Letter to the Editor

Where Do Rodents Fit? Evidence from the Complete Mitochondrial Genome of Sciurus vulgaris

Aurelio Reyes, Carmela Gissi, Graziano Pesole, François M. Catzeflis, and Cecilia Saccone

*Centro Studio Mitcondri e Metabolismo Energetico, Consiglio Nazionale delle Ricerche, Bari, Italy; †Dipartimento di Biochimica e Biologia Molecolare, Università di Bari, Bari, Italy; ‡Dipartimento di Fisiologia e Biochimica Generali, Università di Milano, Milano, Italy; §Laboratoire de Paléontologie, Paléobiologie et Phylogénie, Institut des Sciences de l’Évolution, Centre National de la Recherche Scientifique, Université Montpellier II, Montpellier, France

Nowadays, the order Rodentia represents almost half of all living mammalian species, classified into 30–33 families (Hartenberger 1998), and shows high levels of variability in morphology, habitat utilization, behavior, life history strategy, and geographic distribution (Eisenberg 1981; Wilson and Reeder 1993).

Morphological classifications have considered the order Rodentia a monophyletic group on account of dental, cranial, postcranial, and soft anatomical attributes (Luckett and Hartenberger 1993; Hartenberger 1998). Nevertheless, this view was challenged in different molecular surveys at the beginning of this decade (Graur, Hide, and Li 1991; Graur et al. 1992; Li et al. 1992). More recent studies have relied on the analysis of complete mitochondrial genomes of four rodent species, namely, rat, mouse, guinea pig, and dormouse, encompassing three major lineages: Muridae, Caviidae, and Gliridae (D’Erchia et al. 1996; Reyes, Pesole, and Saccone 1998). Irrespective of the methodological approach, the results obtained in these studies showed the existence of two well-supported rodent clades, one including murid rodents (rat and mouse) and the other including nonmurid rodents (guinea pig and dormouse). These surveys would support rodent paraphyly; however, some authors claim rodent monophyly based on different mitochondrial data sets or methodological approaches (Cao, Okada, and Hasegawa 1997; Philippe 1997; Sullivan and Swofford 1997; Philippe and Laurent 1998). On the whole, molecular studies are still scarce and do not resolve the relationships among the major lineages of this order (see Huchon, Catzeflis, and Doutrey [1999] and references therein), let alone between rodent families. Thus, the relationships among rodents and the issue of rodent monophyly versus paraphyly/polyphyly are far from being settled, and only additional data could help solve them. In this context, we have sequenced, according to standard procedures described elsewhere (e.g., Reyes, Pesole, and Saccone 1998), the complete mitochondrial (mt) genome of the European red squirrel, Sciurus vulgaris (specimen number V-784 of F. Catzeflis), a representative of the Sciuridae family which represents one of the major rodent lineages.

The mtDNA sequence of the squirrel S. vulgaris has been deposited in the EMBL database under accession number AJ238588. The control region is 1,059 nt long and shows some peculiar features, such as the lack of repeated motifs, the lack of two out of the three conserved sequence blocks (CSBs) (namely, CSB2 and CSB3), and the existence of a 58-bp region next to the tRNA-Pro which shows high similarity (average value 83.5%) to the corresponding region of rabbit, dormouse, and guinea pig but not with that of rat and mouse.

Phylogenetic analyses were carried out on the complete mammalian mtDNA sequences available in the EMBL database (release 61): human (Homo sapiens, V00662), common chimpanzee (Pan troglodytes, D38116), pigmy chimpanzee (Pan paniscus, D38113), gorilla (Gorilla gorilla, D38114), orangutan (Pongo pygmaeus, D38115), gibbon (Hylobates lar, X99256), baboon (Papio hamadryas, Y18001), horse (Equus caballus, X79547), donkey (Equus asinus, X97373), Indian rhinoceros (Rhinoceros unicornis, X97336), white rhinoceros (Ceratherium simum, Y07726), harbor seal (Phoca vitulina, X63726), gray seal (Halichoerus grypus, X72004), cat (Felis catus, U20753), dog (Canis familiaris, U96639), fin whale (Balaenoptera physalus, X61145), blue whale (Balaenoptera musculus, X72204), cow (Bos taurus, V00654), sheep (Ovis aries, AF010406), pig (Sus scrofa, AJ002189), hippopotamus (Hippopotamus amphibius, AJ101957), Neotropical fruit bat (Arteius jamaciensis, AF061340), African elephant (Loxodonta africana, AJ224821), aardvark (Orycteropus afer, Y18475), armadillo (Dasypus novemcintus, Y11832), rabbit (Oryctolagus cuniculus, AJ001588), guinea pig (Cavia porcellus, AJ222767), fat dormouse (Glis glis, AJ001562), rat (Rattus norvegicus, X14848), mouse (Mus musculus, V00711), hedgehog (Erinaceus europaeus, X88898), opossum (Didelphis virginiana, Z29573), wallaroo (Macropus robustus, Y10524), and platypus (Ornithorhyncus anatinus, X83427). Noneu- therian supergenes were always used as mammalian outgroups.

