Monophyly of the Order Rodentia Inferred from Mitochondrial DNA Sequences of the Genes for 12S rRNA, 16S rRNA, and tRNA-Valine

Melissa S. Frye and S. Blair Hedges

Department of Biology and Institute of Molecular Evolutionary Genetics, Pennsylvania State University

A recent analysis of amino acid sequence data (Graur et al.) suggested that the mammalian order Rodentia is polyphyletic, in contrast to most morphological data, which support rodent monophyly. At issue is whether the hystricognath rodents, such as the guinea pig, represent an independent evolutionary lineage within mammals, separate from the sciurognath rodents. To resolve this problem, we sequenced a region (2,645 bp) of the mitochondrial genome of the guinea pig containing the complete 12S ribosomal RNA, 16S ribosomal RNA, and transfer RNA^{VAL} genes for comparison with the available sciurognath and other mammalian sequences. Several methods of analysis and statistical tests of the data all show strong support for rodent monophyly (91%–98% bootstrap probability, or BP). Calibration with the mammalian fossil record suggests a Cretaceous date (107 mya) for the divergence of sciurognaths and hystricognaths. An older date (38 mya) for the controversial *Mus-Rattus* divergence also is supported by these data. Our neighbor-joining analyses of all available sequence data (25 genes) confirm that some individual genes support rodent polyphyly but that tandem analysis of all data does not. We propose that the conflicting results are due to several compounding factors. The unique biochemical properties of some hystricognath metabolic proteins, largely responsible for generating this controversy, may have a single explanation: a cascade effect resulting from inactivation of the zinc-binding abilities of insulin. After excluding six genes possibly affected by insulin inactivation, analyses of all available sequence data (7,117 nucleotide sites, 3,099 amino acid sites) resulted in strong support for rodent monophyly (94% BP for DNA sequences, 90% for protein sequences), which lends support to the insulin-cascade hypothesis.

Introduction

Nearly half (2,021) of all living mammal species are rodents (Wilson and Reeder 1993). Among these, the guinea pigs, chinchillas, porcupines, and their relatives are placed in a separate suborder (Hystricognathi) from the majority (89%) of rodent species (Sciurognathi). Until recently, there has been no dispute as to whether the order Rodentia forms a single natural group (i.e., is monophyletic), primarily because hystricognaths and sciurognaths share a suite of presumably derived morphological characters (Luckett and Hartenberger 1993). Some of these characters are associated with gnawing (e.g., single pair of ever-growing incisors with enamel restricted to the buccal surface), while others are not (e.g., the unique pattern of fetal membrane development).

Key words: phylogeny, systematics, molecular, evolution, mammals, *Cavia porcellus*.

Address for correspondence and reprints: Melissa S. Frye, Department of Biology, 208 Mueller Laboratory, Pennsylvania State University, University Park, Pennsylvania 16802.

Mol. Biol. Evol. 12(1):168–176. 1995.

0737-4038/95/1201-0016$02.00

Graur et al. (1991) examined published amino acid sequence data from 15 proteins and, using a maximum-parsimony analysis, obtained a phylogeny that suggested that the order Rodentia is not monophyletic. Based on these results, they proposed that the hystricognaths be placed in a separate order. Hasegawa et al. (1992) could not justify that conclusion with a maximum-likelihood analysis, although Li et al. (1992) argued that such analysis supports rodent polyphyly. Li et al. (1991), Graur et al. (1992), and Ma et al. (1993) pursued this problem with additional protein and nucleotide sequences and continued to find phylogenetic support, although not statistically significant, for rodent polyphyly.

Portions of the mitochondrial 12S rRNA gene have been sequenced in the chinchilla (Allard et al. 1991) and two other South American hystricognaths (capybara and Patagonian cavy) (Springer and Kirsch 1993) and compared with published sequences of other mammals. However, those analyses were inconclusive, possibly because of the small number of sites. Analyses of the amino acid sequences of copper-zinc superoxide dimutase showed some support for polyphyly of rodents; however, no confidence tests were done (Wolf et al. 1993). A molecular marker encoded by a retroposon, BC1 RNA, was

detected in the guinea pig and sciurognath rodents, but not the other mammalian orders examined, which led Martignetti and Brosius (1993) to conclude that it supports the monophyly of Rodentia. However, it also could be argued that a transposable element, because of its ability to move around the genome, may not be an ideal indicator of evolutionary history. Also, Thomas (1994) pointed out that their methods would not have detected the marker in cow, human, or rabbit if the retroposon had been inactivated. While a recent reanalysis of the morphological data bearing on rodent phylogeny found overwhelming support for monophyly (Luckett and Hartenberger 1993), reanalyses of available DNA and protein data have been unable to provide strong support for either rodent monophyly or polyphyly (Honeycutt and Adkins 1993; Cao et al. 1994).

In light of the inconclusive nature of the present molecular evidence bearing on rodent monophyly, we approached this problem by obtaining new mtDNA sequence data from genes evolving at appropriate rates for resolving such divergence events. The region of mitochondrial DNA containing the two ribosomal RNA genes and intervening $tRNA^{VAL}$ is sufficiently large (2.6–2.7 kb) and evolves slowly enough to provide statistical support for older divergences, such as higher-level relationships within amniotes (Hedges et al. 1993; Hedges 1994). Sequences of these genes from six orders of mammals, including two sciurognath rodents (rat and mouse), presently are available. Therefore, we chose to sequence these genes in the guinea pig (*Cavia porcellus*) for the purpose of examining the issue of rodent monophyly. In addition to the new mitochondrial sequence data and analyses, we also have chosen to reexamine the published sequence data bearing on this problem by employing the neighbor-joining method of analysis and additional statistical testing.

