Gene 531 (2013) 388–397

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Jumping and gliding rodents: Mitogenomic affinities of Pedetidae and Anomaluridae deduced from an RNA-Seq approach

Pierre-Henri Fabre [a,b], Knud A. Jønsson [b], Emmanuel J.P. Douzery [a,*]

[a] Institut des Sciences de l'Evolution (ISEM, UMR 5554 UM2-CNRS-IRD), Université Montpellier II, Place Eugène Bataillon - CC 064 - 34095 Montpellier Cedex 5, France
[b] Center for Macroecology Evolution and Climate at the Natural History Museum of Denmark, University of Copenhagen, Universitetsparken, 15, DK-2100 Copenhagen Ø, Denmark

ARTICLE INFO

Article history:
Accepted 16 July 2013
Available online 23 August 2013

Keywords:
Graphiurus
Pedetes
Phylogeny
Mitochondrial DNA
Next generation sequencing
Rodentia

ABSTRACT

An RNA-Seq strategy was used to obtain the complete set of protein-coding mitochondrial genes from two rodent taxa. Thanks to the next generation sequencing (NGS) 454 approach, we determined the complete mitochondrial DNA genome from *Graphiurus kelleni* (Mammalia: Rodentia: Gliridae) and partial mitogenome from *Pedetes capensis* (Pedetidae), and compared them with published rodent and outgroup mitogenomes. We finished the mitogenome sequencing by a series of amplicons using conserved PCR primers to fill the gaps corresponding to tRNA, rRNA and control regions. Phylogenetic analyses of the mitogenomes suggest a well-supported rodent phylogeny in agreement with nuclear gene trees. *Pedetes* groups with *Anomalurus* into the clade Anomaluromorpha, while *Graphiurus* branches within the squirrel-related clade. Moreover, *Pedetes* + *Anomalurus* branch with *Castor* into the mouse-related clade. Our study demonstrates the utility of NGS for obtaining new mitochondrial genomes as well as the importance of choosing adequate models of sequence evolution to infer the phylogeny of rodents.

1. Introduction

Among placental mammals, rodents constitute a highly diverse group in terms of taxonomy (Carleton and Musser, 2005; Kay and Hoekstra, 2008), morphology (Hautier et al., 2008), molecular evolution (Nabholz et al., 2008a, 2009), macroevolution (Fabre et al., 2012) and life history traits (Eisenberg, 1981). However, high rates of speciation (Steppan et al., 2004) and high levels of homoplasy in morphological and molecular characters (Hartenberger, 1985; Hautier et al., 2008, 2011; Nabholz et al., 2008a; Philippe, 1997; Wood, 1965) have hampered delimitation of inter-familial relationships within rodents (Fig. 1). Consequently, the systematics of rodents has displayed an array of differing phylogenetic hypotheses (Luckett and Hartenberger, 1993; Montgelard et al., 2008; Blanga-Kanfi et al., 2009). Poor taxon sampling (Lecointre et al., 1993) and variable evolutionary rates among lineages (Philippe et al., 2005) are additional difficulties that have faced molecular systematists interested in the inference of the rodent Tree of Life. In particular, the "Guinea pig is not a rodent" controversy (D'Erchia et al., 1996; Graur et al., 1991; Philippe, 1997) has helped to identify the biases which can affect molecular phylogenies (Delsuc et al., 2005; Felsenstein, 2004).

During the last decade, there has been a growing interest in the inference of rodent phylogeny using molecular markers like nuclear genes (Adkins et al., 2001, 2003; Blanga-Kanfi et al., 2009; DeBry, 2003; Huchon et al., 1999, 2002, 2007; Meredith et al., 2011; Montgelard et al., 2008), mitogenomes (Arnason et al., 2008; Horn et al., 2011; Horner et al., 2007; Reyes et al., 1998, 2000) and retroposons (Churakov et al., 2010; Kramerov et al., 1999; Veniaminova et al., 2007). These studies concurred with a new hypothesis of suprafamilial relationships among rodents, and proposed three main monophyletic assemblages: (i) a Guinea pig-related clade: Ctenohystrica (Hystricognathi + Ctenodactylidae/Diatomyidae), (ii) a squirrel-related clade: Sciuroidea (=Sciuridae + Aplodontidae) + Gliridae, and (iii) a mouse-related clade: Myodonta (=Muridae + Dipodidae), Geomyoidea (=Geomyidae + Heteromyidae) +Castoridae, and Anomaluridae + Pedetidae. While Ctenohystrica and Sciuroidea + Gliridae are well-supported by independent markers (Adkins et al., 2001, 2003; Huchon and Douzery, 2001; Huchon et al., 1999; Montgelard et al., 2002), this is not the case for the mouse-related clade. In fact only recently published phylogenies using nuclear markers in combination (Blanga-Kanfi et al., 2009; Montgelard et al., 2008) have provided support for this hypothesis. Moreover, only one mitogenome analysis has recovered the monophyly of the mouse-related group (Horn et al., 2011) whereas another mitogenomic analysis supports *Anomalurus* as a closer relative to the Hystricognathi (Horner et al., 2007). In addition, within the mouse-related clade, relationships among

Abbreviations: AIC, Akaike Information Criterion; AU test, Approximately Unbiased test; BP, Bootstrap Percentage; GTR, General Time Reversible; ML, Maximum Likelihood; NGS, Next Generation Sequencing; PP, Posterior Probability; RNA-Seq, RNA Sequencing.

* Corresponding author: Tel.: +33 467144863.
E-mail address: emmanuel.douzery@univ-montp2.fr (E.J.P. Douzery).

0378-1119/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.gene.2013.07.059

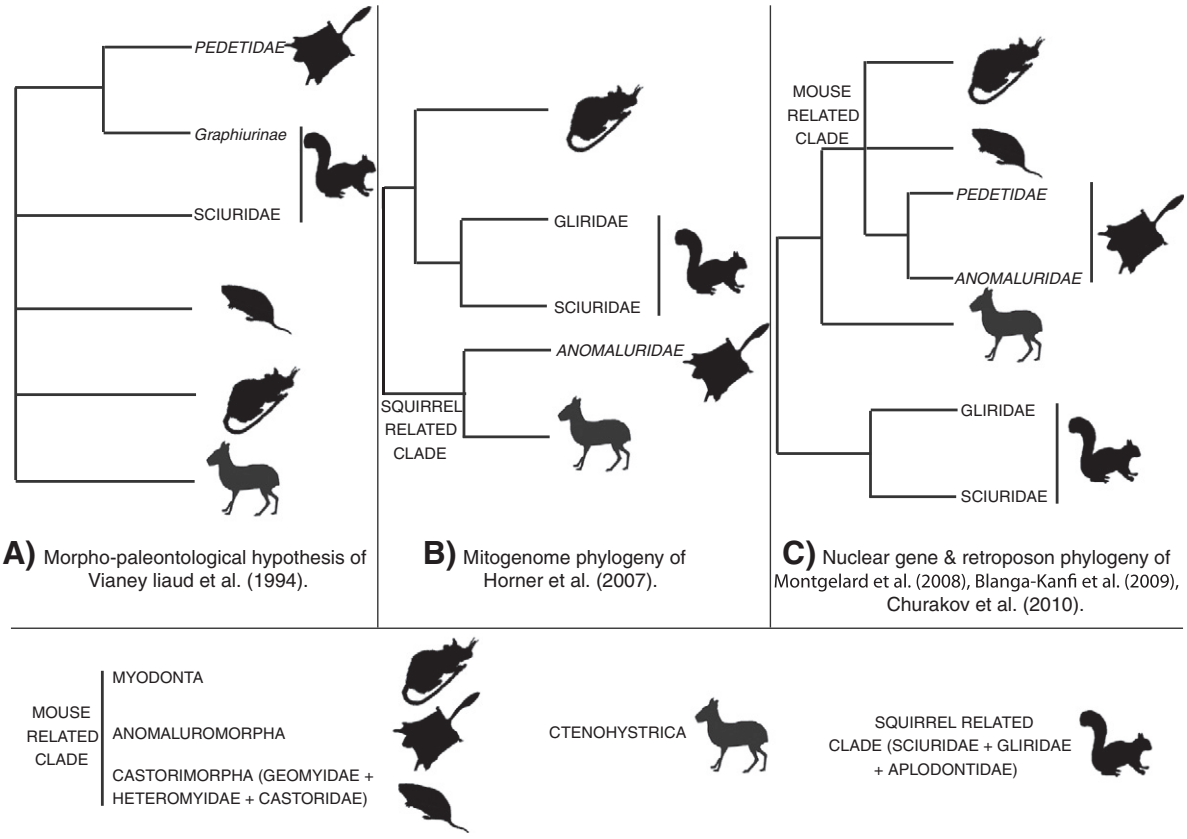

Fig. 1. Different phylogenetic hypotheses for rodents based on morpho-paleontological (A), mitogenome (B) and nuclear gene plus retroposon datasets. Systematic names and icons for the major rodent clades are given along the phylogenies.

Castorimorpha (=Castoridae, and Geomyidae + Heteromyidae), Anomaluromorpha (=Anomaluridae + Pedetidae) and Myodonta are still unresolved (Blanga-Kanfi et al., 2009; Churakov et al., 2010; Montgelard et al., 2008).

