THE ROYAL SOCIETY

# The mitochondrial DNA molecule of the aardvark, *Orycteropus afer*, and the position of the Tubulidentata in the eutherian tree

**Ulfur Arnason*, Anette Gullberg and Axel Janke**

*Division of Evolutionary Molecular Systematics, University of Lund, Sölvegatan 29, S-223 62 Lund, Sweden*

An outstanding problem in mammal phylogeny is the relationship of the aardvark (*Orycteropus afer*), the only living species of the order Tubulidentata, to the extant eutherian lineages. In order to examine this problem the complete mitochondrial DNA (mtDNA) molecule of the aardvark was sequenced and analysed. The aardvark tRNA-Ser(UCN) differs from that of other mammalian mtDNAs reported and appears to have reversed to the ancestral secondary structure of non-mammalian vertebrates and mitochondrial tRNAs in general. Phylogenetic analysis of 12 concatenated protein-coding genes (3325 amino acids) included the aardvark and 15 additional eutherians, two marsupials and a monotreme. The most strongly supported tree identified the aardvark as a sister group of a clade including the armadillo (Xenarthra) and the Cetferungulata (carnivores, perissodactyls, artiodactyls and cetaceans). By applying three molecular calibration points the divergence between the aardvark and armadillo–cetferungulates was estimated at *ca.* 90 million years before present.

**Keywords:** mammalian evolution; mitochondrial DNA; Tubulidentata; aardvark; African clade

## 1. INTRODUCTION

During the past few decades new methods of phylogenetic analysis and the advent of molecular data have reinvigorated higher level mammalian systematics. The availability of large numbers of phylogenetically informative characters provided by molecular data has made systematists optimistic that it will be possible to recover the dichotomous eutherian tree (for a review, see Graur 1993). Several ordinal mammalian relationships have been resolved by means of phylogenetic analyses of complete mitochondrial DNA (mtDNA) molecules. The three mammalian infraclasses, Monotremata, Marsupialia and Eutheria, are all represented by complete mtDNAs. The eutherian representation includes the orders Lipotyphla, Rodentia, Lagomorpha, Primates, Artiodactyla, Cetacea, Carnivora, Perissodactyla and Xenarthra (Edentata). The rationale for establishing the large data sets of complete mtDNAs is that short sequences, such as those of individual genes, may yield different topologies for the same assembly of taxa (Arnason & Johnsson 1992), whereas statistical analyses have shown that the stochastic effects occurring in small data sets, mitochondrial as well as nuclear, are gradually reduced with increasing lengths of alignments (Cao *et al.* 1994).

Besides resolving mammalian relationships, analyses of complete mtDNAs have made it possible to estimate the time of several divergences among mammals. The calibration points applied in these analyses have been eutherian in origin set at 130 million years before present (Myr BP) (Janke *et al.* 1994) and the interordinal divergence between artiodactyls (as represented by the cow) and cetaceans set at 60 Myr BP (A/C-60) (Arnason & Gullberg 1996). More recently, these references have been complemented with a third molecular calibration point, the intraordinal perissodactyl reference E/R-50, the divergence between Equidae and Rhinocerotidae set at 50 Myr BP (Xu *et al.* 1996). The palaeontological support for the two latter references is greater than that of any other comparable eutherian references.

Application of A/C-60 and E/R-50 has made it possible to estimate the time of various mammalian divergences, both inter- and intraordinal, with considerable consistency, the datings of which had been conjectural due to the incompleteness of the fossil record (Arnason *et al.* 1996*a*, 1997, 1998; Janke *et al.* 1997). The most basal eutherian divergence so far recognized is that between Lipotyphla, as represented by the hedgehog, *Erinaceus europaeus* and remaining eutherians (Krettek *et al.* 1995). Application of A/C-60 on various mammalian divergences suggests that interordinal eutherian divergences took place over an extended period of time between 60 and 120 Myr BP. Superficially, the wide time span may suggest that most ordinal mammalian divergences will be resolved once molecular data sets of sufficient sizes become available. This may be an over-optimistic view, however, since the divergence of many orders took place within the relatively narrow time span between 90 and 110 Myr BP, i.e. in the mid-Cretaceous period (Janke *et al.* 1994, 1997; Xu *et al.* 1996; Arnason *et al.* 1996*a*, 1997), thereby making it difficult to resolve their relationships due to the short evolutionary distance between lineages and the long time since their divergence.