Material and Methods

DNA was extracted from the liver of a *Cavia porcellus* using methods described elsewhere (Hedges et al. 1991). A combination of 25 primers was used to amplify and sequence contiguous, overlapping portions of mtDNA (both strands) comprising the entire genes for 12S rRNA, 16S rRNA, and $tRNA^{VAL}$, and a portion of the gene for $tRNA^{LEU}$; these data correspond to sites 648–3229 in the complete human sequence (Anderson et al. 1981). The primer sequences are listed elsewhere (Hedges 1994), except for an additional one used, 16H12 (laboratory name, heavy strand), which is (5′–3′) TTA GGG AGA GGA TTT GAA CCT CTG and is located (3′ end) at site 3279 on the complete human sequence. Amplification and sequencing followed methods described elsewhere (Hedges et al. 1991; Hedges and Bezy 1993).

The new guinea pig sequence was compared with other mammal sequences in the databases: Virginia opossum (*Didelphis virginiana,* GenBank accession number Z29573), mouse (*Mus musculus,* J01420), rat (*Rattus norvegicus,* X14848), cow (*Bos taurus,* representing several artiodactyl sequences available, J01394), whale (*Balaenoptera physalus,* representing the two cetacean sequences available, X61145), seal (*Phoca vitulina,* representing the two carnivore sequences available, X63726), and human (*Homo sapiens,* J01415). The chicken (*Gallus gallus,* X52392) sequence was used to root the tree. Sequences were aligned with ESEE (Cabot and Beckenbach 1989). Phylogenetic and statistical analyses of the sequence data were accomplished with MEGA (Molecular Evolutionary Genetic Analysis), version 1.01 (Kumar et al. 1993).

The Jukes-Cantor (1969) distance correction was used with neighbor joining (Saitou and Nei 1987) after establishing that the transition/transversion ratio (which, for comparisions of all taxa, ranged from 0.90 to 1.88 with an average of 1.22) and base compositional frequencies (35.2%A, 23.4%C, 21.5%G, 19.8%T) for these data were not sufficiently biased to warrant the more complicated corrections available (Kumar et al. 1993). The rate of substitution often varies among sites, following a gamma distribution (Uzzell and Corbin 1971), and therefore we examined the effect of this variation by also applying a gamma distance (Kumar et al. 1993) in conjunction with the Jukes-Cantor distance. The gamma parameter (a = 1.2) was calculated from the data set. Maximum-parsimony analyses were performed using both MEGA and PAUP (Phylogenetic Analysis Using Parsimony), version 3.1 (Swofford 1993).

In addition, we examined the sequences for consistency in the rates of substitution between species of eutherians with the use of the Muse and Weir (1992) test (opossum was used as an outgroup for some of the pairwise comparisons). Times of divergence (± 1 SE) were estimated based on a UPGMA tree constructed for all taxa. The date of 130 mya for the eutherian-metatherian divergence (Carroll 1988) was used to calibrate this clock.

Sequences also were examined from additional genes, available in the databases, that pertained to the question of rodent monophyly. Nucleotide sequences were analyzed as above, and a Poisson correction was employed with the amino acid sequence data to correct for multiple replacements at the same site. Aligned sequences from each gene were analyzed separately and in tandem. Gamma distances also were used in these analyses to examine the effect of site-to-site variation in substitution or replacement rate. For the tandem analyses, gamma parameters of a = 1 for nucleotide sequence

data and $a = 2$ for amino acid sequence data were used, as recommended by Kumar et al. (1993).

The statistical significance of groups in the phylogenetic trees was assessed by two methods: the bootstrap probability (BP; Felsenstein 1985) with 2,500 replications (Hedges 1992) and, for the neighbor-joining analyses, a *t*-test for branch-length significance from zero, expressed as the complement of the probability or confidence probability (CP; Rzhetsky and Nei 1992; Kumar et al. 1993). A result of 95% or greater was considered significant.

Results

New mtDNA Sequence Data

Monophyly of the order Rodentia (*Cavia* + *Mus* + *Rattus*) was supported in all neighbor-joining and maximum-parsimony analyses of the contiguous mitochondrial sequence data set (12S rRNA, 16S rRNA, $tRNA^{VAL}$, and a portion of $tRNA^{LEU}$). The sequence alignment consisted of 2,187 bp, of which 946 were variable and 526 were parsimony sites. In the neighbor-joining analyses, statistical support for rodent monophyly was 98% (BP) and 99% (CP) when only four taxa were compared: hystricognath (*Cavia porcellus*), sciurognath (either *Rattus norvegicus* or *Mus musculus*), primate (*Homo sapiens*), and outgroup (*Gallus gallus*). When all eight mammalian taxa were included, support for rodent monophyly was 91% (BP) and 94% (CP). In the parsimony analyses, support for rodent monophyly was 96% (BP) in the four-taxon case and 90% (BP) when all taxa were compared (if both bird and opossum are used as outgroups, the support drops to 88%). However, when sequence data from each of these three genes were examined separately (table 1), none of the genes strongly supports rodent monophyly. When the only other mtDNA sequences available for these taxa (cytochrome *b*) (Ma et al. 1993) are included in a tandem analysis (nine taxa), support decreases to 67% (BP).

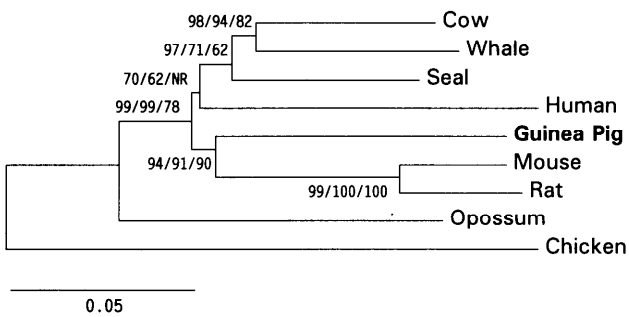

FIG. 1.—Phylogenetic tree of mammals produced by neighbor-joining analysis (Jukes-Cantor distance) of a contiguous section of mtDNA containing the complete genes for 12S rRNA, 16S rRNA, and $tRNA^{VAL}$ (2,187 aligned sites, 946 variable sites, 526 parsimony sites). Numbers at nodes are the CP of the standard error of branch length/bootstrap probability using Jukes-Cantor with neighbor joining (2,500 replicates)/bootstrap probability using a maximum-parsimony heuristic search (2,500 replicates). NR indicates a node not resolved by parsimony.