Congruence among mitochondrial and nuclear gene trees suggests that molecular phylogenies describe the underlying species-level relationships (Reyes et al., 2004). It is therefore important to build mitochondrial and nuclear phylogenies for the same set of taxa. Here, we investigate the relationships of the three main rodent clades using complete mitogenomes focusing on the deep branching patterns within the mouse-related clade. Because of its atypical pattern of molecular evolution (Galtier et al., 2009; Nabholz et al., 2008a,b), the mitochondrial genome often displays a limited resolving power to answer questions about higher level relationships within animals (Arnason et al., 2008; Foster et al., 1997; Rota-Stabelli et al., 2009). However, the rise of next generation sequencing (NGS) methods (*e.g.* 454, Illumina; Meyer et al., 2008) and new probabilistic models of sequence evolution (Lartillot and Philippe, 2004; Lartillot et al., 2007; Rota-Stabelli et al., 2009) has resurrected the interest in mitogenomes for inferring phylogenetic relationships (*e.g.* Douglas et al., 2011; Jex et al., 2008; Nabholz et al., 2010; Rowe et al., 2011). The new, high-throughput sequencing methodologies have increased the amount of available complete mitogenomes of mammals (*e.g.*, Horn et al., 2011; Mason et al., 2011; Miller et al., 2009, 2012). Consequently, several phylogenetic analyses of mitogenomes have allowed for a detailed study of higher level rodent systematics (Horner et al., 2007), the *Rattus* species level systematics (Robins et al., 2008, 2010) and molecular evolution within Castoridae (Horn et al., 2011). Using a combination of an RNA-Seq approach (Wang et al., 2009) with NGS and a series of PCR amplicons, we here sequenced two new mitochondrial genomes for the following taxa: *Pedetes capensis* (springhare; Pedetidae) and *Graphiurus kelleni* (Kellen's African dormouse; Gliridae). This deep sequencing transcriptomic approach allowed us to detect functional mitochondrial protein-coding genes despite their potential differential levels of expression (Nabholz et al., 2013).

Springhares are represented by two extant African species (*Pedetes capensis* and *P. surdaster*), with adaptations to arid habitat and jumping locomotion. Hypotheses about their evolutionary affinities are controversial (Fig. 1; see Luckett and Hartenberger, 1993; Vianey-Liaud et al., 1994) and led morphologists to relate them to either Anomaluridae (Bugge, 1974; Gill, 1872; Lavocat and Parent, 1985) or Graphiurinae (Vianey-Liaud et al., 1994). African dormice of the subfamily Graphiurinae belong to Gliridae together with Glirinae and Leithiinae, and represent the most diversified of the three main lineages of this family (Montgelard et al., 2003; Nunome et al., 2007). Based on the comparison of two individual mitochondrial genes, it has been suggested that springhares are related to African scaly-tailed flying squirrels (Anomaluridae) (Montgelard et al., 2002). The latter point suggests that the inclusion of a Pedetidae taxon would break the long, isolated branch of Anomaluridae observed in mitogenomic trees (Horner et al., 2007).

Here, we provide new molecular data for these two undersampled rodent families Pedetidae and Gliridae, and by sampling these lineages we test the Anomaluromorpha hypothesis using complete mitogenome data. Using model-based approaches (maximum likelihood, Bayesian inference) and topological tests of alternative hypotheses we also explored the potential relationship between Pedetidae, Anomaluridae, and their putative close relatives Castoridae and Myodonta in the mouse-related clade. Our superfamily level rodent mitogenomic phylogeny allows us to address the following questions: (i) Are the rodent

mitogenome phylogenetic results congruent with nuclear gene and retroposon analyses? (ii) What are the closest relatives of Pedetidae among rodents? (iii) How does the choice of model of sequence evolution affect phylogenetic inference?

2. Materials and methods

2.1. Mitogenome sequencing

To test for the phylogenetic affinities of Pedetidae with either Anomaluridae or Graphiurinae, we focused on two species for which biological material was available for RNA-Seq experiments. Nucleic acid extracts were isolated from cell cultures (fibroblasts) for *Graphiurus kelleni* (Gauthier Dobigny, Centre de Biologie et de Gestion des Populations, Montferrier-sur-Lez, France) and *Pedetes capensis* (Terry Robinson, Stellenbosch, South Africa). Total RNA was extracted following the RNeasy protocol (Qiagen). After first and second strand cDNA synthesis, tagged libraries were prepared for the RNA-Seq 454 sequencing by GATC-Biotech (www.gatc-biotech.com). The 454 reads—with an average length of 250 nt—were produced under the GS-FLX protocol by the same company.

De novo assembly of the 454-reads was conducted with Geneious (Drummond et al., 2011), version 5.5. The resulting contigs were mined for mitochondrial sequences with a similarity-based BLASTN approach (Altschul et al., 1990) against the phylogenetically related mitogenomes of *Glis glis* (accession NC_001892; Reyes et al., 1998) and *Anomalurus* sp. (accession NC_009056; Horner et al., 2007) for *Graphiurus kelleni* and *Pedetes capensis*, respectively, and using a 10^{-5} significance threshold for the expected (e) value.

Mitogenomic fragments not obtained from this 454 approach were amplified and sequenced using primers anchored in the 454 contigs (SI Table 1). Total DNA extracted by the phenol–chloroform procedure was used as PCR template. PCR conditions started with an initial heating to 95 °C for 5 min, followed by 5 cycles of 95 °C for 30 s, 61 °C for 30 s and 72 °C for 1 min. This was followed by three repeats of the same 5 cycles but with annealing temperature decreasing from 59 °C to 57 °C, and then 55 °C. These touchdown cycles were followed by 20 cycles of 95 °C for 30 s, 53 °C for 30 s and 72 °C for 1 min, with a final extension of 72 °C for 5 min. PCR products were purified from 1% agarose gel using Amicon Ultrafree-DNA columns (Millipore). They were sequenced on both strands using Sanger automatic sequencing (BigDye® Terminator v3.1 kit) on an Applied ABI Prism 3130 XL sequencer.

All sequences were checked for the presence of signatures that may be indicative of pseudogenes. We found no stop codons or indels within the coding gene sequences. We also built alignments for each individual gene and constructed the corresponding phylogenetic trees (see below). We did not find any phylogenetic incongruence or spurious branch length patterns among individual genes.

2.2. Dataset assembly

The molecular dataset presented here include all Rodentia taxa for which mitochondrial genome sequences were available in public databases (see SI Table 2). During the course of our study some additional mitogenomes became available (*e.g.*, Ryu et al., 2013) but were not included in the analyses. Here, we focused on 37 mitochondrial genomes for 34 rodent species (available as of March 2011). Outgroups were chosen among the Euarchontoglires (Janecka et al., 2007; Meredith et al., 2011) for which mitogenomes were available (see SI Table 2).

DNA sequences were aligned with MUSCLE 3.7 (Edgar, 2004). From these mitogenomes, we built 4 different datasets by combining: (i) the 13 protein-coding genes (partition "123"), (ii) the 13 protein-coding genes, the two rRNA and all the 22 tRNA (partition "123RNA"), (iii) the 13 protein-coding genes but excluding the third codon positions (partition "12"), and (*iv*) the 13 protein-coding genes excluding the third codon positions, the two rRNA and all the 22 tRNA (partition "12RNA"). We also concatenated the 13 protein-coding genes translated into amino acid sequences in a super-protein dataset (partition "AA"). Two different *a priori* partitioning strategies were applied to the nucleotide data sets: (a) each gene/rRNA/tRNA was assigned its own partition ("gene-partitioned"), and (b) each codon position within each protein-coding gene and each rRNA and tRNA was assigned its own partition ("codon-RNA-partitioned").

2.3. Maximum likelihood phylogenetic analyses

Phylogenetic tree inference was performed on the concatenated datasets (partitions 123, 123RNA, 12, 12RNA, and AA) under the maximum likelihood (ML) criterion, with RAxML version 7.0.4 (Stamatakis, 2006). We used Modeltest 3.07 (Posada and Crandall, 1998) to determine the best fitting model of DNA sequence evolution following the Akaike Information Criterion (AIC). We applied to each partition the GTRGAMMA option, *i.e.*, the general time reversible (GTR) model (Rodriguez et al., 1990) plus among-site rate heterogeneity accommodated by a gamma (Γ) distribution (Yang, 1996), together with a fraction of invariant sites. For the bootstrap analyses, we used the GTRMIX option of RAxML which assumes the RAxML GTRCAT model for the topological search, but then uses the GTRGAMMA model when computing the likelihood. Each RAxML run comprised 10,000 tree search replicates (with the default parameters). We applied the same methodology to each gene dataset. The robustness of nodes was estimated with ML bootstrap percentages (BP_{RAxML}) after 10,000 pseudoreplicates.

2.4. Bayesian analyses

Bayesian inferences were conducted on three of the concatenated datasets (partitions 123, 123RNA, and AA). To account for the potential heterogeneity of substitution patterns among partitions, we used the CAT mixture model (Lartillot and Philippe, 2004), implemented in PhyloBayes 3 (Lartillot et al., 2009). To account for among-site heterogeneity in the nucleotide and amino acid substitution rates, we used a Γ distribution with 4 discrete categories (Γ_4). Relative exchangeabilities among nucleotides and amino acids were described under the GTR model. For the protein dataset, we also explored a range of empirical exchangeability matrices to evaluate their potential effect on the phylogenetic inference. In addition to GTR exchangeabilities, we therefore used the following matrices: POI (Poisson, *i.e.*, all exchangeabilities among amino acids are equal), WAG (Whelan and Goldman, 2001), LG (Le and Gascuel, 2008), mtREV (Adachi and Hasegawa, 1996), and mtZoa (Rota-Stabelli et al., 2009). For each mitochondrial dataset, two Markov chains Monte Carlo analyses were run with PhyloBayes for 10,000 cycles (ca. 8,000,000 generations) with trees sampled every 5 cycles after discarding the first 1000 as a burn-in. Convergence was ensured when the maximum difference in bipartition frequencies as estimated by the two chains was below 0.1. Node support was estimated by posterior probabilities (PP).