*Author for correspondence (ulfur.arnason@gen.lu.se).


*Received* 17 September 1998 *Accepted* 15 October 1998

One particularly vexing lineage for systematists is the order Tubulidentata, represented by one recent species, the aardvark, *Orycteropus afer*. Some anatomical studies have aligned the aardvark with the Paenungulata (Le Gros Clark & Sonntag 1926; Carroll 1988) and an African origin of this clade, including Tubulidentata, Proboscidea, Sirenia and Hyracoidea, has been proposed (Carroll 1988). Earlier molecular studies demonstrated affinities between the aardvark and the paenungulates (de Jong *et al.* 1981; Rainey *et al.* 1984; Miyamoto & Goodman 1986) and these findings have been supported in more recent molecular studies which have extended the clade to also include golden moles and elephant shrews (de Jong *et al.* 1993; Lavergne *et al.* 1996; Stanhope *et al.* 1996; Madsen *et al.* 1997; Springer *et al.* 1997*a*,*b*), but, relative to other orders, the position of the African clade in the eutherian tree remains unknown.

In the present study we have explored the phylogenetic position of the Tubulidenta based on analyses of complete mtDNAs and by rooting the eutherian tree with a non-eutherian mammalian outgroup, the platypus and two marsupials. Non-eutherian outgroups are necessary since the order Xenarthra (Edentata), as represented by the armadillo, *Dasypus novemcinctus*, is not basal in the eutherian tree (Arnason *et al.* 1997), an assumption made in many previous phylogenetic analyses of eutherian relationships.

## 2. METHODS

Mitochondrial DNA was isolated from frozen tissue samples of the aardvark following previously established procedures (Arnason *et al.* 1991). The tissue was supplied by Dr J. Wensing (Burgers Zoo, Arnheim, The Netherlands). The mtDNA was digested separately with *Hin*dIII and *Bcl*I. The digested DNA was either ligated directly into phage M13mp18 or run on a preparative agarose gel from which the restriction fragments were excised and electroeluted before ligation. The entire mtDNA molecule was covered with natural (not PCR-amplified) and, in most instances, overlapping clones. Sequencing was carried out manually on single-stranded DNA applying the dideoxy termination technique (Sanger 1981) with $^{35}$SdATP, using both universal and numerous specific sequencing primers. In addition to the natural clones, the sequence of the *tRNA-Ser*(*UCN*) gene of the aardvark was determined in ten cloned PCR products, covering both strands.

The complete sequence of the mtDNA of the aardvark has been deposited at EMBL with accession number Y18475. Users of the sequence are kindly requested to refer to the present paper and not only to the accession number of the sequence.

The phylogenetic alignment included the following species: platypus, *Ornithorhynchus anatinus* (Janke *et al.* 1996), opossum, *Didelphis virginiana* (Janke *et al.* 1994), wallaroo, *Macropus robustus* (Janke *et al.* 1997), hedgehog, *Erinaceus europaeus* (Krettek *et al.* 1995), mouse, *Mus musculus* (Bibb *et al.* 1981), rat, *Rattus norvegicus* (Gadaleta *et al.* 1989), rabbit, *Oryctolagus cuniculus* (Gissi *et al.* 1998), guinea-pig, *Cavia porcellus* (D'Erchia *et al.* 1996), armadillo, *D. novemcinctus* (Arnason *et al.* 1997), cat, *Felis catus* (Lopez *et al.* 1996), harbour seal, *Phoca vitulina* (Arnason & Johnsson 1992), horse, *Equus caballus* (Xu & Arnason 1994), white rhinoceros, *Ceratotherium simum* (Xu & Arnason 1997), cow, *Bos taurus* (Anderson *et al.* 1982), fin whale, *Balaenoptera physalus* (Arnason *et al.* 1991), blue whale, *Balaenoptera musculus* (Arnason & Gullberg 1993), gibbon, *Hylobates lar* (Arnason *et al.* 1996*b*) and human, *Homo sapiens* (Arnason *et al.* 1996*c*). Thus, whenever available, two mtDNAs of each eutherian order were included in the analyses. Even though a greater number of complete mtDNAs are available in some instances, the present selection of two complete mitochondrial sequences of each order was sufficient for maintaining constancy in phylogenetic reconstruction.

