hares and allies). Nature 379:333–335

Graybeal A (1993) The phylogenetic utility of cytochrome b: Lessons from bufonid frogs. Mol Phylogenet Evol 2:256–269

Halanych KM (1996) Testing hypotheses of chaetognath origins: Long branches revealed by 18S rDNA. Syst Biol 45:223–246

Halanych KM (1998) Lagomorphs misplaced by more data and fewer taxa. Syst Biol 47:138–146

Halanych KM, Robinson TJ (1997) Phylogenetic relationships of cottontails (*Sylvilagus,* Lagomorpha): Congruence of cytogenetic and 12S rDNA data. Mol Phylogenet Evol 7:294–302

Halanych KM, Demboski JR, Cook J, Van Vuuren BJ, Klein DR, Cook JA (1999) Cytochrome b phylogeny of North American hares and jackrabbits (*Lepus,* Lagomorpha) and the effects of substitutional saturation in outgroup rooting Mol Phylogenet Evol (in press)

Hibbard C (1963) The origin of the P3 pattern of *Sylvilagus, Caprolagus, Oryctolagus,* and *Lepus.* J Mammal 44:1–15

Hillis DM (1996) Inferring complex phylogenies. Nature 383:130–131

Hillis DM, Huelsenbeck JP (1992) Signal, noise, and reliability in molecular phylogenetic analysis. J Hered 83:189–195

Hillis DM, Larson A, Davis SK, Zimmer EA (1990) Nucleic acids III: Sequencing. In: Hillis DM, Moritz C (eds) Molecular systematics. Sinauer Associates, Sunderland, MA, pp 318–370

Hoffmann RS (1993) Order Lagomorpha. In: Wilson DE, Reeder DM (eds) Mammal species of the world: A taxonomic and geographic reference. Smithsonian Institution Press, London, pp 807–827

Huelsenbeck JP, Hillis DM, Jones R (1996) Molecular approaches and the growth of phylogenetic biology. In: Ferraris, JD, Palumbi SR (eds) Molecular zoology: Advances, strategies, and protocols. Wiley–Liss, New York, pp 19–45

Irwin DM, Árnason U (1994) Cytochrome b gene of marine mammals: phylogeny and evolution. J Mammal Evol 2:37–55

Irwin DM, Kocher TD, Wilson AC (1991) Evolution of the cytochrome b gene of mammals. J Mol Evol 32:128–144

Janczewski DN, Modi WS, Clairborne Stephens J, O'Brien SJ (1995) Molecular evolution of mitochondrial 12S RNA and cytochrome b sequences in the Pantherine lineage of Felidae. Mol Biol Evol 12:690–707

Jin L, Nei M (1991) Relative efficiencies of the maximum-parsimony and distance-matrix methods of phylogeny construction for restriction data. Mol Biol Evol 8:356–365

Kishino H, Hagesawa M (1989) Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. J Mol Evol 29:170–179

Kocher TD, Thomas WK, Meyer WK, Edwards SV, Pääbo S, Villablanca FX, Wilson AC (1989) Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. Proc Natl Acad Sci USA 86:6196–6200

Lara MC, Patton JL, da Silva MNF (1996) The simultaneous diversification of South American echimyid rodents (Hystricognathi) based on complete cytochrome b sequences. Mol Phylogenet Evol 5:403–413

Li CK, Wilson RW, Dawson MR, Krishtalka L (1987) The origin of rodents and lagomorphs. In: Genoways HH (ed) Current mammalogy. Plenum Press, New York, pp 97–108

Maddison WP, Maddison DR (1992) MacClade: Analysis of phylogeny and character evolution, Version 3. Sinauer, Sunderland, MA

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Matthee CA, Robinson TJ (1997) Molecular phylogeny of the springhare, *Pedetes capensis,* based on mitochondrial DNA sequences. Mol Biol Evol 14:20–29

Meng J, Wyss AR, Dawson MR, Zhai R (1994) Primitive fossil rodent from Inner Mongolia and its implications for mammalian phylogeny. Nature 370:134–136

Meyer A (1994) Shortcomings of the cytochrome b gene as a molecular marker. TREE 9:378–380

Meyer A, Wilson AC (1990) Origin of tetrapods inferred from their mitochondrial DNA affiliation to lungfish. J Mol Evol 31:359–364

Mignotte F, Gueride M, Champagne AM, Mounolou JC (1990) Direct repeats in the non-coding region of rabbit mitochondrial DNA. Involvement in the generation of intra- and inter-individual heterogeneity. Eur J Biochem 194:561–571

Moritz C, Dowling TE, Brown WM (1987) Evolution of animal mitochondrial DNA: Relevance for population biology and systematics. Annu Rev Ecol Syst 18:269–292

Myers K, MacInnes CD (1981) Proceedings of the World Lagomorph Conference. University of Guelph Press, Guelph, Ontario.