2.5. Testing alternative topologies

The different topologies found in the various analyses, as well as a number of other topologies proposed in other studies, were tested using the approximately unbiased (AU—Shimodaira, 2002) test with 10 batches of 10^6 bootstrap replicates as implemented in CONSEL (Shimodaira and Hasegawa, 2001). The 4 nucleotide datasets were used for these tests. The site likelihoods for each of the test topologies were calculated with PAUP* version 4.0b10 (Swofford, 2002), under the corresponding gene partitioning scheme and appropriate model of sequence evolution for each dataset.

3. Results

3.1. Next-generation sequencing of Pedetes and Graphiurus mitogenomes using 454

After the 454 run, we obtained NGS transcriptomic data for a total of 62,424 reads for *Graphiurus* and 51,668 reads for *Pedetes*. After *de novo* assembly, we collected 5119 and 4241 contigs for the dormouse and the springhare, respectively. Twelve *Graphiurus* contigs showed significant similarity with the mitogenome of *Glis*, while 9 *Pedetes* contigs were similar to *Anomalurus*. To recapitulate, 4111 reads with an average length of 239 nt were mapped onto the *Graphiurus* mitogenome, providing 14,679 nt with a mean coverage of 69 ×. For *Pedetes*, 690 reads with an average length of 236 nt were mapped onto the mitogenome, providing 10,343 nt with a mean coverage of 11 ×.

To quantify variations of the transcription level along the mitogenome, we reported the per-nucleotide coverage of the RNA-Seq 454 reads for *Graphiurus kelleni*, as its mitochondrial DNA is the best covered of the two sequenced in this study (Fig. 2). If we assume that the protocol of 454 library construction has not biased the relative initial amounts of mitochondrial transcripts, then the RNA-Seq coverage can be used as a proxy for the transcription level of each mitochondrial region. All protein-coding genes display a transcription level above the median for mitochondrial transcripts, except ND3 and ND6. For ribosomal regions, the expression level of the 12S rDNA is close to the median, whereas the 16S rDNA displays a very strong over-expression (*i.e.*, over the 90% to 95% quantiles). For the non-coding region, the second half of the control region (*i.e.*, on the $tRNA^{Phe}$ side) is expressed at a level close to the median.

Finally, we did not detect expression of any of the 22 tRNA. A standard PCR + Sanger protocol has therefore been used to finish the mitogenome sequence. With this approach, we were able to fill all the mitochondrial gaps for *Graphiurus*, and all but one for *Pedetes* (*i.e.*, the area corresponding to the control region). After PCR completion, the mitogenome of *Graphiurus kelleni* is 16,596 nt long. The mitogenome of *Pedetes capensis* is 15,366 nt long, and displays a similar genome arrangement, except that, despite sequencing efforts, we did not succeed in obtaining the control region and the two surrounding $tRNA^{Pro}$ and $tRNA^{Phe}$ gene sequences. Both mitogenomes contained 13 protein-coding genes (ND1 to 6, ND4L, CytB, COI to III, ATP6, and ATP8), two rRNA genes (12S rRNA and 16S rRNA), 22 tRNA genes, and a major non-coding region (Fig. 2). The gene arrangement pattern was identical to those of typical placentals (Boore, 1999). Most genes were encoded on the H-strand, except the ND6 and eight tRNA genes ($tRNA^{Gln}$, $tRNA^{Ala}$, $tRNA^{Asn}$, $tRNA^{Cys}$, $tRNA^{Tyr}$, $tRNA^{Ser(UGA)}$, $tRNA^{Glu}$, and $tRNA^{Pro}$). These mitogenomic sequences were deposited in the EMBL-EBI European Nucleotide Archive (ENA) database under accession numbers HE978360 (*Graphiurus kelleni*) and HE983623 (*Pedetes capensis*).

3.2. Maximum likelihood phylogenetic inference with a priori site partitioning

The phylogenetic position of *Pedetes* and *Graphiurus* was inferred under the ML method from the mitogenomics information—13 protein-coding genes, 2 rRNA, and 22 tRNA—and under different *a priori* partition schemes of the aligned sites. Fig. 3 displays the highest-likelihood trees based on the 123RNA and 12RNA partitions. Table 1 demonstrates how the ML bootstrap support for the main rodent clades varies according to the partitioning scheme: inclusion or exclusion of third codon positions of protein-coding genes, incorporation or not of the rRNA and tRNA alignments, and analysis of the 13 proteins. All phylogenies are in agreement with the current view of rodent molecular phylogenetics (*e.g.*, Churakov et al., 2010). In particular, with the additional contribution of the two new mitogenomes here analyzed, there is unambiguous support for the branching of *Pedetes* with *Anomalurus* (= Anomaluromorpha), and *Graphiurus* with *Glis* (= Gliridae).

Seven clades are recovered with maximal bootstrap support regardless of the partitioning scheme: Hystricognathi, Sciuroidea (= Sciuridae + Gliridae), Anomaluromorpha, Myodonta, Muroidea,

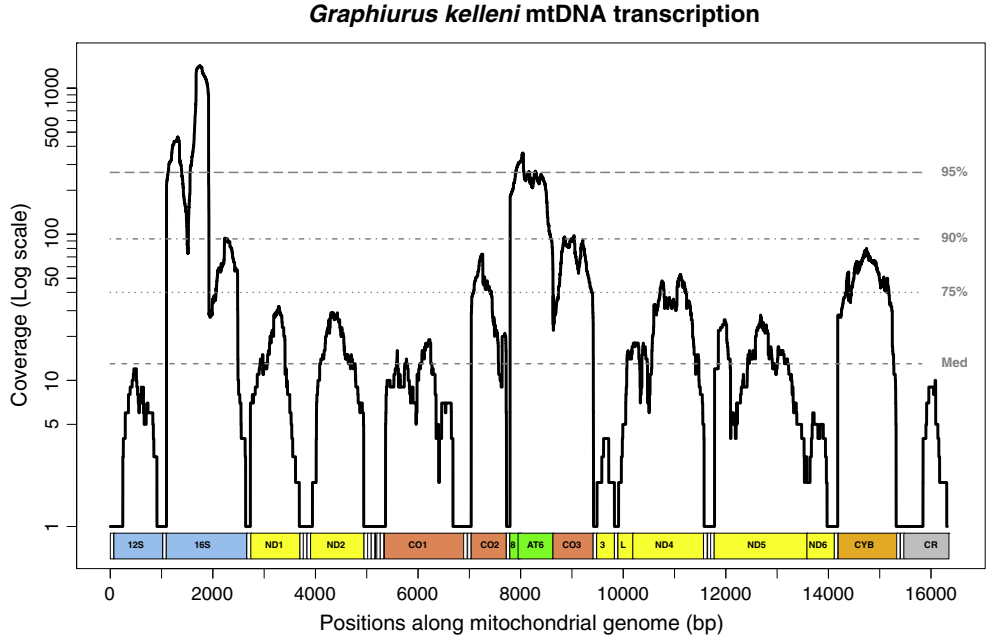

Fig. 2. Per-nucleotide RNA sequencing coverage of the mitogenomic transcripts of *Graphiurus kelleni*. The Y-axis provides the coverage with a logarithmic scale, and its maximum value exceeded 1400× in the 16S rRNA region. Horizontal grey dashed lines indicate the median (Med) coverage, and the 75%, 90%, and 95% quantiles. The map of the mitogenome is provided on the X-axis together with the nucleotide coordinates. Sequence names are provided, with tRNA left blank, rRNA in blue, NADH-dehydrogenases in yellow (3 = ND3; L = ND4L), cytochrome oxidases in coral, ATP-ases in green (8 = ATP8), and control region in grey.