The phylogenetic analyses were primarily based on amino acid (aa) sequences. The analyses were performed with maximum likelihood (ML) (Felsenstein 1981), maximum parsimony (MP) (Fitch 1971) and neighbour-joining (NJ) (Saitou & Nei 1987). The ML analyses were implemented by the PUZZLE program v. 4.0 for establishing Q/P (quartet/puzzle) support values (Strimmer & von Haeseler 1996) and the MOLPHY program package v. 2.3 (Adachi & Hasegawa 1996*a*). The MP and NJ analyses were performed with the PHYLIP package (Felsenstein 1993). ML analysis of aa sequences as well as distance calculations adopted the mtREV-24 rate matrix of aa sequence evolution (Adachi & Hasegawa 1996*b*). Furthermore, a uniform rate of substitution of sequence evolution and the observed aa frequencies of the data set was assumed. Nucleotide sequences were analysed by applying the HKY model (Hasegawa *et al.* 1985), but otherwise under the same conditions as described for aa sequences.

## 3. RESULTS

### (a) *Features of the genome*

Tandemly organized repeats occurring in variable numbers (heteroplasmy) have been shown to characterize the control regions of many eutherian mtDNAs which have been sequenced in their entirety. For this reason the lengths of complete mtDNAs are in many cases not absolute. The length of the presently reported mtDNA of the aardvark is 16 816 nucleotides (nt). This length includes 33 repeated motifs (CGCATA). In the two other clones sequenced the number of repeats were 10 and 20, respectively.

The organization of the mtDNA of the aardvark and the location of individual genes conform with that of other eutherian mtDNAs which have been sequenced in their entirety. Two genes, *NADH1* and *NADH3*, do not have a canonical methionine start codon. The start codon of *NADH1* is GTG, while that of *NADH3* is ATC. In mammals GTG has been described as a start codon in *NADH4L* in the blue whale (Arnason & Gullberg 1993); GTG is also the start codon of the *NADH1* gene of the elephant shrew (X. Xu, A. Janke and U. Arnason, unpublished data), the rat and the mouse. ATT or ATC have been reported as start codons in the *NADH3* gene of several mammals, the mouse, rat, Sumatran orang-utan and Indian and white rhinoceroses. The occurrence of ATT/ATC as a start codon is consistent with the notion (Fearnley & Walker 1987) that any ATN start codon may specify methionine. COIII, *NADH3* and *NADH4* have an incomplete stop codon (T). The stop codon of *ATPase6* is also incomplete (TA). The 3′ end nucleotide of these genes is contiguous with the 5′ terminal of a *tRNA* gene and it has been proposed that the transcripts of such protein-coding mtDNA genes contain a stop codon created by post-transcriptional polyadenylation (Ojala *et al.* 1981). It is also probable that *NADH2* has an incomplete stop codon (T) rather than a complete stop codon (TAG). If this is so, the nucleotides AG constitute a part