Recent work has shown an acceleration in the rate of substitutions in the human lineage in cytochrome *b* and cytochrome oxidase subunit II genes (Ramharack and Deeley 1987; Ma et al. 1993). For the region of the mitochondrial genome examined here, a consistent rate of evolution was rejected for 11 of the 15 pairwise comparisons that included human data, using the Muse and Weir test (1992), which indicates a faster rate of evolution in the human lineage than in the other eutherian lineages studied (table 2). However, nine of these comparisons involved the Carnivora-Artiodactyla-Cetacea lineage, which appears to have a slower rate of substitution than in the rodent or human lineages. The time of divergence of the hystricognath and sciurognath lineages is estimated from these data to be 107 ± 7.2 mya, which is older than the divergence of many eutherian orders. Using only transversions, this split was estimated to be 106 ± 10 mya. Another controversy in mammalian systematics involves the time of divergence between the mouse and rat lineages. Paleontological evidence suggests a date of 10–15 mya (Jaeger et al. 1986), while molecular studies indicate that an earlier date, 22–35 mya, may be correct (Sarich 1985; Janke et al. 1994). Our analyses lend support to the earlier date, giving a divergence time of 38 ± 2.4 mya, when both transitions and transversions are considered (33 ± 3.2 mya when only transversions are used).

All Available Sequence Data

Four-taxon analyses of the available sequence data yielded mixed results (table 1). In the analyses of DNA sequence data (13 genes), six genes (alpha-lactalbumin, islet amyloid protein, lipoprotein lipase, pancreatic polypeptide, transglutaminase, 16S rRNA) supported a primates + sciurognaths grouping, four (factor IX, ILGF I, preproglucagon, preproinsulin) supported primates + hystricognaths, and three (cytochrome *b*, 12S rRNA, $tRNA^{VAL}$) supported hystricognaths + sciurognaths. However, only two genes, lipoprotein lipase and pancreatic polypeptide, showed significant bootstrap support (for primates + sciurognaths). A tandem analysis of all nucleotide data (10,515 sites) supported both rodent monophyly (hystricognaths + sciurognaths) and polyphyly (primates + sciurognaths) at 46% BP. The use of a gamma distance ($a = 1.0$) in the tandem analysis supported rodent monophyly with 71% BP. Thus, the results from analyses of all available DNA sequence data were inconclusive.

In the analyses of available amino acid sequence data (22 genes, including those translated from nucleo-

Table 1
Results from Neighbor-Joining Analyses (Four-Taxon) of Previously Published Gene Sequences[a]

Gene	Outgroup[b]	Aligned Sites		Topology Supported[c]		
				Polyphyly		Monophyly
		Total	Variable	P-S	P-H	H-S
DNA sequences:						
Alpha-lactalbumin	Mru	450	238	84%	6%	9%
Cytochrome *b*	Gga	1,197	601	4%	1%	93%
Factor IX	X	888	249	14%	55%	29%
Insulin-like growth factor I	Gga	357	77	18%	78%	2%
Islet amyloid protein	Gga	210	94	69%	11%	18%
Lipoprotein lipase	Gga	1,401	476	96%	2%	1%
Pancreatic polypeptide	Gga	435	237	99%	0%	0%
Preproglucagon	Gga	666	313	1%	76%	23%
Preproinsulin	Gga	393	197	8%	73%	18%
Transglutaminase	Gga	2,112	903	47%	35%	17%
12S rRNA	Dvi,Gga	839	329	2%	14%	78%
16S rRNA	Dvi,Gga	1,279	510	53%	5%	38%
$tRNA^{VAL}$	Dvi,Gga	69	39	14%	1%	84%
12S and 16S rRNA and $tRNA^{VAL}$	Dvi,Gga	2,187	946	0%	19%	88%
Tandem	...	10,515	4,669	46%	4%	46%
Tandem without metabolic sequences	...	7,116	2,994	5%	0%	94%
Protein sequences:[d]						
Alpha-crystallin A chain	Dvi	173	26	1%	18%	80%
Alpha-globin	Tac	141	60	6%	0%	92%
Alpha-lactalbumin	Mru	164	89	30%	27%	42%
Beta-globin	Mgi	148	58	51%	5%	43%
Beta-nerve growth factor	Gga	243	109	7%	4%	88%
"Big" gastrin	Dvi	46	14	31%	0%	68%
Copper-zinc superoxide dimutase	Cca	167	54	16%	5%	77%
Cytochrome *b*	Gga	399	166	5%	6%	88%
Factor IX	X	505	279	89%	7%	3%
Insulin	Gga	107	73	6%	84%	9%
Insulin-like growth factor	Gga	154	31	37%	57%	4%
Islet amyloid protein	Gga	136	51	10%	15%	73%
Lipocortin	Cli	347	125	79%	3%	16%
Lipoprotein lipase	Gga	467	139	81%	13%	4%
Myelin basic protein	Gga	202	71	66%	0%	32%
Myoglobin	Mru	153	42	17%	0%	82%
Pancreatic polypeptide	Gga	101	67	51%	36%	11%
Pancreatic ribonuclease	Mru	128	71	22%	53%	24%
Preproglucagon	Gga	180	93	54%	0%	45%
Transglutaminase	Gga	703	289	51%	22%	25%
Vasoactive intestinal peptide	Gga	46	10	65%	32%	1%
Vasopressin neurophysin precursor	Aan	171	68	58%	41%	0%
Tandem	...	4,199	1,639	19%	2%	78%
Tandem without metabolic sequences	...	3,099	1,144	5%	3%	90%

[a] Three ingroup taxa were used: *Cavia porcellus, Homo sapiens,* and *Rattus norvegicus* (or in the case of transglutaminase, beta-nerve growth factor, and myelin basic protein, *Mus musculus,* and preproglucagon, *Mesocricetus auratus*).