Concatenated supergenes for ribosomal 12S and 16S RNA genes and for protein genes coded by the H-strand were used for phylogenetic analyses. For protein genes, only first and second codon positions (P12) were considered, due to the substitution saturation and base composition heterogeneity observed in third codon positions (data not shown). Phylogenetic analyses were conducted using different approaches: the Markov model (Saccone et al. 1990; also called general time reversible [GTR] in PAUP*) for the calculation of genetic distances, and the minimum-evolution (ME; Rzhetsky and
Fig. 1.—Phylogenetic tree obtained by the maximum-likelihood (PROTML) method using the mtREV24-F model on concatenated H-stranded amino acid sequences. The lengths of the branches are proportional to the number of amino acid replacements per site. Bootstraps are based on 100 replications. Noneutherians were used as outgroups.

Nei 1992) or neighbor-joining (NJ; Saitou and Nei 1987) method for tree reconstruction. The GTR analyses were also performed on the two supergenes assuming a gamma distribution for substitution rate across sites, where the parameter $\alpha$ (Yang 1994) and the proportion of invariant sites ($I$) were estimated with the maximum-likelihood method assuming the GTR-ME phylogeny using PAUP*. Maximum-parsimony (MP) and maximum likelihood with the mtREV24-F model were used and a heuristic search (PROTML) was conducted on concatenated H-strand encoded amino acid sequences using the PAUP* (Swofford 1998) and MOLPHY (Adachi and Hasegawa 1996) packages, respectively. Bootstrap values were based on 1,000 replicates, except in the case of PROTML, where only 100 replicates were performed. Furthermore, log-likelihood ratio tests (Kishino and Hasegawa 1989) were performed on both protein-coding and ribosomal genes in order to determine the degree of support for alternative hypotheses of relationships among rodents and other mammals.