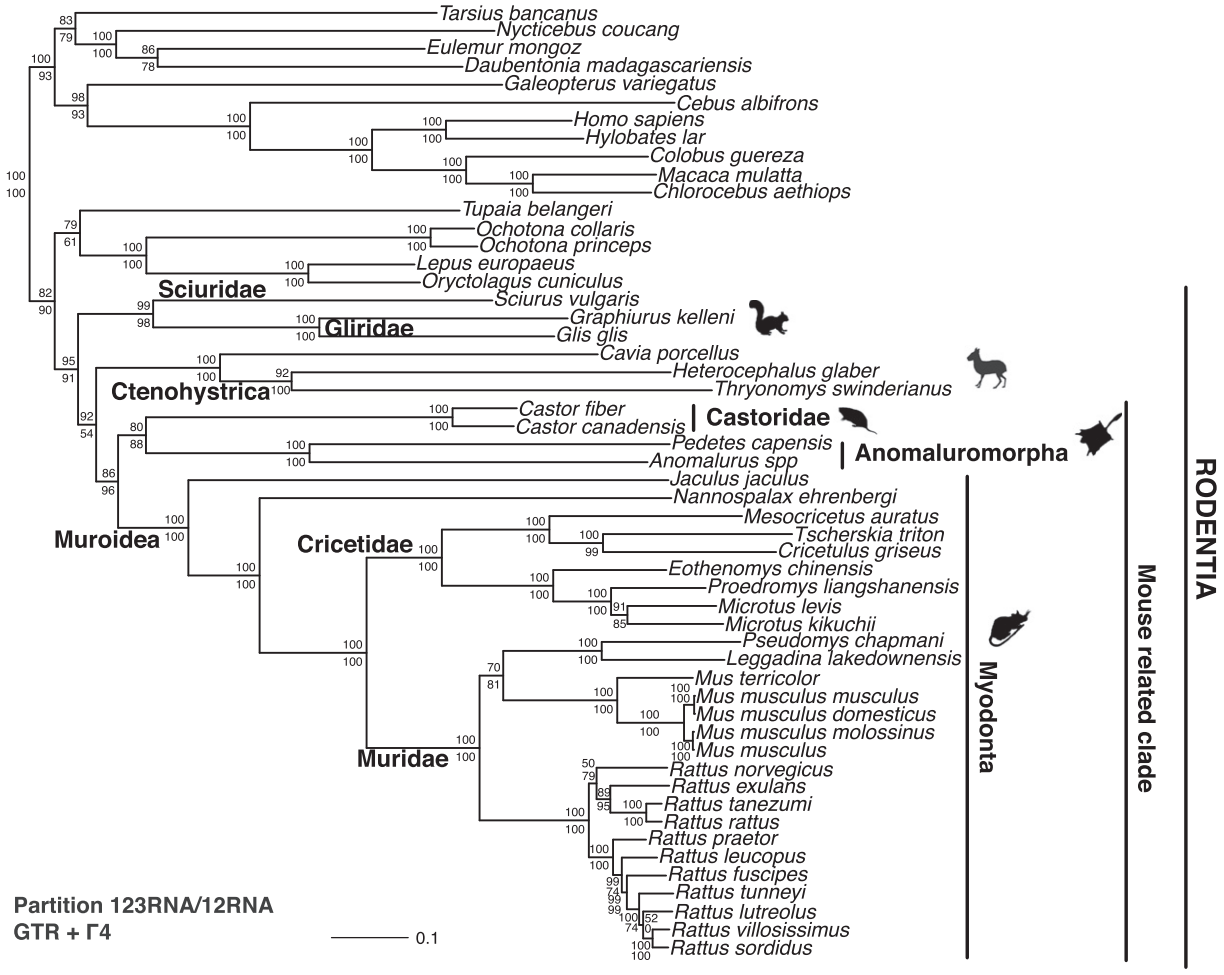

Fig. 3. Maximum likelihood topology obtained from the dataset concatenation of the 13 protein-coding, the two rRNA and the 22 tRNA genes, with inclusion (123RNA) or exclusion (12RNA) of third codon positions. Bootstrap percentages are indicated on each node and correspond to the 123RNA (above branches) and 12 RNA (below) partitions. The tree is rooted by primate taxa. All calculations have been conducted under RAxML. Systematic names and icons for the major rodent clades are given along the phylogeny.

Cricetidae, and Muridae (Fig. 3; Table 1). The support for 4 other major clades depends upon the character partition analyzed. (i) The grouping of Hystricognathi with the mouse-related clade is supported by the combination of all mitogenomics partitions, but the support decreases either with the exclusion of third codon positions or rRNA + tRNA genes. This grouping is not recovered at the protein level. (ii) The monophyly of the mouse-related clade is strongly supported by the rRNA + tRNA genes, and by their combination with the protein-coding genes. (iii) The Anomaluromorpha clusters with Castoridae with rather strong support ($80 < BP < 98$)

Table 1
Maximum likelihood bootstrap support for the main rodent clades under different character partitions. Third codon positions of the 13 protein-coding genes have been included (123) or excluded (12), and the two rRNA + 22 tRNA have been included (RNA) or not. All calculations have been conducted with RAxML under the GTR + Γ_4 (nucleotides) or WAG + Γ_4 (amino acids: AA) models. A dash indicates a clade not recovered under the corresponding partition analysis.

Clades	Maximum likelihood models and partitions					
	GTR + Γ_4	GTR + Γ_4	GTR + Γ_4	GTR + Γ_4	GTR + Γ_4	WAG + Γ_4
	123RNA	12RNA	123	12	RNA	AA
Rodentia	95	67	64	–	93	75
Ctenohystrica	100	100	100	100	100	100
Gliridae + Sciuridae	100	100	100	100	100	100
Ctenohystrica + mouse-related clade	92	78	72	54	–	–
Mouse-related clade	86	96	–	–	98	80
Anomaluromorpha	100	100	100	100	100	100
Anomaluromorpha + Castoridae	80	88	90	88	–	94
Myodonta	100	100	100	100	100	100
Muroidea	100	100	100	100	100	100
Cricetidae	100	100	100	100	100	100
Muridae	100	100	100	100	100	100

for all partitions but the rRNA + tRNA genes. (iv) The monophyly of the rodent mitogenomes here sampled is strongly supported by the partitions 123RNA and 12RNA, and by the amino acid matrix. However, the support decreases when third codon positions or rRNA + tRNA genes are excluded (partitions 12RNA, 123, and 12).

3.3. Bayesian phylogenetic inference with a posteriori site partitioning

To circumvent the potential problem of *a priori* defining the partitions of sites that are supposed to evolve according to the same substitution pattern, we used the CAT mixture model in a Bayesian framework. This model distributes *a posteriori* into distinct categories of nucleotide or amino acid composition the different alignment sites sharing similar evolutionary patterns. The corresponding maximum posterior probability phylograms are reported in Figs. 4A (partition 123RNA) andB (partition AA). Topologies of the ML and Bayesian trees are very similar, both at the DNA and protein levels.

Table 2 shows how posterior probabilities for the main rodent clades vary according to the choice of exchangeability matrices among amino acids. We distinguished two categories of clades following their statistical support. First, we found 7 clades (Hystricognathi, Gliridae + Sciuroidea, Anomaluromorpha, Myodonta, Muroidea, Cricetidae, Muridae) for which all the analyses yield the same topology with the highest posterior probabilities. Second, four nodes displayed different topologies depending on the model. This is the case for (i) Rodentia, (ii) Hystricognathi + mouse-related clade, (iii) the mouse-related clade and (iv) the Anomaluromorpha + Castoridae lineages for which some amino acid models (CAT + LG + Γ_4; CAT + POI + Γ_4, CAT + WAG + Γ_4) did not recover the same topology (Table 2). The monophyly of the mouse-related clade is not recovered under the CAT + POI + Γ_4 and CAT + LG + Γ_4 models. Anomaluromorpha is the sister group of Castoridae under all models but the CAT + LG + Γ_4 and CAT + WAG + Γ_4 models.

3.4. Evaluating the likelihood of alternative branching patterns

Our main topology supported a Hystricognathi + mouse-related clade. However, as suggested by ML topological tests on DNA data sets (see Table 3), we cannot exclude a Hystricognathi + Gliridae + Sciuroidea relationship ($P_{AU} > 0.05$), and a Gliridae + Sciuroidea + mouse-related relationship is only rejected for the 123RNA and the 123 partitions ($P_{AU} < 0.05$). Within the mouse-related clade, an Anomaluromorpha + Myodonta relationship is rejected ($P_{AU} < 0.05$) for the coding gene partitions only (123 and 12 partitions). This highlights that tRNA genes support a sister-relationship between Anomaluromorpha and Myodonta. All the AU tests rejected a Castorimorpha + mouse-related clade relationship for all the partitions ($P_{AU} < 0.05$). We also tested the hypothesis *Pedetes* + *Graphiurus*

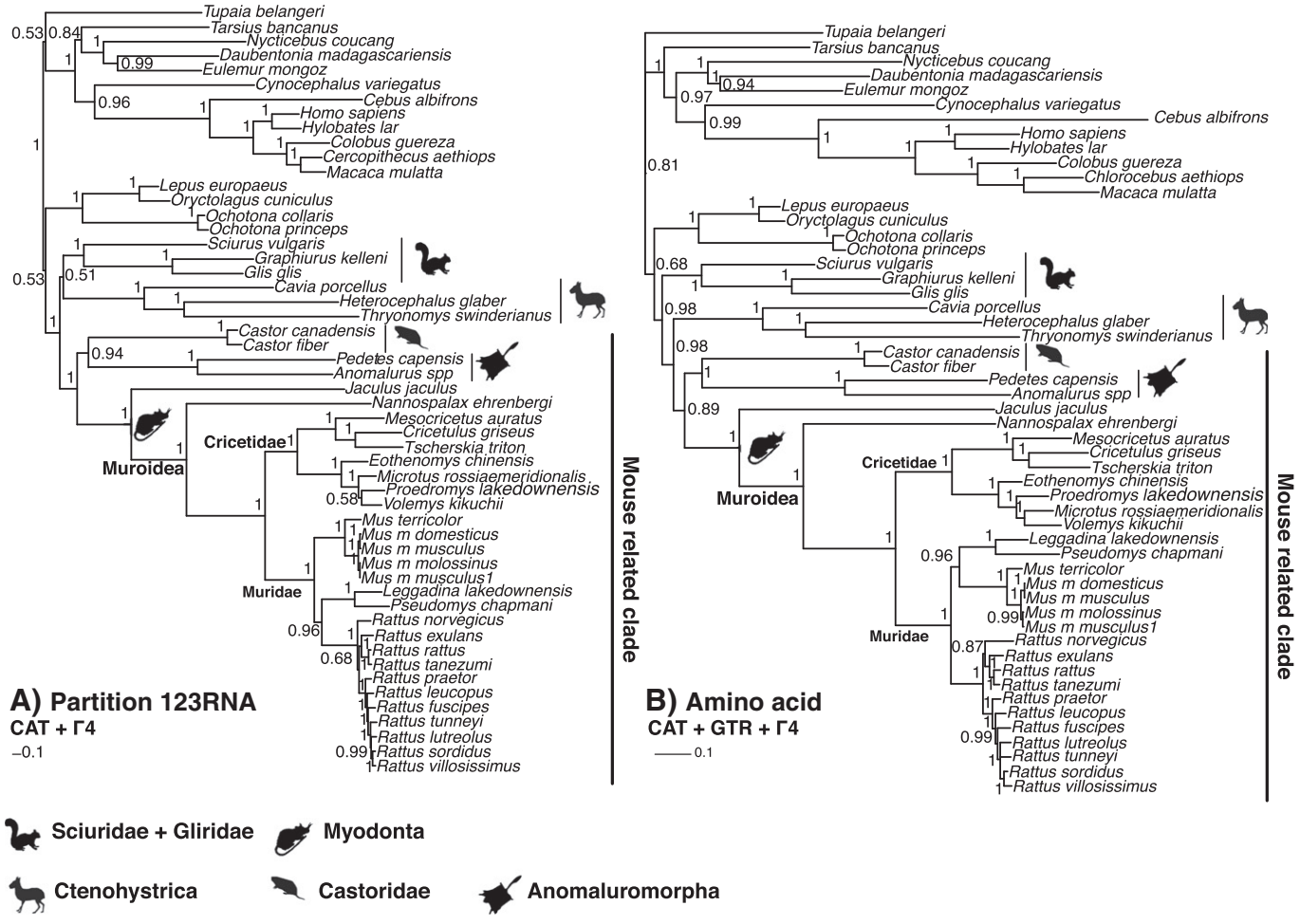

Fig. 4. Maximum posterior probability phylogram obtained from the Bayesian analyses of the nucleotide (A) and the amino acid (B) matrices. The DNA dataset includes the 13 protein-coding, the two rRNA and the 22 tRNA genes. The protein dataset includes the 13 mitochondrial proteins. Posterior probabilities are indicated on each node. Calculations have been conducted under the CAT + GTR + Γ (for DNA) and CAT + mtREV + Γ (for proteins) mixture models implemented in PhyloBayes. Clades are labeled with icons and text along the phylogeny.