[b] Gga, *Gallus gallus;* Mru, *Macropus rufus* (kangaroo); Dvi, *Didelphis virginiana;* Aan, *Anser anser* (goose); Tac, *Tachyglossus aculeatus* (echidna); Cli, *Columba livia* (pigeon); Cca, *Caretta caretta;* and X, bovine and human factor X.

[c] P, Primate; S, Sciurognathi; and H, Hystricognathi; expressed as bootstrap probability based on 2,500 replicates.

[d] Includes amino acid translations of DNA sequences.

Table 2
Significant χ^2 Values for Overall Pairwise Comparisons among Taxa

Species Compared	Outgroup	Overall χ^2 Values[a]
Bos taurus, Balaenoptera physalus	*Rattus norvegicus*	11.3
Bos taurus, Homo sapiens	*Mus musculus*	11.4
Bos taurus, Homo sapiens	*Rattus norvegicus*	13.7
Bos taurus, Homo sapiens	*Didelphis virginiana*	19.2
Bos taurus, Rattus norvegicus	*Didelphis virginiana*	7.33
Phoca vitulina, Balaenoptera physalus	*Mus musculus*	8.66
Phoca vitulina, Balaenoptera physalus	*Rattus norvegicus*	9.97
Phoca vitulina, Homo sapiens	*Mus musculus*	23.8
Phoca vitulina, Homo sapiens	*Rattus norvegicus*	14.3
Phoca vitulina, Homo sapiens	*Cavia porcellus*	10.4
Phoca vitulina, Homo sapiens	*Didelphis virginiana*	19.7
Phoca vitulina, Rattus norvegicus	*Didelphis virginiana*	9.30
Phoca vitulina, Cavia porcellus	*Didelphis virginiana*	8.27
Balaenoptera physalus, Homo sapiens	*Cavia porcellus*	6.64
Balaenoptera physalus, Homo sapiens	*Didelphis virginiana*	22.4
Balaenoptera physalus, Rattus norvegicus	*Didelphis virginiana*	7.10
Mus musculus, Rattus norvegicus	*Didelphis virginiana*	9.85
Mus musculus, Homo sapiens	*Didelphis virginiana*	16.4
Rattus norvegicus, Homo sapiens	*Didelphis virginiana*	12.7

[a] Two degrees of freedom.

tide sequences), 10 genes supported primates + sciurognaths (beta-globin, factor IX, lipocortin, lipoprotein lipase, myelin, pancreatic polypeptide, preproglucagon, transglutaminase, vasoactive intestinal peptide, and vasopressin neurophysin precursor), three supported primates + hystricognaths (insulin, insulin-like growth factor, and pancreatic ribonuclease), and nine supported hystricognaths + sciurognaths (alpha-crystallin A chain, alpha-globin, alpha-lactalbumin, beta-nerve growth factor, "big" gastrin, copper-zinc superoxide dimutase, cytochrome *b*, islet amyloid protein, and myoglobin). However, no genes showed significant bootstrap support. A tandem analysis of all amino acid data (4,199 residues) supported rodent monophyly (hystricognath + sciurognath) at 78% BP. The use of a gamma distance ($a = 2.0$) in the tandem analysis resulted in the same topology, with BP = 85%.

Results using larger numbers of taxa were even less conclusive and usually resulted in decreased statistical confidence (the Appendix contains species used for each gene). For DNA sequence data, three genes (islet amyloid protein, preproinsulin, and 12S rRNA) supported rodent monophyly, and nine supported polyphyly (although not necessarily a sciurognath-primate clade). Only one gene (preproinsulin) supported monophyly significantly at BP = 96%. For amino acid sequences, eight genes (alpha-globin, beta-globin, beta-nerve growth factor, "big" gastrin, insulin, copper-zinc superoxide dimutase, islet amyloid protein, preproglucagon, and transglutaminase) supported monophyly (none significantly). The addition of taxa altered some trees, but only in the case of one gene, preproinsulin (DNA sequences), did support increase to a significant level.

Discussion

Monophyly of the mammalian order Rodentia now has additional support from a large molecular data set (12S rRNA, 16S rRNA, $tRNA^{VAL}$). In four-taxon analyses, this support is statistically significant (99% CP, 98% BP), and in the full nine-taxon analysis, it is nearly significant (94% CP, 91% BP). Nonetheless, our analyses of the available molecular data bearing on this problem (25 genes) confirmed the results of Graur et al. (1991) that some genes significantly support a polyphyletic Rodentia. It is therefore of interest to look more closely at all of the molecular evidence and attempt to find an explanation for this discordance.

In all of the separate gene analyses (table 1), only two showed statistically significant support for any grouping (primate + sciurognath = rodent polyphyly). In those cases (lipoprotein lipase and pancreatic polypeptide), the results were obtained using DNA sequence data, while analyses of the corresponding amino acid sequences for each gene did not show significant support (BP = 51% and 81%). None of the amino acid sequences or tandem analyses showed significant support for either rodent monophyly or polyphyly. These findings, together with the maximum-likelihood and more recent maximum-parsimony analyses, suggest that those previously published data are unable to resolve this controversy.