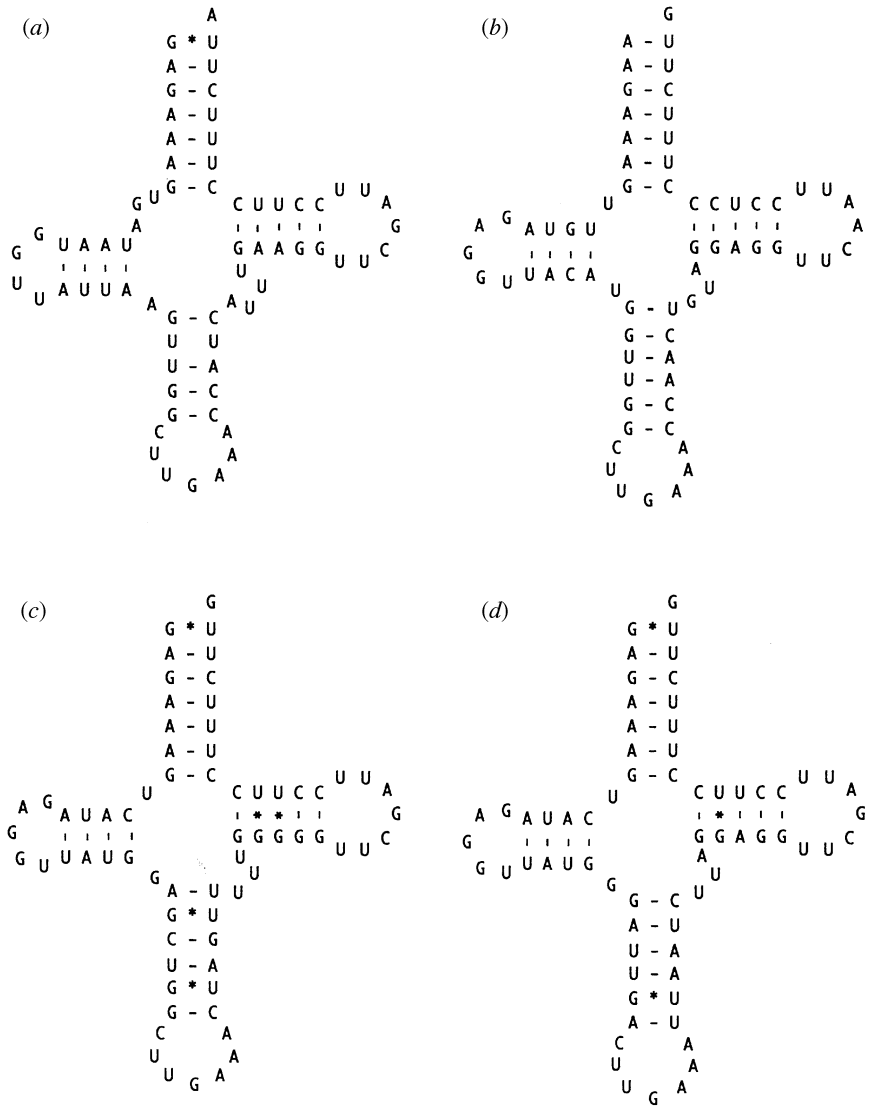


Figure 1. Postulated secondary structure of the *tRNA-Ser*(*UCN*) of the aardvark compared to that of the elephant shrew, armadillo and wallaroo.

of the adjacent *tRNA-Trp* gene, an arrangement that would be consistent with the sequence of *tRNA-Trp* in other eutherians.

The *tRNA-Ser*(*UCN*) of the aardvark differs from that of other mammals studied so far. Figure 1 shows the potential secondary structures of the *tRNA-Ser*(*UCN*) of the aardvark, two other eutherians, the elephant shrew and armadillo and one marsupial, the wallaroo. The latter three species have the typical mammalian secondary structure for this particular tRNA. The optimal stem alignment suggests that the acceptor and DHU (dihydrouridine) stems of the aardvark are connected by three nucleotides as compared with one nucleotide in other eutherians and that the anticodon stem of the aardvark is five base pairs long as compared with six in other eutherians. The *tRNA-Ser*(*UCN*) of the aardvark is shown with maximal complementarity in the acceptor and DHU stems, resulting in three nucleotides connecting the acceptor and DHU stems, as compared with one nucleotide in this particular tRNA in other mammals studied so far. An alternative secondary structure with two connecting nucleotides would result in two non-complementary base pairs in the acceptor stem.

It has been proposed that the common mammalian structure of *tRNA-Ser*(*UCN*), characterized by one nucleotide in the connecting region, has served to compensate for a larger anticodon/a distance, when the anticodon stem became extended from five to six nucleotide pairs (Steinberg & Cedergren 1994). The *tRNA-Ser*(*UCN*) of the aardvark suggests a reversal to the ancestral secondary structure of the *tRNA-Ser*(*UCN*) such as occurs in non-mammalian vertebrates and in mitochondrial tRNAs in general. The findings contradict the hypothesis that reversals are highly improbable in mitochondrial tRNAs (Lynch 1996; Börner *et al.* 1997).