Table 2
Bayesian posterior probabilities for the main rodent clades under different models of sequence evolution and character partitions. All calculations have been conducted with PhyloBayes under the CAT + Γ_4 model. For the DNA partitions 123RNA and 123, GTR exchangeabilities among nucleotides have been used. For the 13 proteins (partition AA), different exchangeabilities among amino acids have been used, with either estimation of all rates (GTR), or use of empirical rates: Poisson (POI), WAG, LG, mtZoa and mtREV. A dash indicates a clade not recovered under the corresponding model and partition analysis.

Clades	Bayesian models (PhyloBayes) and partitions							
	CAT +GTR +Γ_4 123	CAT +GTR +Γ_4 123 + RNA	CAT +GTR +Γ_4 AA	CAT +POI +Γ_4 AA	CAT +WAG +Γ_4 AA	CAT +LG +Γ_4 AA	CAT +mtZoa +Γ_4 AA	CAT +mtREV +Γ_4 AA
Rodentia	–	1	0.54	–	–	–	0.51	0.98
Ctenohystrica	1	1	1	1	1	1	1	1
Gliridae + Sciuridae	0.99	1	1	1	1	1	1	1
Ctenohystrica + mouse-related clade	–	–	0.57	–	–	–	0.53	0.98
Mouse related clade	–	0.94	0.98	–	1	–	0.98	0.89
Anomaluromorpha	1	1	1	1	1	1	1	1
Anomaluromorpha + Castoridae	1	1	1	1	–	–	1	1
Myodonta	1	1	1	1	1	1	1	1
Muroidea	1	1	1	1	1	1	1	1
Cricetidae	1	1	1	1	1	1	1	1
Muridae	1	1	1	1	1	1	1	1

(Vianey-Liaud et al., 1994) which was also rejected by all the AU tests ($P_{AU} < 0.05$).

4. Discussion

4.1. Mitogenome assembly based on 454 RNA-Seq

Deep transcriptome sequencing has proven useful to retrieve mitochondrial genomes in animals (Douglas et al., 2011; Jex et al., 2008; Nabholz et al., 2010; Rowe et al., 2011). Here, we used GS-FLX 454 reads to obtain *Graphiurus kelleni* and *Pedetes capensis* mitochondrial transcripts expressed in adult rodent fibroblasts (see Sections 2.1 and 3.1). From these analyses we obtained all 12 mtDNA protein-coding genes encoded from the heavy strand and the ND6 gene encoded on the light strand. A limitation of the high throughput RNA-Seq protocol is that some regions were not covered, and have to be filled through a standard PCR and Sanger protocol. This gap filling limitation of our mitogenomic RNA-Seq approach might reflect three mutually non-exclusive factors. (i) The RNA precursor resulting from the transcription of each mitochondrial DNA strand is polycistronic. This primary transcript is then processed to release individual mRNA, rRNA, and tRNA sequences. Folded tRNA likely act as secondary structure punctuations which are excised during mitochondrial RNA processing (Falkenberg et al., 2007). If tRNA sequences are rapidly eliminated during transcript maturation, they are less likely to be captured by the RNA-Seq approach. (ii) Different regions of the same mitogenome display differential transcription levels (Nabholz et al., 2013), meaning that the lower the expression, the lower the probability of recovering the corresponding transcripts. (iii) The coverage power of the 454 sequencing approach is not sufficiently deep to ensure the detection of all mitogenomic transcripts. Ultra deep NGS approaches, for example by the Illumina platform, may help to better detect transient tRNA molecules. NGS is a promising avenue to assemble mitogenomes from genomic and/or transcriptomic data, and the higher the coverage, the easier the task (Mason et al., 2011; Rowe et al., 2011).

4.2. Models of sequence evolution to analyze mitogenomic data

Several difficulties may hamper the extraction of evolutionary signal from complete mitogenomes when attempting to infer the rodent phylogeny. First, multiple substitutions on third codon positions of protein-coding genes introduce compositional heterogeneity and saturation. Second, the combination of mitochondrial sequences subjected to different selective pressures has to be analyzed under an adequate model. Moreover, owing to the marked compositional asymmetry of the two strands of the mitochondrial circle, the ND6 gene is generally excluded from phylogenetic analyses because it is encoded on the L-strand, while the 12 other mitochondrial protein-coding genes—all encoded on the H-strand—are concatenated into a single super-gene. The same remark holds at the amino acid level for which all proteins apart from ND6 are concatenated into a super-protein (*e.g.*, Horner et al., 2007). Instead of using more sophisticated models of sequence evolution to accommodate these features, often an important piece of information is excluded from most analyses.

To reduce the impact of multiple substitutions at the DNA level, a possibility is the removal of third codon positions. This does not change the nodal support for most nodes, including the affinities between Anomaluromorpha and Castoridae, but decreases the support for the monophyly of rodents, and the affinity between Ctenohystrica and the mouse-related clade (Table 1, Fig. 3). By contrast, the nucleotide

Table 3
Likelihood-based tests of alternative topologies within Rodentia. Results are computed under the GTR + Γ + INV model with different concatenated mitogenome DNA partitions: all codons with (123) or without (12) third positions, and with (RNA) or without rRNA + tRNA genes. The best ML topology is given in Fig. 3. Δln*L*: difference of log-likelihood with respect to the best tree. AU, Approximately Unbiased test (significant P-values in bold).

Hypotheses	123RNA		12RNA		123		12	
	Δln*L*	AU	Δln*L*	AU	Δln*L*	AU	Δln*L*	AU
Gliridae + Sciuroidea + Ctenohystrica	2.70	0.65	10.60	0.47	7.00	0.42	25.20	0.14
Gliridae + Sciuroidea + mouse-related clade	52.50	**0.01***	28.40	0.09	40.50	**0.01***	25.80	0.11
Anomaluromorpha + Myodonta	48.50	0.06	9.20	0.46	72.90	**0.01***	48.20	**0.01***
Castorimorpha + mouse-related clade	65.20	**0.02***	38.30	**0.03***	106.50	**0.01***	48.60	**0.01***

* Statistically significant at the 5% level.

information contained in the rRNA and tRNA genes supports the monophyly of rodents and of the mouse-related clade (Tables 1 and 2). This suggests that the combined use of partitions 123 and RNA is required to provide a better resolved rodent phylogeny: even if some sites are saturated by multiple substitutions, others may retain the evolutionary signal. We emphasize that it is important to incorporate the Γ distribution in the tree-building models in order to take into account this heterogeneity of substitution rates among sites within and also between the different mitochondrial partitions.

The use of amino acids provides an alternative way to reduce compositional heterogeneity and saturation problems with third codon positions. To circumvent the estimation of the 190 entries of amino acid rate exchangeability matrices, several empirical models have been designed. For example, models like WAG (Whelan and Goldman, 2001) and LG (Le and Gascuel, 2008), which were built on a wide spectrum of proteins, have been proposed. The analysis of expanded mitochondrial datasets led to the definition of models more focused on mitochondrial proteomics, such as mtREV (Adachi and Hasegawa, 1996; Yang et al., 1998) and mtZoa (Rota-Stabelli et al., 2009) for vertebrates and metazoans respectively. However, these models assume homogeneous exchangeabilities among sites, hence identically analyzing all positions of any partition of the super-gene or super-protein alignment. In addition, the use of a rate exchangeability matrix involves that at each site, each amino acid has the potential to be replaced by any of the 19 remaining amino acids, proportionally to the exchange rates. However, in practice, it is unlikely that every amino acid can occur at a given site. This is the reason why standard site-homogeneous models tend to have a biased estimate of the number of amino acid replacements (Douzery, 2011; Lartillot and Philippe, 2009).

As an alternative to reduce the above-mentioned problems of saturation and among-site substitution pattern heterogeneity, mixture models like CAT have been developed (Lartillot and Philippe, 2004). Here, the site-heterogeneous CAT model allows the convenient analysis of datasets combining different mitochondrial markers, including ND6. Because the CAT model groups sites into a number of independent categories defined *a posteriori*, it accounts for different site-specific nucleotide or amino acid preferences. Moreover, it has been shown that the CAT model is less impacted by systematic errors because it better detects multiple substitutions than an empirical model like WAG (Lartillot and Philippe, 2008).