It is important to examine why these data fail to

resolve this issue. Our results (fig. 1) suggest that the guinea pig, and presumably the Hystricognathi, diverged soon after the origin of the order Rodentia, which will increase the difficulty of determining the true tree. Graur et al. (1992) suggested that guinea pig genes appear to evolve more rapidly only because the phylogenetic assumptions made about rodents were incorrect. However, selection constraints apparently were relaxed in the hystricognath lineage, at least in the case of insulin, which would result in rapid rates of evolution in that lineage. Kimura (1983) pointed out that the high substitution rate of insulin probably is due to removal of the zinc-binding constraints, although some other researchers suggest that it is the result of positive evolutionary pressures (Nishi and Steiner 1990). Insulin in most hystricognaths is very different from the insulins of other mammals because of the lack of zinc in insulin-producing cells. This results in an inability to form stable hexamers and in greatly decreased activity; guinea pig insulin shows as little as 2% of the biological activity of bovine insulin (Blundell and Wood 1975). Still, when all available insulin sequences were examined, the guinea pig clustered (although not significantly) with other rodents (Hedges et al. 1990).

The effects of having an insulin that is, at best, only a third as effective as that of other mammals could result in compensatory changes occurring in other enzymes and hormones involved in the glucose, fatty acid, and ketone body cycles. Insulin is involved in the regulation of all three of these cycles. For example, guinea pigs produce other insulin-related molecules in some tissues that resemble normal mammalian insulin (Rosenzweig et al. 1985). However, there is only one copy of insulin in the guinea pig genome, so this is not simply the result of a duplication of the insulin gene (Chan et al. 1984). Glucagon in guinea pigs has been shown to have an unusual C-terminal region that probably causes reduced function (Conlon et al. 1985), and, since glucagon acts as an antagonist to insulin, a glucagon with reduced potency would be advantageous. Insulin is a powerful antilipolytic hormone. Lipoprotein lipase activity increases 10- to 20-fold in the adipose tissue of fasted guinea pigs, while it increases only 4- to 5-fold in fasted rats and mice, and the enzyme is turned over very rapidly in guinea pig adipose tissue (Semb and Olivecrona 1986).

The change in insulin leading to reduced activity could cause the selectional constraints of many other enzymes and hormones to be altered and thus could result in a cascade of changes as other proteins adapted to the new insulin. If this hypothesis is correct, then enzymes or hormones that could be affected by the change in insulin activity should not be used to determine phylogenetic relationships. A neighbor-joining analysis of the tandem set of amino acid sequences, excluding any that might be influenced by changes in insulin activity (alpha-lactalbumin, insulin, islet amyloid protein, glucagon, lipoprotein lipase, and pancreatic polypeptide), resulted in higher bootstrap support for monophyly: 90% BP with Poisson correction (versus 78%) and 93% BP with gamma correction (versus 85%). The tandem set of nucleotide sequences, excluding the same genes, also resulted in strong support for monophyly: 94% BP with Jukes-Cantor distance (versus 46%), 93% with Jukes-Cantor/gamma distance (versus 71%).

Another possible explanation for the unusual biochemical traits of the guinea pig is its 5,000-yr history of domestication (Woods 1982). Domestication alters selection pressures, which could be a factor if some of these unusual biochemical traits are not present in other hystricognaths (for many genes, only the guinea pig has been sequenced). If domestication was an influence, then it is important to choose, if possible, genes that may not have been affected by such altered selection pressures. Domestication is unlikely to cause large rate differences in the ribosomal RNA genes because they perform a more general function in the cell. This feature, and the large number of sites (2.7 kb), may explain why these new mtDNA sequence data are able to resolve rodent monophyly.

The new DNA sequence data set presented here provide strong support for the monophyly of rodents, a conclusion that already has considerable support from morphology (Honeycutt and Adkins 1993; Luckett and Hartenberger 1993). When all molecular sequence data are considered, including tandem analyses of the nucleotide data and the amino acid data, the original claim of rodent polyphyly (Graur et al. 1991) can no longer be supported. While additional molecular data bearing on this question will be welcomed, there is presently no compelling reason to assume that the Hystricognathi is not the sister group to the Sciurognathi or that the guinea pig is not a rodent.

Sequence Availability

The new mtDNA sequence for guinea pig 12S rRNA, 16S rRNA, and $tRNA^{VAL}$ and portions of the $tRNA^{PHE}$ and $tRNA^{LEU}$ has been deposited in the sequence databases under accession number L35585.

Acknowledgments

We thank Savante Pääbo for an advance copy of the sequence data for *Didelphis virginiana,* Sudhir Kumar for statistical advice, Spencer Muse for his technical advice with the Muse and Weir test, and Carla Hass for comments on the manuscript. Work was supported by funds provided by The Pennsylvania State University.