### (b) *Phylogeny*

The phylogenetic analyses were primarily based on the concatenated aa sequences of the 12 H-strand-encoded protein-coding genes, excluding the L-strand-encoded *NADH6* gene due to its deviating nucleotide and aa composition. After excluding gaps and ambiguous sites adjacent to gaps, 9975 nt positions (3325 aa) remained for analysis. The platypus and two marsupials, the opossum and wallaroo, were used as outgroups to root the eutherian tree. The aa composition of all species included in the data set conform, as tested by a 5% level $\chi^2$ test,

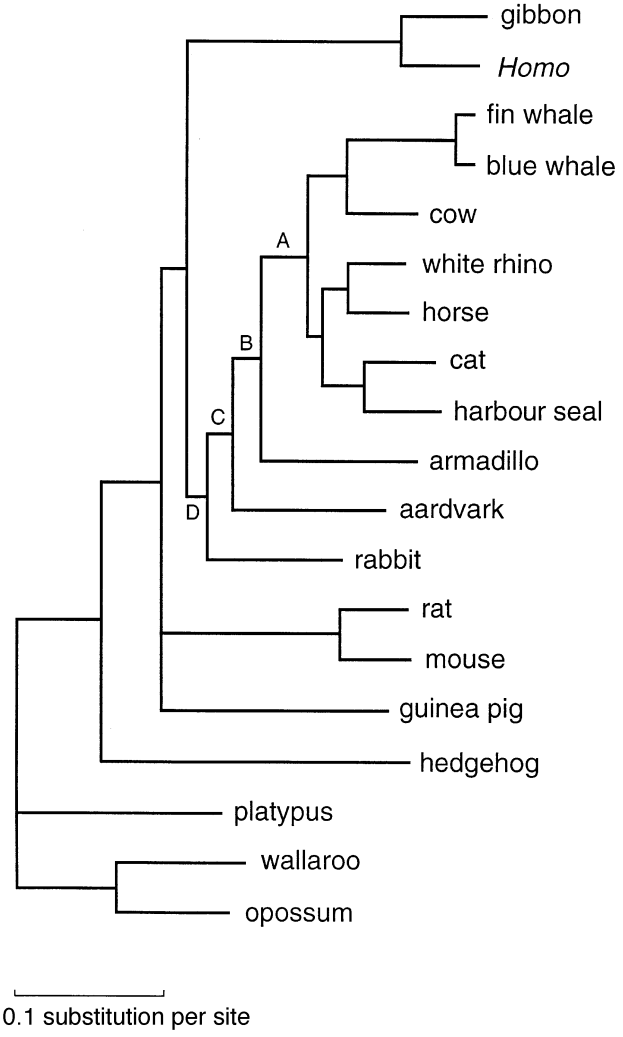


Figure 2. ML tree of concatenated aa sequences of 12 H-strand-encoded mitochondrial genes of 19 mammals. For scientific names see § 2. The analysis identified the aardvark as a sister group to the armadillo and cetferungulates. Branch lengths are according to genetic distance. Support values for branches A–D are given in table 1.

with the aa frequency used by the ML model. The hedgehog had the lowest probability ($p = 9.89\%$), followed by the gibbon ($p = 50.59\%$) and *Homo* ($p = 65.35\%$). The probabilities for all other species were well above $p = 90\%$. The inclusion or exclusion of the hedgehog did not alter the topology of the eutherian tree (figure 2) and it was therefore maintained in the data set, even though its aa composition deviated somewhat from that of the other species.