To avoid a compositional bias in the analysis, the hedgehog sequence was excluded from the phylogenetic reconstruction, because it deviates significantly from the mean nucleotide frequency on the P12 sites ($\chi^2 = 74.18$, df = 3, $P < 0.001$). Figures 1 and 2 show the tree obtained by means of PROTML and GTR-ME on amino acid and protein-coding genes, respectively. In both cases, relationships within ferungulates (carnivores, perissodactyls, cetaceans, artiodactyls) and their clustering with Chiroptera are supported by high bootstrap values, in agreement with previous studies (e.g., Pumo et al. 1998; Reyes, Pesole, and Saccone 1998; Ursing and Arnason 1998). Elephants and aardvarks, both members of the African clade (Springer et al. 1997), are closely related to one another (bootstrap values 91% and 93%), Armadillo clusters with the African clade, when re-
solved, with a high bootstrap value (92%). With regard to rodent species, two different clades supported by high bootstrap values were observed: one containing rat and mouse (100%) and the other containing guinea pig, squirrel, and dormouse (86% and 90%). Within the latter group, squirrel is most closely related to dormouse (bootstrap values 96% and 83%). Rabbit was found to cluster with nonmurid rodents, leading to a “quasi-Glires” clade, even though bootstrap support is poor (52% and 66%). The main difference between figures 1 and 2 is that the former suggests rodent polyphyly and the latter suggests rodent paraphyly if the position of rabbit is not considered. The phylogenetic trees obtained with the other methods and data sets (i.e., GTR-ME and GTR-NJ either assuming constant rates of evolution or with $\alpha = 0.77$, $I = 0.42$ for CDS and $\alpha = 0.64$, $I = 0.24$ for rRNA supergenes, and MP on the protein supergene) show the same general topology (data not shown). Such trees exclude rodent monophyly, but they do not discriminate between paraphyly and polyphyly. Furthermore, log-likelihood ratio tests for different rodent tree topologies revealed that monophyly is significantly rejected in all cases and that the best tree corresponds to the one showing rodent polyphyly and the clustering of nonmurids with rabbit (fig. 1). However, the alternative tree, assuming rodent paraphyly and rabbit clustering with nonmurids (fig. 2), is not significantly different (best tree in fig. 1 shows $\ln L = 83,193.8$, while the corresponding value for the alternative tree shown in fig. 2 is $\ln L = 83,203.2$, based on P12 sites).
The results show that, irrespective of the methodological approach used for phylogenetic reconstruction, squirrel cluster first with dormouse (e.g., figs. 1 and 2) with high bootstrap support. The clustering of Sciuridae (squirrels) with Gliridae (dormice) has been previously reported on the basis of mitochondrial 12S rRNA analysis (Nebdal, Honeycutt, and Schlitter 1996), the vWF gene (Huchon, Catzeflis, and Douzery 1999), and the presence of the retroposon B1-did (Kramerov, Vassetzky, and Serdobova 1999). This association has also been proposed on the basis of paleontological and morphological evidence (Bugge 1985; Lavocat and Parent 1985; Hartenberger 1998). Thus, the close relationship between these two lineages is supported by both molecular and morphological data, even though other conflicting opinions based on comparative morphology have also been put forward (McKenna and Bell 1997). The existence of two completely separated clades of rodents, one containing nonmurid rodents (squirrel, dormouse, and guinea pig) and the other containing murid rodents (rat and mouse) has been supported by high bootstrap values (>80% in most cases) with different methodological approaches (e.g., figs. 1 and 2). The relative positions of the two rodent clades in the mammalian tree and the result of the log-likelihood ratio tests lead to two hypotheses: (1) that rodents are polyphyletic, with murids being an early branch of the tree and nonmurids being a sister group of nonprimate eutherians, and (2) that rodents are paraphyletic and placed at a basal position of the mammalian tree, with murid rodents representing an earlier branch than nonmurid rodents. The first hypothesis would be in great disagreement with morphological and paleontological data (Luckett and Hartenberger 1993; Hartenberger 1996) and would suggest that the degree of convergent and/or parallel evolution between murid and nonmurid rodents might be higher than once believed (Hartenberger 1985; Li et al. 1992). The second hypothesis obtained support from previous analyses based on complete mitochondrial genomes (D’Erchia et al. 1996; Pumol et al. 1998; Reyes, Pesole, and Saccone 1998), and in the present study, it also remains one of the most likely hypotheses.

It is noteworthy that using the largest available mitochondrial data set (34 mammalian species, of which 5 are rodents) and different methodological approaches (homogeneous- and heterogeneous-site rate models), we never found support for rodent monophyly, in contrast to what is reported in previous studies (Cao, Okada, and Hasegawa 1997; Philippe 1997; Sullivan and Swoford 1997; Philippe and Laurent 1998). Finally, it has also been suggested that rodent monophyly is obscured by the higher evolutionary rate of murids and subsequent long-branch attraction with noneneither species at the basal position of the tree (Philippe and Laurent 1998). Nevertheless, we confirmed by relative-rate test using the CODRATES program (Muse and Gaut 1994) that, except for squirrels, all rodent species considered here show the same or similar evolutionary rates on ribosomal and P12 sites. Thus, long-branch attraction should affect both murid and nonmurid species in the same way (with the exception of squirrel) and would not explain why murids remain at a basal position while nonmurids are placed as a sister group of the nonprimate eutherians.

All in all, the position of rodents in the mammalian tree remains an open question, as data are in agreement with both paraphyly and polyphyly of rodents. The confirmation of one of these hypotheses or even a completely different picture may arise from the analysis of new rodent sequences. In particular, the sampling of murid-related taxa, such as Spalacidae, Rhizomyidae, and Dipodidae, or of representatives of other rodent lineages (e.g., Ctenodactylidae, Bathyergidae, and Anomaluridae) would be of great interest.

Acknowledgments

We thank M. Lonigro for revision of the manuscript. This work was supported by TMR European Project ERB-FMRX-CT98-0221 and by MURST, Italy. A.R. and C.G. contributed equally to this paper.

LITERATURE CITED


DAN GRAUR, reviewing editor

Accepted February 16, 2000