APPENDIX

Table A1

Species Used in Trees with Greater Numbers of Taxa

Sequence	Species
DNA Sequences:	
Alpha-lactalbumin	Mru, Bta, Chi, Cpo, Hsa, Mmu, Oar, Rno, Sus
Cytochrome *b*	Gga, Aam, Bta, Cdo, Cpo, Dbi, Egr, Gca, Haf, Hsa, Laf, Mdo, Mmu, Rno, Slo, Ssc
Factor IX	X, Cfa, Cpo, Hsa, Mmu, Oar, Rno, Sus
Insulin-like growth factor I	Gga, Cpo, Hsa, Oar, Rno, Sus
Islet amyloid protein	Gga, Cpo, Hsa, Mau, Mmu, Ode, Rno
Lipoprotein lipase	Gga, Cpo, Hsa, Oar, Rno
Pancreatic polypeptide	Gga, Cdo, Cpo, Hsa, Mmu, Rno
Preproglucagon	Gga, Bta, Cpo, Hsa, Mau
Preproinsulin	Gga, Bta, Cae, Cfa, Cpo, Hsa, Mmu, Pta
Transglutaminase	Gga, Cpo, Hsa, Mmu
12S rRNA	Gga, Bph, Bta, Cpo, Dvi, Hsa, Mmu, Pvi, Rno
16S rRNA	Gga, Bph, Bta, Cpo, Dvi, Hsa, Mmu, Pvi, Rno
$tRNA^{VAL}$	Gga, Bph, Bta, Cpo, Dvi, Hsa, Mmu, Pvi, Rno
12S and 16S rRNA and $tRNA^{VAL}$	Gga, Bph, Bta, Cpo, Dvi, Hsa, Mmu, Pvi, Rno
Protein sequences:	
Alpha-crystallin A chain	Dvi, Aja, Bac, Bta, Bva, Cdo, Cfa&Fca, Cho, Cpo&Pca, Csi, Eca, G, Gca&Ham, Hsa, Laf, Lfu, Mja, Mmu*, Mru, Dvi, Oaf, Ocu, Opr, Pca*, Pph, Ppo, Rno&Mru&Mun&Mau, Ssc, Tin, Tme, Uur, Zca
Alpha-globin	Tac, Bta, Cpo, Dvi, Ema, Hsa, Mly, Rno
Alpha-lactalbumin	Mru, Bta, Chi, Cpo, Hsa, Mmu, Oar, Rno, Ssc
Beta-hemoglobin	Tac, Bta, Hsa, Mgi, Mly, Rno
Beta-nerve growth factor	Gga, Cpo, Hsa, Mmu, Pna
"Big" gastrin	Dvi, Bta&Oar, Cbr, Cfa, Chi, Cpo, Fca, Hsa, Rno, Ssc
Copper-zinc superoxide dimutase	Cca, Cpo, Eca, Hsa, Mmu, Oar, Ocu, Rno, Ssc
Cytochrome *b*	Gga, Mdo, Cpo, Haf, Aam, Gca, Bta, Ssc, Hsa, Mmu, Slo, Dbi, Egr, Rno, Laf, Cdo
Factor IX	X, Hsa, Mmu, Rno, Ssc, Cpo, Bta, Cfa, Oar
Insulin	Gga, Bta, Cae, Cfa, Cpo, Hsa, Mmu, Pta
Insulin-like growth factor	Gga, Chi, Ssc, Oar, Rno, Bta, Cpo, Hsa
Islet amyloid protein	Gga, Rno, Hsa, Fca, Mmu, Cpo
Lipocortin	Cli, Mmu, Rno, Hsa, Cpo
Lipoprotein lipase	Gga, Hsa, Ssc, Rno, Cpo
Myelin basic protein	Gga, Cpo, Hsa, Mmu, Ptr
Myoglobin	Gga, Atr, Bac, Bph, Bta, Cac, Cap, Cel, Cfa&OmeCfi, Cgu, Cpo, Dvi, Eca&Ebr, Eeu, Egi, Ema, Gcr, Ggo, Gme, Hag, Hgr&Pvi, Hsa, Hsa, Ige, Ksi, Laf, Lla, Llu, Lma, Lmu, Lpi, Mca, Mfa, Mme, Mmu, Mno, Mru, Nco, Oaf, Oan, Oar, Ocu, Oor, Opr, Ozi, Pan&Epa&Pen, Pca**, Pgu, Pph&Pda, Ppo, Ppy, Psi, Ptr, Rae, Rno, Sle, Ssc, Tac, Tgl, Ttr&Dde, Vch, Zca, Zca*
Pancreatic polypeptide	Gga, Cfa, Cpo, Hsa, Mmu, Rno
Pancreatic ribonuclease	Mru, Aac, Aam, Ame, Bar, Bta, Cbr, Cca*, Cdo&Cac, Cel, Cho, Clo, Cpo, Cta, Dda, Eca, Gca, Gmu, Ham, Hcr, Hsa, Mau, Mco, Mmu, Oar, Pen, Pgu, Rno, Rta, Sle, Ssc
Preproglucagon	Gga, Bta, Cpo, Hsa, Mau, Ode, Rno
Transglutaminase	Gga, Cpo, Hsa, Mmu
Vasoactive intestinal peptide	Gga, Cpo, Mmu, Oar&Chi&Cfa, Ocu, Rno, Ssc
Vasopressin neurophysin precursor	Aan, Bph, Bta, Chi, Eca, Hsa, Oar, Rno, Sus