The ML analysis was performed with the mtREV-24 model of aa evolution (Adachi & Hasegawa 1996*b*). The positions of the guinea-pig and rabbit were unstable in the tree. Exclusion of these two species reconstructed the same relationship in the three analytical approaches applied (ML–QP, NJ and MP), with the aardvark basal to the armadillo and the cetferungulates: (aardvark (armadillo, cetferungulates)). The bootstrap and QP support values for this relationship were *ca.* 90%. ML–QP and NJ analyses including the guinea-pig and the rabbit identified the rabbit as the sister group of

Table 1. *Support values and branch lengths for selected branches*

| branch | NJ | QP | ML length | s.e. |
|---|---|---|---|---|
| A | 100 | 95 | 0.03142 | 0.0038 |
| B | 54 | 74 | 0.01929 | 0.0031 |
| C | 74 | 69 | 0.01741 | 0.0031 |
| D | 89 | 59 | 0.01430 | 0.0031 |

NJ bootstrap and QP support values and ML branch lengths and standard errors were calculated according to the mtREV-24 model of aa evolution (Adachi & Hasegawa 1996*b*).

(aardvark (armadillo, cetferungulates)), while MP analysis of the same data set reconstructed two equally parsimonious solutions: (aardvark, (rabbit, (armadillo, cetferungulates))) and (aardvark, ((rabbit, armadillo), cetferungulates)). In the ML–QP analysis the guinea-pig and the myomorph rodents were on a common branch and in the NJ analysis the guinea-pig was basal to the divergence between primates and (rabbit, (aardvark, (armadillo, cetferungulates))), while the MP analysis identified the guinea-pig as the sister group of the primates. The ML–QP tree in figure 2 depicts the (rabbit, (aardvark, (armadillo, cetferungulates))) relationship identified in the ML–QP and NJ analyses. The position of the guinea-pig has been left unresolved since the three analytical approaches yielded different positions for that taxon. Table 1 shows the bootstrap support for NJ and the QP support values and ML branch lengths with their standard errors for selected branches related to the phylogenetic position of the aardvark in the eutherian tree.

The position of the aardvark was analysed further by calculating the ML values for different positions of the aardvark in the eutherian tree (table 2). The two most probable topologies are the aardvark as the sister group to the armadillo–cetferungulate clade and the aardvark as the sister group of the armadillo outside the cetferungulates. Relative to these two topologies other topologies were highly improbable.

The phylogenetic analysis of nucleotide sequences was performed on non-synonymous changes at the first codon position and all changes at the second codon position (all non-synonymous). Analysis of the combined data set of both codon positions yielded the same results as the aa analysis, even though analysis of either codon position did not conclusively resolve the position of the aardvark. NJ analysis of distances at the second position reconstructed the same tree as the aa analysis, but MP identified four equally parsimonious trees, leaving the position of the aardvark unresolved.

The origin of the African clade, as represented by the tubulidentate lineage, was calculated by applying three molecular reference points: A/C-60 (artiodactyl/cetacean divergence 60 Myr BP), E/R-50 (equid/rhinocerotid divergence 50 Myr BP) and M/E-130 (marsupial/eutherian divergence 130 Myr BP). After calibration for differences in evolutionary rates, using the mt REV-24 model of aa evolution, ML distances and branch lengths yielded similar datings (88–94 Myr BP) for the age of the tubulidentate lineage. Rate calibrations and estimates of

Table 2. *ML analysis of different positions of aardvark (TUB) in the eutherian tree*

(The ML analysis was based on the mtREV-24 model of sequence evolution. The value in angle brackets shows the log-likelihood (lnL) value of the best tree. ΔlnL indicates the difference in lnL to that of the best tree, followed by the standard error (s.e.) (Kishino & Hasegawa 1989), and bootstrap probability (pboot) for this particular topology (Kishino *et al* 1990). MON, monotremes (platypus); MAR, marsupials (opossum and wallaroo); LIP, lipotyphlans (hedgehog); ROD, myomorph rodents (rat and mouse); CAV, caviomorph rodents (guinea-pig); PRIM, primates (*Homo* and gibbon), LAG, lagomorphs (rabbit); TUB, tubulidentates (aardvark); XEN, xenarthrans (armadillo); FER, ferungulates and cetaceans (cow, white rhinoceros, horse, harbour seal, domestic cat, fin whale and blue whale).)