The Bayesian analysis of rodent mitochondrial proteins along with the site-heterogeneous CAT + Γ mixture model, allowed us to infer a mitochondrial phylogeny congruent with published nuclear topologies, together with strong support for the Anomaluromorpha, Anomaluromorpha + Castorimorpha, and mouse-related clades. However, the different exchangeability matrices used in association with CAT + Γ provided different results for tricky nodes (Table 2). Whereas some models provide strong support for the mouse-related clade (CAT + Γ with GTR, WAG, mtZOA, or mtREV), other models do not support it (CAT + Γ with POI or LG). We found that the best resolved phylogeny (Fig. 4) is inferred from the empirical mtREV exchangeability matrix because it likely better describes the specificities of the replacement rate of the amino acids of the 13 mitochondrial proteins, together with CAT accounting for the among site heterogeneity in stationary amino acid frequencies.

4.3. Implications for phylogenetics of Anomaluridae and Pedetidae

The phylogenetic affinities of Anomaluridae (scaly-tailed flying squirrels) and Pedetidae (springhares) among rodents have been controversial, especially from the morpho-paleontological point of view (see Hartenberger and Luckett, 1985; Vianey-Liaud et al., 1994). Each family has for example been linked to different clades. Fossil data have related Anomaluridae to the North African Zegdoumyidae, together with the Graphiurinae subfamily of Gliridae (Vianey-Liaud and Jaeger, 1996; Vianey-Liaud et al., 1994). Based on cranio-dental and protein characters, several authors (de Jong, 1985; Doran, 1879; Flynn et al., 1985; Marivaux et al., 2004; Martin, 1995; von Koenigswald, 1985) have pointed out a possible relationship between Pedetidae and Hystricognathi or Ctenodactylidae rodents. By contrast, other morphological (Bugge, 1974; Gill, 1872; Lavocat and Parent, 1985; Marivaux et al., 2011; Ruf et al., 2010) and molecular (*e.g.*, Montgelard et al., 2002) studies suggest the grouping of Anomaluridae plus Pedetidae into the Anomaluromorpha clade.

Our mitogenomics results strongly support the affinities between Anomaluridae and Pedetidae. Complete mitogenomes therefore increase the body of molecular characters that supports the monophyly of Anomaluromorpha: individual mitochondrial genes (cytochrome *b* and 12S rRNA: Montgelard et al., 2002), nuclear markers (Blanga-Kanfi et al., 2009; Montgelard et al., 2008), retroposon data (Churakov et al., 2010), and a combination of molecular data (Huchon et al., 2007). In addition, morphological characters (Ruf et al., 2010) and paleontological data (Marivaux et al., 2011; Sallam et al., 2010a, 2010b) support this phylogenetic hypothesis. In particular, Ruf et al. (2010) studied the chorda tympani morphology and showed that Anomaluridae and Pedetidae share two synapomorphies (*i.e.*, the epitensoric character state, and the structure of the mallear channel). The cranio-dental analysis of Marivaux et al. (2011) also highlighted synapomorphic character states between these two families and the extinct African family Zegdoumyidae from the Early Eocene (45–50 Mya).

4.4. Implications for affinities of Anomaluromorpha among rodents

Our mitochondrial genome comparative analysis provides strong support for three major clades among rodents: (i) the squirrel-related clade (Sciuridae + Aplodontidae + Gliridae), (ii) the Ctenohystrica clade here represented by Hystricognathi and (iii) the mouse-related clade (Anomaluromorpha, Castoridae + Geomyoidea, and Myodonta). These phylogenetic results are in agreement with previous inferences based on nuclear genes (Adkins et al., 2001; Adkins et al., 2003; Blanga-Kanfi et al., 2009; DeBry, 2003; Huchon et al., 1999, 2002, 2007; Montgelard et al., 2008) and retroposon characters (Churakov et al., 2010). Moreover, mitogenome data provide moderate support for a relationship between Hystricognathi + mouse-related clade as suggested by the recent retroposon analysis of Churakov et al. (2010).

Based on the mitogenome analysis, Anomaluromorpha is shown to belong to the mouse-related clade, in agreement with nuclear gene phylogenies (Blanga-Kanfi et al., 2009; Huchon et al., 2002; Montgelard et al., 2008) and retrotransposon insertion patterns (Churakov et al., 2010). This result contrasts with another mitogenomic study suggesting that *Anomalurus* is a sister taxon to Hystricognathi (Horner et al., 2007). Such a branching pattern may have been caused by one of the factors known to affect phylogenetic inference: nucleotide/amino acid composition (Bernt et al., 2012), reduced taxon sampling in the previous mitogenomic analysis (Lecointre et al., 1993), differential evolutionary rates among lineages (Philippe et al., 2005), and model choice (Delsuc et al., 2005). In particular, we postulate that taxon sampling plays a key role, as the addition of *Pedetes* breaks the long, isolated branch of *Anomalurus*, and hence stabilizes the phylogenetic position of Anomaluromorpha with rodents.

The mouse-related clade includes three subclades: (1) Anomaluromorpha, (2) Castorimorpha which are divided into Castoridae + Geomyidae + Heteromyidae, and (3) Myodonta which are divided into Dipodoidea and Muroidea (Michaux and Catzeflis, 2000; Michaux et al., 2001; Steppan et al., 2004). Relationships among these three lineages are controversial and previous molecular analyses were either not conclusive (Blanga-Kanfi et al., 2009; Churakov et al., 2010) or provided moderate support to link Anomaluromorpha with Myodonta (Huchon et al., 2002; Montgelard et al., 2008). Our DNA and protein mitogenomics characters support a relationship between Castorimorpha and Anomaluromorpha (Figs. 3 and 4; Tables 1 to 3), in agreement with the phylogenetic analyses of Horn et al. (2011). Mitogenomic sequences of

Geomyidae and Heteromyidae are however required to confirm this evolutionary hypothesis.

5. Perspectives

Next generation DNA sequencing provides the opportunity to gather mitogenomes of non-model organisms. An RNA-Seq approach is convenient for identifying protein-coding genes for phylogenetic and expression purposes, although it has the potential limitation that all genes are not expressed at the same level (Nabholz et al., 2013). Conversely, a NGS approach based on the genomic DNA is more appropriate if one is interested in collecting rRNA and tRNA genes and building mitochondrial maps to study gene order and gene content (D'Onorio de Meo et al., 2012), as this method will ensure a more even coverage of all regions of the mitochondrial circle.

The present update of rodent mitogenomes highlights the importance of denser taxonomic sampling as we recovered a mitochondrial phylogeny in better agreement with nuclear gene trees (see also Reyes et al., 2004). Comparative mitogenomics provides phylogenetic information for the deep nodes of rodent phylogeny and also further suggests the branching of Anomaluridae + Pedetidae with Castoridae within the mouse-related clade. Perspectives into the comparative mitogenomics and phylogeny of rodents would be (i) to use time-heterogeneous models to take into account potential branch specific variations of compositional and substitution patterns (Blanquart and Lartillot, 2008; Dutheil et al., 2012; Zhou et al., 2007), and (ii) to improve taxonomic sampling by using deep sequencing methodologies. Resolution of the phylogenetic relationships among the major mammalian clades will benefit from growing mitogenome collections which contain valuable information to be extracted in a probability framework with adequate, non-homogeneous models.

Conflict of interest

The authors have no conflicts of interest to declare.

Acknowledgments

We thank Gauthier Dobigny and Terry Robinson for providing *Graphiurus* and *Pedetes* resources, François Catzeflis for access to the mammalian tissue collection of Montpellier, and Marie-ka Tilak for her invaluable help in initiating the project. We thank Dmitri Kramerov, Kenneth Puliafico, and two anonymous reviewers for helpful suggestions on the manuscript. This work has been supported by the Agence Nationale de la Recherche "Bio-Informatique" (ANR-10-BINF-01 "Ancestrome"), and by the bioinformatics cluster of the ISE-M. This publication is contribution No. 2013-097 of the Institut des Sciences de l'Evolution de Montpellier (UMR 5554—CNRS). P-HF and KAJ acknowledges the Danish National Research Foundation for providing funding to the Center for Macroecology, Evolution and Climate.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gene.2013.07.059.