Novacek MJ (1990) Morphology, paleontology, and the higher clades of mammals. In: Genoways HH (ed) Current mammalogy. Plenum Press, New York, pp 507–543

Novacek MJ (1993) Reflections on higher mammalian phylogenetics. J Mammal Evol 1:3–30

Novacek MJ, Wyss AR, McKenna MC (1988) The major groups of eutherian mammals. In: Benton MJ (ed) The phylogeny and classification of the tetrapods, Vol 2. Mammals, Systematics Association Special Volume No 35B. Clarendon Press, Oxford, pp 31–71

Pääbo S, Wilson AC (1988) Polymerase chain reaction reveals cloning artifacts. Nature 334:387–388

Palumbi SR (1989) Rates of molecular evolution and the fraction of nucleotide positions free to vary. J Mol Evol 29:180–187

Robinson TJ (1980) Comparative chromosome studies in the family Leporidae (Lagomorpha, Mammalia). Cytogenet Cell Genet 28:64–70

Robinson TJ, Skinner JD (1983) Karyology of the riverine rabbit, *Bunolagus monticularis,* and its taxonomic implications. J Mammal 64:678–681

Robinson TJ, Elder FFB, Lopez-Formant W (1981) Banding studies in

the volcano rabbit, *Romerolagus diazi*, and Crawshay's hare, *Lepus crawshayi*. Evidence of the leporid ancestral karyotype. Can J Genet Cytol 23:469–474

Robinson TJ, Elder FFB, Chapman JA (1983a) Karyotypic conservatism in the genus *Lepus* (order Lagomorpha). Can J Genet Cytol 25:540–544

Robinson TJ, Elder FFB, Chapman JA (1983b) Evolution of chromosomal variation in cottontails, genus *Sylvilagus* (Mammalia: Lagomorpha): *S. aquaticus*, *S. floridanus*, and *S. transitionalis*. Cytogenet Cell Genet 35:216–222

Robinson TJ, Elder FFB, Chapman JA (1984) Evolution of chromosomal variation in cottontails, genus *Sylvilagus* (Mammalia: Lagomorpha). II. *Sylvilagus audubonii*, *S. idahoensis*, *S. nuttallii*, and *S. palustris*. Cytogenet Cell Genet 38:282–289

Springer MS, Douzery E (1996) Secondary structure and patterns of evolution among mammalian mitochondrial 12S rRNA molecules. J Mol Evol 43:357–373

Springer MS, Hollar LJ, Burk A (1995) Compensatory substitutions and the evolution of the mitochondrial 12S rRNA gene in mammals. Mol Biol Evol 12:1138–1150

Sullivan J, Holsinger KE, Simon C (1995) Among-site rate variation and phylogenetic analysis of 12S rRNA in Sigmodontine rodents. Mol Biol Evol 12:988–1001

Swofford DL (1993) PAUP: Phylogenetic analysis using parsimony, Version 3.1.2d5. Illinois Natural History Survey, Champaign

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

van der Loo W, James A, Schröder J (1981) Chromosome evolution in Leporids. In: Myers K, MacInnes CD (eds) Proceedings of the World Lagomorph Conference. University of Guelph Press, Guelph, Ontario, pp 28–36

Whiteford CEM (1995) Molecular phylogeny of the genus *Pronolagus* (Mammalia: Lagomorpha) and the use of morphological and molecular characters in the delineation of *P. rupestris*, M.Sc. thesis. Department of Zoology and Entomology, University of Pretoria, Pretoria, South Africa

Yoder AD, Vilgalys R, Ruvolo M (1996) Molecular evolutionary dynamics of cytochrome b in strepsirrhine primates: The phylogenetic significance of third-position transversions. Mol Biol Evol 13: 1339–1350

Zang Y, Ryder EA (1995) Different rates of mitochondrial DNA sequence evolution in Kirk's dik-dik (*Madoqua kirkii*) populations. Mol Phylogenet Evol 4:291–297