| topology | ΔlnL | s.e. | pboot |
|---|---|---|---|
| (MON,MAR,(LIP,(ROD,CAV,(PRIM,(LAG,(**TUB**,(XEN,FER))))))) | | < −45259.8 | > 0.4820 |
| (MON,MAR,(LIP,(ROD,CAV,(PRIM,(LAG,(XEN,(**TUB**,FER))))))) | −8.3 | 23.5 | 0.2800 |
| (MON,MAR,(LIP,(ROD,CAV,(PRIM,(LAG,((**TUB**,XEN),FER)))))) | −19.9 | 21.9 | 0.1060 |
| (MON,MAR,(LIP,(ROD,CAV,(PRIM,(**TUB**,(LAG,(XEN,FER))))))) | −25.3 | 18.2 | 0.0220 |
| (MON,MAR,(LIP,(ROD,CAV,((**TUB**,PRIM),(LAG,(XEN,FER)))))) | −38.1 | 26.0 | 0.0220 |
| (MON,MAR,(LIP,(ROD,CAV,(**TUB**,(PRIM,(LAG,(XEN,FER))))))) | −31.9 | 27.4 | 0.0380 |
| (MON,MAR,(LIP,(**TUB**,(ROD,CAV,(PRIM,(LAG,(XEN,FER))))))) | −52.2 | 40.2 | 0.0460 |
| (MON,MAR,(**TUB**,(LIP,(ROD,CAV,(PRIM,(LAG,(XEN,FER))))))) | −127.4 | 47.3 | 0.0000 |

various other divergence times have been described previously (Arnason *et al*. 1996*a*, 1998) and are therefore not detailed here.

## 4. DISCUSSION

For the first time, the present analyses have allowed identification of the phylogenetic position of the African clade, as represented by the aardvark, within reasonably narrow limits. The analyses have also suggested that the temporal origin of the African clade is close to that of several other eutherian lineages, making full phylogenetic resolution difficult even in analyses of complete mtDNA molecules. The analyses underlined the somewhat unstable position of the the guinea-pig and rabbit in the eutherian tree. When the rabbit (Lagomorpha) was excluded, all phylogenetic analyses placed the aardvark as a sister group of the armadillo and cetferungulates. After inclusion of the rabbit the same relationship was supported in ML–QP and NJ analyses, while MP reconstructed two other arrangements, one in which the position of the aardvark and the rabbit was switched and one with the rabbit and the aardvark on a common branch as a sister group of the armadillo–cetferungulates. The position of the guinea-pig was identified differently by the three analytical approaches applied. This suggests that it is unlikely that the position of the guinea-pig will become stabilized in the eutherian tree until additional hystricomorph taxa become available. It is also conceivable that the phylogenetic position of the African clade will become better defined as more mtDNAs of that clade are described.

Previous studies of the phylogenetic position of the Tubulidentata and the African clade relative to other eutherian orders have not provided conclusive answers (Stanhope *et al*. 1996; Madsen *et al*. 1997; Springer *et al*. 1997*a*,*b*). Earlier analyses of nuclear sequences commonly placed the Tubulidentata basal among higher Eutheria, the so-called Epiteria, while other similar studies have reported unresolved relationships among 13 eutherian clades including the African clade. Some of the conclusions on the phylogenetic position of the aardvark have been based on the assumption that xenarthrans (edentates) have a basal position in the eutherian tree. However, more recent studies have shown that the xenarthrans, as represented by the armadillo, are the sister group of the cetferungulates, thereby rendering the xenarthrans (edentates) unsuitable as an outgroup to root the eutherian ordinal tree and, at the same time, invalidating the Epiteria as a systematic designation (Arnason *et al*. 1997).