References

Adachi, J., Hasegawa, M., 1996. Model of amino acid substitution in proteins encoded by mitochondrial DNA. J. Mol. Evol. 42, 459–468.
Adkins, R.M., Gelke, E.L., Rowe, D., Honeycutt, R.L., 2001. Molecular phylogeny and divergence time estimates for major rodent groups: evidence from multiple genes. Mol. Biol. Evol. 18, 777–791.
Adkins, R.M., Walton, A.H., Honeycutt, R.L., 2003. Higher-level systematics of rodents and divergence time estimates based on two congruent nuclear genes. Mol. Phylogenet. Evol. 26, 409–420.
Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.
Arnason, U., Adegoke, J.A., Gullberg, A., Harley, E.H., Janke, A., Kullberg, M., 2008. Mitogenomic relationships of placental mammals and molecular estimates of their divergences. Gene 421, 37–51.
Bernt, M., Braband, A., Middendorf, M., Misof, B., Rota-Stabelli, O., Stadler, P.F., 2012. Bioinformatics methods for the comparative analysis of metazoan mitochondrial genome sequences. Mol. Phylogenet. Evol. http://dx.doi.org/10.1016/j.ympev.2012.09.019.
Blanga-Kanfi, S., Miranda, H., Penn, O., Pupko, T., DeBry, R.W., Huchon, D., 2009. Rodent phylogeny revised: analysis of six nuclear genes from all major rodent clades. BMC Evol. Biol. 9, 71.
Blanquart, S., Lartillot, N., 2008. A site- and time-heterogeneous model of amino acid replacement. Mol. Biol. Evol. 25, 842–858.
Boore, J.L., 1999. Animal mitochondrial genomes. Nucleic Acids Res. 27, 1767–1780.
Bugge, J., 1974. The cephalic arterial system in insectivores, primates, rodents and lagomorphs, with special reference to the systematic classification. Acta Anat. 87, 1–159.
Carleton, M., Musser, G.G., 2005. Order Rodentia, In: Wilson, E.D., Reeder, D.M. (Eds.), Mammal Species of the World: A Taxonomic and Geographic Reference, 3rd ed. Johns Hopkins University Press, Baltimore, pp. 895–1531.
Churakov, G., Sadasivuni, M.K., Rosenbloom, K.R., Huchon, D., Brosius, J., Schmitz, J., 2010. Rodent evolution: back to the root. Mol. Biol. Evol. 27, 1315–1326.
de Jong, W., 1985. Superordinal affinities of Rodentia studied by sequence analysis of eye lens protein. In: Luckett, W.P., Hartenberger, J.-L. (Eds.), Evolutionary Relationships Among Rodents: A Multidisciplinary Analysis. Plenum Press, New York and London, pp. 211–226.
DeBry, R.W., 2003. Identifying conflicting signal in a multigene analysis reveals a highly resolved tree: the phylogeny of Rodentia (Mammalia). Syst. Biol. 52, 604–617.
Delsuc, F., Brinkmann, H., Philippe, H., 2005. Phylogenomics and the reconstruction of the tree of life. Nat. Rev. Genet. 6, 361–375.
D'Erchia, A.M., Gissi, C., Pesole, G., Saccone, C., Arnason, U., 1996. The guinea-pig is not a rodent. Nature 381, 597–600.
D'Onorio de Meo, P., et al., 2012. MitoZoa 2.0: a database resource and search tools for comparative and evolutionary analyses of mitochondrial genomes in Metazoa. Nucleic Acids Res. 40, D1168–D1172.
Doran, A.H.G., 1879. The mammalian ossicula auditus. Trans. Linn. Soc. Lond. Zool. 2nd Ser. 1, 371–497.
Douglas, K.C., Halbert, N.D., Kolenda, C., Childers, C., Hunter, D.L., Derr, J.N., 2011. Complete mitochondrial DNA sequence analysis of *Bison bison* and bison–cattle hybrids: function and phylogeny. Mitochondrion 11, 166–175.
Douzery, E.J.P., 2011. Molecular phylogeny: inferring the patterns of evolution. In: Gargaud, M., Lopez-Garcia, P., Martin, H. (Eds.), Origins and Evolution of Life: An Astrobiological Perspective. Cambridge University Press, Cambridge, pp. 291–312.
Drummond, A.J., et al., 2011. Geneious v5.4. (available from http://www.geneious.com/).
Dutheil, J.Y., Galtier, N., Romiguier, J., Douzery, E.J.P., Ranwez, V., Boussau, B., 2012. Efficient selection of branch-specific models of sequence evolution. Mol. Biol. Evol. 29, 1861–1874.
Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797.
Eisenberg, J.F., 1981. The Mammalian Radiations: Analysis of Trends in Evolution, Adaptation, and Behaviour. The University of Chicago Press, Chicago.
Fabre, P.-H., Hautier, L., Dimitrov, D., Douzery, E.J.P., 2012. A glimpse on the pattern of rodent diversification: a phylogenetic approach. BMC Evol. Biol. 12, 88.
Falkenberg, M., Larsson, N.G., Gustafsson, C.M., 2007. DNA replication and transcription in mammalian mitochondria. Annu. Rev. Biochem. 2007 (76), 679–699.
Felsenstein, J., 2004. Inferring Phylogenies. Sinauer Associates, Sunderland, Massachusetts.
Flynn, J.J., Jacobson, L., Lindsay, E.H., 1985. Problems to other rodents and origin of major groups. In: Luckett, W.P., Hartenberger, J.-L. (Eds.), Evolutionary Relationships Among Rodents: A Multidisciplinary Analysis. Plenum Press, New York and London, pp. 589–616.
Foster, P.G., Jermiin, L.S., Hickey, D.A., 1997. Nucleotide composition bias affects amino acid content in proteins coded by animal mitochondria. J. Mol. Evol. 44, 282–288.
Galtier, N., Nabholz, B., Glemin, S., Hurst, G.D.D., 2009. Mitochondrial DNA as a marker of molecular diversity: a reappraisal. Mol. Ecol. 18, 4541–4550.
Gill, T., 1872. Arrangement of the Families of Mammals with Analytical Tables. Smithsonian Miscellaneous Collections, Washington, 230 1–98.
Graur, D., Hide, W.A., Li, W.-H., 1991. Is the guinea pig a rodent? Nature 351, 649–652.
Hartenberger, J.-L., 1985. The order Rodentia: major questions on their evolutionary origin, relationships and suprafamilial systematics. In: Luckett, W.P., Hartenberger, J.-L. (Eds.), Evolutionary Relationships Among Rodents: A Multidisciplinary Analysis. Plenum Press, New York and London, pp. 1–33.
Hartenberger, J.-L., Luckett, W.P., 1985. Evolutionary Relationships Among Rodents: A Multidisciplinary Analysis. Plenum Press, New York and London.
Hautier, L., Michaux, J., Marivaux, L., Vianey-Liaud, M., 2008. The evolution of the zygomasseteric construction in Rodentia, as revealed by a geometric morphometric analysis of the mandible of *Graphiurus* (Rodentia, Gliridae). Zool. J. Linn. Soc. 154, 807–821.
Hautier, L., Lebrun, R., Saksiri, S., Michaux, J., Vianey-Liaud, M., Marivaux, L., 2011. Hystricognathy vs sciurognathy in the rodent jaw: a new morphometric assessment of hystricognathy applied to the living fossil *Laonastes* (Diatomyidae). PLoS One 6, e18698.
Horn, S., et al., 2011. Mitochondrial genomes reveal slow rates of molecular evolution and the timing of speciation in beavers (*Castor*), one of the largest rodent species. PLoS One 6, e14622.
Horner, D.S., Lefkimmiatis, K., Reyes, A., Gissi, C., Saccone, C., Pesole, G., 2007. Phylogenetic analyses of complete mitochondrial genome sequences suggest a basal divergence of the enigmatic rodent Anomalurus. BMC Evol. Biol. 7, 16.