NOTE.—Species abbreviations used were as follows (asterisks are used to indicate different species that had the same abbreviation): Aal, *Alces alces alces;* Aam, *Antilocapra americana;* Amc, *Aepyceros melampus;* Atr, *Aotus trivirgatus;* Aja, *Artibeus jamaicensis;* Bac, *Balaenoptera acutorostrata;* Bar, *Bubalus arnee bubalis;* Bta, *Bos taurus;* Bph, *Balaenoptera physalus;* Bva, *Bradypus variegatus;* Cae, *Cercopithecus aethiops;* Cap, *Cebus apella;* Cba, *Camelus bactrianus;* Cbe, *Chinchilla brevicaudata;* Cca, *Caretta caretta;* Cca*, *Capreolus capreolus;* Cdo, *Camelus dromedarius;* Cel, *Cervus elaphus;* Cfa, *Canis familiaris;* Cfi, *Castor fiber;* Cgu, *Ctenodactylus gundi;* Chi, *Capra hircus;* Cho, *Choloepus hoffmanni;* Cja, *Callithrix jacchus;* Clo, *Cricetulus longicaudatus;* Cpo, *Cavia porcellus;* Csi, *Ceratotherium simum;* Cta, *Connochaetes taurinus;* Dbi, *Diceros bicornis;* Dda, *Dama dama;* Dde, *Delphinus delphis;* Dvi, *Didelphis virginiana;* Ebu, *Equus burchelli;* Eca, *Equus caballus;* Eeu, *Erinaceus europaeus;* Egi, *Eschrichtius gibbosus;* Egr, *Equus grevyi;* Ema, *Elephas maximus;* Epa, *Erythrocebus patas:* Fca, *Felis catus:* G, *Galago* sp.; Gca, *Giraffa camelopardalis;* Gcr, *Galago crassicaudatus;* Ggo, *Gorilla gorilla;* Gme, *Globicephala melaena;* Gmu, *Galea musteloides;* Haf, *Hystrix africaeaustralis;* Hag, *Hylobates agilis;* Ham, *Hippopotamus amphibius;* Hcr, *Hystrix cristata;* Hgr, *Halichoerus grypus;* Hhy, *Hydrochoerus hydrochaeris;* Hsa, *Homo sapiens;* Ige, *Inia geoffrensis;* Ksi, *Kogia simus;* Laf, *Loxodonta africana;* Lfu, *Lemur fulvus fulvus;* Lla, *Lagothrix lagothricha;* Llu, *Lutra lutra;* Lme, *Lagostromus maximus;* Lmu, *Lepilemur mustelinus;* Lpi, *Lycaon pictus,* Mau, *Mesocricetus auratus;* Mca, *Mesoplodon carlhubbsi;* Mco, *Myocastor coypus;* Mdo, *Monodelphis domestica;* Mfa, *Macaca fascicularis;* Mgi, *Macropus gigantus;* Mja, *Manis javanica;* Mly, *Megoderma lyra;* Mme, *Meles meles;* Mmu, *Mus musculus;* Mmu*, *Macaca mulatta;* Mno, *Megaptera novaeangliae;* Mru, *Macropus rufus;* Mun, *Meriones unguiculatus;* Mvi, *Mustela vison;* Nco, *Nycticebus coucang;* Oaf, *Orycteropus afer;* Oan, *Ornithorhynchus anatinus;* Ocu, *Oryctolagus cuniculus;* Ode, *Octodon degus;* Ome, *Otocyon megalotis;* Oor, *Orcinus orca;* Opr, *Ochotona princeps;* Ozi, *Ondatra zibethicus;* Pan, *Papio anubis;* Pca, *Pedetes capensis;* Pca*, *Procavia capensis;* Pca**, *Physeter catodon;* Pda, *Phocoenoides dalli dalli;* Pen, *Presbytis entellus;* Pgu, *Proechimys guairae;* Pna, *Praomys natalensis;* Pph, *Phocoenoides phocoena;* Ppo, *Perodicticus potto edwarsi;* Ppy, *Pongo pygmaeus;* Psi, *Phoca sibirica;* Ptr, *Pan troglodytes;* Pvi, *Phoca vitulina;* Rae, *Rousettus aegyptiacus;* Rno, *Rattus norvegicus;* Rta, *Rangifer tarandus;* Sle, *Spalax leucodon;* Slo, *Stenella longirustris;* Ssc, *Sus scrofa;* Ssc*, *Saimiri sciureus;* Tac, *Tachyglossus aculeatus;* Tgl, *Tupaia glis;* Tin, *Tapirus indicus;* Tme, *Tamandua mexicana;* Ttr, *Tursiops truncatus;* Uur, *Ursus ursinus;* Vch, *Vulpes chama;* Zca, *Zalophus californianus;* Zca*, *Ziphius cavirostris.*

LITERATURE CITED

ALLARD, M. W., M. M. MIYAMOTO, and R. L. HONEYCUTT. 1991. Tests for rodent polyphyly. Nature **353**:610–611.

ANDERSON, S., A. T. BANKIER, B. G. BARREL, M. H. L. DE BRUIJN, A. R. COULSON, J. DROUIN, I. C. EPERSON, D. P. NIERLICH, B. A. ROE, F. SANGER, P. H. SCHREIER, A. J. H. SMITH, R. STANDEN, and I. G. YOUNG. 1981. Sequence and organization of the human mitochondrial genome. Nature **290**:457–465.

BLUNDELL, T. L., and S. P. WOOD. 1975. Is the evolution of insulin Darwinian or due to selectively neutral mutation? Nature **257**:197–203.

CABOT, E. L., and A. T. BECKENBACH. 1985. Simultaneous editing of multiple nucleic acid and protein sequences with ESEE. Comput. Appl. Biosci. **5**:233–234.

CAO, Y., J. ADACHI, T. YANO, and M. HASEGAWA. 1994. Phylogenetic place of guinea pigs: no support of the rodent-polyphyly hypothesis from maximum-likelihood analyses of multiple protein sequences. Mol. Biol. Evol. **11**:593–604.

CARROLL, R. L. 1988. Vertebrate paleontology and evolution. W. H. Freeman, New York.

CHAN, S. J., V. EPISKOPOV, S. ZEITLIN, S. K. KARATHANASIS, A. MACKRELL, D. F. STEINER, and A. EFSTRATIADIS. 1984. Guinea pig preproinsulin gene: an evolutionary compromise? Proc. Natl. Acad. Sci. USA **81**:5046–5050.

CONLON, J. M., H. F. HANSEN, and T. W. SCHWARTZ. 1985. Primary structure of glucagon and a partial sequence of oxyntomodulin from the guinea pig. Regul. Pept. **11**:309–320.

FELSENSTEIN, J. 1985. Confidence limits in phylogenies: an approach using the bootstrap. Evolution **39**:783–791.

GRAUR, D., W. A. HIDE, and W.-H. LI. 1991. Is the guinea-pig a rodent? Nature **351**:649–652.

GRAUR, D., W. A. HIDE, A. ZARKIKH, and W.-H. LI. 1992. The biochemical phylogeny of guinea pigs and gundis and the paraphyly of the order Rodentia. Comp. Biochem. Physiol. **101B**:495–498.

HASEGAWA, M., Y. CAO, J. ADACHI, and T. YANO. 1992. Rodent polyphyly? Nature **355**:595.

HEDGES, S. B. 1992. The number of replications needed for accurate estimation of the bootstrap p-value in phylogenetic studies. Mol. Biol. Evol. **9**:366–369.

———. 1994. Molecular evidence for the origin of birds. Proc. Natl. Acad. Sci. USA **91**:2621–2624.

HEDGES, S. B., and R. L. BEZY. 1993. Phylogeny of xantusiid lizards: concern for data and analysis. Mol. Phylogenet. Evol. **2**:76–87.