Huchon, D., Douzery, E.J.P., 2001. From the Old World to the New World: a molecular chronicle of the phylogeny and biogeography of hystricognath rodents. Mol. Phylogenet. Evol. 20, 239–251.
Huchon, D., Catzeflis, F.M., Douzery, E.J.P., 1999. Molecular evolution of the nuclear von Willebrand Factor gene in mammals and the phylogeny of rodents. Mol. Biol. Evol. 16, 577–589.
Huchon, D., et al., 2002. Rodent phylogeny and a timescale for the evolution of Glires: evidence from an extensive taxon sampling using three nuclear genes. Mol. Biol. Evol. 19, 1053–1065.
Huchon, D., Chevret, P., Jordan, U., Kilpatrick, C.W., 2007. Multiple molecular evidences for a living mammalian fossil. Proc. Natl. Acad. Sci. U. S. A. 104, 7495–7499.
Janecka, J.E., et al., 2007. Molecular and genomic data identify the closest living relative of primates. Science 318, 792–794.
Jex, A.R., Hu, M., Littlewood, D.T.J., Waeschenbach, A., Gasser, R.B., 2008. Using 454 technology for long-PCR based sequencing of the complete mitochondrial genome from single *Haemonchus contortus* (Nematoda). BMC Genomics 9, 11.
Kay, E.H., Hoekstra, H.E., 2008. Rodents. Curr. Biol. 18, R406–R410.
Kramerov, D., Vassetzky, N., Serdobova, I., 1999. The evolutionary position of dormice (Gliridae) in Rodentia determined by a novel short retroposon. Mol. Biol. Evol. 16, 715–717.
Lartillot, N., Philippe, H., 2004. A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. Mol. Biol. Evol. 21, 1095–1109.
Lartillot, N., Philippe, H., 2008. Improvement of molecular phylogenetic inference and the phylogeny of Bilateria. Philos. Trans. R. Soc. Lond. B Biol. Sci. 363, 1463–1472.
Lartillot, N., Philippe, H., 2009. Improvement of molecular phylogenetic inference and the phylogeny of Bilateria. In: Telford, M.J., Littlewood, D.T.J. (Eds.), Animal Evolution: Genomes, Fossils, and Trees. Oxford University Press, Oxford, pp. 127–138.
Lartillot, N., Brinkmann, H., Philippe, H., 2007. Suppression of long-branch attraction artefacts in the animal phylogeny using a site-heterogeneous model. BMC Evol. Biol. 7 (1), S4.
Lartillot, N., Lepage, T., Blanquart, S., 2009. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. Bioinformatics 25, 2286–2288.
Lavocat, R., Parent, J.-P., 1985. Phylogenetic analysis of the middle ear features in fossil and living rodents. In: Luckett, W.P., Hartenberger, J.-L. (Eds.), Evolutionary Relationships Among Rodents: A Multidisciplinary Analysis. Plenum Press, New York and London, pp. 333–354.
Le, S.Q., Gascuel, O., 2008. An improved general amino acid replacement matrix. Mol. Biol. Evol. 25, 1307–1320.
Lecointre, G., Philippe, H., Van Le, H.L., Le Guyader, H., 1993. Species sampling has a major impact on phylogenetic inference. Mol. Phylogenet. Evol. 2, 205–224.
Luckett, W.P., Hartenberger, J.-L., 1993. Monophyly or polyphyly of the order Rodentia: possible conflict between morphological and molecular interpretations. J. Mammal. Evol. 1, 127–147.
Marivaux, L., Vianey-Liaud, M., Jaeger, J.J., 2004. High-level phylogeny of early Tertiary rodents: dental evidence. Zool. J. Linn. Soc. 142, 105–134.
Marivaux, L., et al., 2011. Zegdoumyidae (Rodentia, Mammalia), stem anomaluroid rodents from the early to middle Eocene of Algeria (Gour Lazib, Western Sahara): new dental evidence. J. Syst. Palaeontol. 9, 563–588.
Martin, T., 1995. Incisor enamel microstructure and phylogenetic interrelationships of Pedetidae and Ctenodactylidae (Rodentia). Berl. Geowiss. Abh. Reihe E. Paläobiol. 16, 693–707.
Mason, V.C., Li, G., Helgen, K.M., Murphy, W.J., 2011. Efficient cross-species capture hybridization and next-generation sequencing of mitochondrial genomes from noninvasively sampled museum specimens. Genome Res. 21, 1695–1704.
Meredith, R.W., et al., 2011. Impacts of the Cretaceous Terrestrial Revolution and KPg extinction on mammal diversification. Science 334, 521–524.
Meyer, M., Stenzel, U., Hofreiter, M., 2008. Parallel tagged sequencing on the 454 platform. Nat. Protoc. 3, 267–278.
Michaux, J., Catzeflis, F., 2000. The bushlike radiation of muroid rodents is exemplified by the molecular phylogeny of the LCAT nuclear gene. Mol. Phylogenet. Evol. 17, 280–293.
Michaux, J., Reyes, A., Catzeflis, F., 2001. Evolutionary history of the most speciose mammals: molecular phylogeny of muroid rodents. Mol. Biol. Evol. 18, 2017–2031.
Miller, W., et al., 2009. The mitochondrial genome sequence of the Tasmanian tiger (*Thylacinus cynocephalus*). Genome Res. 19, 213–220.
Miller, J.M., Malenfant, R.M., Moore, S.S., Coltman, D.W., 2012. Short reads, circular genome: skimming SOLiD sequence to construct the bighorn sheep mitochondrial genome. J. Hered. 103, 140–146.
Montgelard, C., Bentz, S., Tirard, C., Verneau, O., Catzeflis, F.M., 2002. Molecular systematics of Sciurognathi (Rodentia): the mitochondrial cytochrome *b* and 12S rRNA genes support the Anomaluroidea (Pedetidae and Anomaluridae). Mol. Phylogenet. Evol. 22, 220–233.
Montgelard, C., Matthee, C.A., Robinson, T.J., 2003. Molecular systematics of dormice (Rodentia: Gliridae) and the radiation of *Graphiurus* in Africa. Proc. R. Soc. Lond. B 270, 1947–1955.
Montgelard, C., Forty, E., Arnal, V., Matthee, C.A., 2008. Suprafamilial relationships among Rodentia and the phylogenetic effect of removing fast-evolving nucleotides in mitochondrial, exon and intron fragments. BMC Evol. Biol. 8, 321.
Nabholz, B., Mauffrey, J.-F., Bazin, E., Galtier, N., Glemin, S., 2008a. Determination of mitochondrial genetic diversity in mammals. Genetics 178, 351–361.
Nabholz, B., Glémin, S., Galtier, N., 2008b. Strong variations of mitochondrial mutation rate across mammals: the longevity hypothesis. Mol. Biol. Evol. 25, 120–130.
Nabholz, B., Glemin, S., Galtier, N., 2009. The erratic mitochondrial clock: variations of mutation rate, not population size, affect mtDNA diversity across birds and mammals. BMC Evol. Biol. 9, 54.
Nabholz, B., Jarvis, E.D., Ellegren, H., 2010. Obtaining mtDNA genomes from next-generation transcriptome sequencing: a case study on the basal Passerida (Aves: Passeriformes) phylogeny. Mol. Phylogenet. Evol. 57, 466–470.
Nabholz, B., Ellegren, H., Wolf, J.B.W., 2013. High levels of gene expression explain the strong evolutionary constraint of mitochondrial protein-coding genes. Mol. Biol. Evol. 30, 272–284.
Nunome, M., Yasuda, S.P., Sato, J.J., Vogel, P., Suzuki, H., 2007. Phylogenetic relationships and divergence times among dormice (Rodentia, Gliridae) based on three nuclear genes. Zool. Scr. 36, 537–546.
Philippe, H., 1997. Rodent monophyly: pitfalls of molecular phylogenies. J. Mol. Evol. 45, 712–715.
Philippe, H., Zhou, Y., Brinkmann, H., Rodrigue, N., Delsuc, F., 2005. Heterotachy and long-branch attraction in phylogenetics. BMC Evol. Biol. 5, 50.
Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817–818.
Reyes, A., Pesole, G., Saccone, C., 1998. Complete mitochondrial DNA sequence of the fat dormouse, *Glis glis*: further evidence of rodent paraphyly. Mol. Biol. Evol. 15, 499–505.
Reyes, A., Pesole, G., Saccone, C., 2000. Long-branch attraction phenomenon and the impact of among-site rate variation on rodent phylogeny. Gene 23, 177–187.
Reyes, A., Gissi, C., Catzeflis, F., Nevo, E., Pesole, G., Saccone, C., 2004. Congruent mammalian trees from mitochondrial and nuclear genes using Bayesian methods. Mol. Biol. Evol. 21, 397–403.
Robins, J.H., McLenachan, P.A., Phillips, M.J., Craig, L., Ross, H.A., Matisoo-Smith, E., 2008. Dating of divergences within the *Rattus* genus phylogeny using whole mitochondrial genomes. Mol. Phylogenet. Evol. 49, 460–466.
Robins, J.H., McLenachan, P.A., Phillips, M.J., McComish, B.J., Matisoo-Smith, E., Ross, H.A., 2010. Evolutionary relationships and divergence times among the native rats of Australia. BMC Evol. Biol. 10, 375.
Rodriguez, F., Oliver, J., Marin, A., Medina, J., 1990. The general stochastic model of nucleotide substitutions. J. Theor. Biol. 142, 485–501.
Rota-Stabelli, O., Yang, Z., Telford, M.J., 2009. MtZoa: a general mitochondrial amino acid substitutions model for animal evolutionary studies. Mol. Phylogenet. Evol. 52, 268–272.
Rowe, K.C., et al., 2011. Museum genomics: low-cost and high-accuracy genetic data from historical specimens. Mol. Ecol. Resour. 11, 1082–1092.
Ruf, I., Frahnert, S., Maier, W., 2010. The chorda tympani and its significance for rodent phylogeny. Mamm. Biol. 74, 100–113.
Ryu, S.H., Kwak, M.J., Hwang, U.W., 2013. Complete mitochondrial genome of the Eurasian flying squirrel *Pteromys volans* (Sciuromorpha, Sciuridae) and revision of rodent phylogeny. Mol. Biol. Rep. 40, 1917–1926.
Sallam, H.M., Seiffert, E.R., Simons, E.L., 2010a. A highly derived Anomalurid rodent (Mammalia) from the earliest late Eocene of Egypt. Palaeontology 53, 803–813.
Sallam, H.M., Seiffert, E.R., Simons, E.L., Brindley, C., 2010b. A large-bodied anomaluroid rodent from the earliest late Eocene of Egypt: phylogenetic and biogeographic implications. J. Vertebr. Paleontol. 30, 1579–1593.
Shimodaira, H., 2002. An approximately unbiased test of phylogenetic tree selection. Syst. Biol. 51, 492–508.
Shimodaira, H., Hasegawa, M., 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics 17, 1246–1247.
Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690.
Steppan, S.J., Adkins, R.M., Anderson, J., 2004. Phylogeny and divergence-date estimates of rapid radiations in muroid rodents based on multiple nuclear genes. Syst. Biol. 53, 533–553.
Swofford, D.L., 2002. PAUP*. Phylogenetic analysis using parsimony (and other methods), 4.
Veniaminova, N.A., Vassetzky, N.S., Kramerov, D.A., 2007. B1 SINEs in different rodent families. Genomics 89, 678–686.
Vianey-Liaud, M., Jaeger, J.-J., 1996. A new hypothesis for the origin of African Anomaluridae and Graphiuridae (Rodentia). Palaeovertebrata 25, 349–358.
Vianey-Liaud, M., Jaeger, J.-J., Hartenberger, J.-L., Mahboubi, M., 1994. Les rongeurs de l'Eocène d'Afrique Nord-Occidentale (Glib Zegdou (Algérie) et Chambi (Tunisie)) et l'origine des Anomaluridae. Palaeovertebrata 23, 93–118.
von Koenigswald, W., 1985. Evolutionary trends in the enamel of rodent incisors. In: Luckett, W.P., Hartenberger, J.-L. (Eds.), Evolutionary Relationships Among Rodents: A Multidisciplinary Analysis. Plenum Press, New York and London, pp. 211–226.
Wang, Z., Gerstein, M., Snyder, M., 2009. RNA-Seq: a revolutionary tool for transcriptomics. Nat. Rev. Genet. 10, 57–63.
Whelan, S., Goldman, N., 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol. Biol. Evol. 18, 691–699.
Wood, A., 1965. Grades and clades among rodents. Evolution 19, 115–130.
Yang, Z., 1996. Among-site rate variation and its impact on phylogenetic analyses. Trends Ecol. Evol. 11, 367–372.
Yang, Z., Nielsen, R., Hasegawa, M., 1998. Models of amino acid substitution and applications to mitochondrial protein evolution. Mol. Biol. Evol. 15, 1600–1611.
Zhou, Y., Rodrigue, N., Lartillot, N., Philippe, H., 2007. Evaluation of the models handling heterotachy in phylogenetic inference. BMC Evol. Biol. 7, 206.