HEDGES, S. B., R. L. BEZY, and L. R. MAXSON. 1991. Phylogenetic relationships and biogeography of the xantusiid lizards, inferred from mitochondrial DNA sequences. Mol. Biol. Evol. **8**:767–780.

HEDGES, S. B., C. A. HASS, and L. R. MAXSON. 1993. Relations of fish and tetrapods. Nature **363**:501–502.

HEDGES, S. B., K. D. MOBERG, and L. R. MAXSON. 1990. Tetrapod phylogeny inferred from 18S and 28S ribosomal RNA sequences and a review of the evidence for amniote relationships. Mol. Biol. Evol. 7:607–633.

HONEYCUTT, R. L., and R. M. ADKINS. 1993. Higher level systematics of eutherian mammals: an assessment of molecular characters and phylogenetic hypotheses. Annu. Rev. Ecol. Systematics **24**:279–305.

JAEGER, J.-J., H. TONG, and C. DENYS. 1986. The age of *Mus-Rattus* divergence: paleontological data compared with the molecular clock. C. R. Acad. Sci. Paris, Série II **14**:917–922.

JANKE, A., G. FELDMMAIER-FUCHS, W. K. THOMAS, A. VON HAESELER, and S. PÄÄBO. 1994. The marsupial mitochondrial genome and the evolution of placental mammals. Genetics **137**:243–256.

JUKES, T. H., and C. R. CANTOR. 1969. Evolution of protein molecules. Pp. 21–132 *in* H. L. MUNROE, ed. Mammalian protein metabolism. Academic Press, New York.

KIMURA, M. 1983. The neutral theory of molecular evolution. Cambridge University Press, New York.

KUMAR, S., K. TAMURA, and M. NEI. 1993. MEGA: molecular evolutionary genetics analysis. Version 1.01. The Pennsylvania State University, University Park.

LI, W.-H., W. A. HIDE, and D. GRAUR. 1992. The origin of rodents and guinea pigs. Nature **359**:277.

LI, W.-H., W. A. HIDE, A. ZHARKIKH, D. P. MA, and D. GRAUR. 1991. The molecular taxonomy and evolution of the guinea pig. J. Hered. **83**:174–181.

LUCKETT, W. P., and J.-L. HARTENBERGER. 1993. Monophyly or polyphyly of the order Rodentia: possible conflict between morphological and molecular interpretations. J. Mammalian Evol. **2**:127–147.

MA, D. P., A. ZHARKIKH, D. GRAUR, J. L. VANDEBERG, and W.-H. LI. 1993. Structure and evolution of opossum, guinea pig and porcupine cytochrome *b* genes. J. Mol. Evol. **36**: 327–334.

MARTIGNETTI, J. A., and J. BROSIUS. 1993. Neural BC1 RNA as an evolutionary marker: guinea pig remains a rodent. Proc. Natl. Acad. Sci. USA **90**:9698–9702.

MUSE, S. V., and B. S. WEIR. 1992. Testing for equality of evolutionary rates. Genetics **132**:269–276.

NISHI, M., and D. F. STEINER. 1990. Cloning of complementary DNA encoding islet amyloid polypeptide, insulin and glucagon precursors from a New World rodent, the degu, *Octodon degus*. Mol. Endocrinol. **4**:1192–1198.

RAMHARACK, R., and R. G. DEELEY. 1987. Structure and evolution of the primate cytochrome c oxidase subunit II gene. J. Biol. Chem. **262**:14014–14021.

ROSENZWEIG, J. L., D. LEROITH, M. A. LESNIAK, C. C. YIP, D. N. ORTH, H. R. NANKIN, P. MURONE, M. BERELOWITZ, and L. A. FROHMAN. 1985. Two distinct insulin-related molecules in the guinea pig: immunological and biochemical characterization of insulin-like immunoactivity from extrapancreatic tissues of guinea pig. Diabetologia **28**:237–243.

RZHETSKY, A., and M. NEI. 1992. A simple method for estimating and testing minimum-evolution trees. Mol. Biol. Evol. **9**:945–967.

SAITOU, N., and M. NEI. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**:406–425.

SARICH, V. M. 1985. Rodent macromolecular systematics. Pp. 423–452 *in* W. P. LUCKETT and J.-L. HARTENBERGER, eds. Evolutionary relationships among the rodents. Plenum, New York.

SEMB, H., and T. OLIVECRONA. 1986. Nutritional regulation of lipoprotein lipase in guinea pig tissues. Biochim. Biophys. Acta **876**:249–255.

SPRINGER, M. S., and J. A. W. KIRSCH. 1993. A molecular perspective on the phylogeny of placental mammals based on mitochondrial 12S rDNA. J. Mammalian Evol. **1**:149–166.

SWOFFORD, D. L. 1993. PAUP: phylogenetic analysis using parsimony, version 3.1. Illinois Natural History Survey, Champaign, Ill.

THOMAS, R. 1994. What is a guinea pig? Trends Ecol. Evol. **9**:159–160.

UZZELL, T., and K. W. CORBIN. 1971. Fitting discrete probability distributions to evolutionary events. Science **172**: 1089–1096.

WILSON, D. E., and D. M. REEDER. 1993. Mammal species of the world: a taxonomic and geographic reference. Smithsonian Institution Press, Washington, D.C.

WOLF, B., K. REINECKE, K. D. AUMANN, R. BRIGELIUS-FLOHE, and L. FLOHE. 1993. Taxonomic classification of the guinea pig based on its copper/zinc superoxide dimutase sequence. Biol. Chem. Hoppe Seyler **374**:641–649.

WOODS, C. A. 1982. The history and classification of the South American hystricognath rodents: reflections on the far away and long ago. Pp. 377–392 *in* M. A. MARES and H. H. GENOWAYS, eds. Mammalian biology in South America, University of Pittsburgh Pymatuning Lab of Ecology, Linesville, Pa.

SHOZO YOKOYAMA, reviewing editor

Received May 9, 1994

Accepted September 23, 1994