While some interpretations of the position of the aardvark have been affected by the rooting applied, it is probable that the limited resolution reported for some deep divergences in the nuclear data sets is related to the sizes of the alignments used, making the analyses vulnerable to stochastic effects (Cao *et al*. 1994). In this context, it is nevertheless noteworthy that the African clade has received high statistical support with detailed resolution in mitochondrial as well as nuclear data sets (Lavergne *et al*. 1996; Stanhope *et al*. 1996; Madsen *et al*. 1997; Springer *et al*. 1997*a*,*b*). The disparity between the philosophies of establishing phylogenetic relationships on the basis of extensive sampling of alignments of limited length relative to intensive sampling of fewer taxa has become evident in two recent contributions (Penny & Hasegawa 1997; Springer *et al*. 1997*b*). In the latter contribution the use of long alignments (such as these obtained from complete mtDNAs) was strongly recommended relative to the use of shorter sequences. It has been estimated that some 3000 nt positions of mitochondrial sequences are the minimum number for resolving the phylogeny among seven mammalian taxa; opossum, mouse, rat, human, harbour seal, cow and fin whale (von Haeseler *et al*. 1993). For resolving distant mammalian divergences that have taken place within limited periods of time, this number of nucleotide positions may even be too low. It remains to be seen what are the minimum sizes of nuclear data sets for providing resolution comparable to that of complete mtDNAs, but considering the resolution provided by several nuclear data sets it is probable that the lengths of the alignments will have to exceed the length of comparable mtDNA alignments.

Resolved topologies including unequivocal outgroups for relative rate tests are essential for all molecular datings. Based on a recently established molecular calibration

point, the divergence between ruminant artiodactyls (cow) and cetaceans (A/C-60) (Arnason & Gullberg 1996), the divergence between armadillo and cetferungulates has been set at 86 Myr BP and that between primates and the armadillo–cetferungulate lineage at *ca*. 95 Myr BP (Arnason *et al*. 1997; Janke *et al*. 1997). The timings are supported by a more recently established calibration point, the intraordinal perissodactyl divergence between Equidae and Rhinocerotidae set at 50 Myr BP (E/R-50) (Xu *et al*. 1996; Arnason *et al*. 1998). Application of these two references to the most strongly supported phylogenetic tree places the divergence between Tubulidentata and the armadillo–cetferungulate clade at *ca*. 90 Myr BP. The divergence time between the African clade and 13 mammalian orders has been estimated at 91 Myr BP (Springer *et al*. 1997*a*). The calibration points applied in that estimation included a *Mus–Rattus* divergence set at 14 Myr BP. That dating is drastically different from datings according to A/C-60, E/R-50 and M/E-130 which place the mouse–rat split at 35–40 Myr BP. The mitochondrial dating (35–40 Myr BP) is in agreement with previous nuclear findings (Wilson *et al*. 1977; Holmes 1991), but remains unsupported by the fossil record (see Janke *et al*. (1994) and references therein). Based on mitochondrial and most nuclear data, a rat–mouse divergence of 10–14 Myr BP would be inconsistent with eutherian origin at *ca*. 130 Myr BP and application of a rat–mouse divergence of 10–14 Myr BP as a calibration point would result in a pronounced underestimate of the time of other divergences. In addition, some of the rate-adjustments made by Springer *et al*. (1997*b*) assumed a basal position of xenarthrans.

The present analyses have placed the aardvark, as a representative of the proposed 'African clade', on a branch including the cetferungulates and some other taxa. The findings are consistent with previous results which found that the major radiation of recent eutherian orders took place within a relatively short span of time, in the Albian and Cenomanian periods, between 110 and 90 Myr BP (Janke *et al*. 1994, 1997; Arnason *et al*. 1997, 1998). During this time span many still existing orders apparently diverged at short intervals from the lineage ultimately resulting in the Cetferungulata. For this reason it is likely that the inclusion of each additional lineage will make it increasingly difficult to resolve all ordinal divergences arising in the mid-/late Cretaceous period, even by means of complete mtDNA genomes or nuclear data sets of comparable sizes.

We express our gratitude to Dr J. Wensing (Burgers Zoo, Arnheim) and Professor W. W. de Jong (Department of Biochemistry, Katholieke Universiteit, Nijmegen, The Netherlands) for tissue samples and to Suzette Mouchaty for valuable comments on the manuscript. The work was supported by the Swedish Natural Sciences Research Council, the Carl Trygger Foundation, the Nilsson-Ehle Foundation and by contract ERBCHR XCT 930254 from the European Commission.

## REFERENCES

Adachi, J. & Hasegawa, M. 1996*a* *MOLPHY version 2.3: programs for molecular phylogenetics based on maximum likelihood*. Tokyo: The Institute of Statistical Mathematics.

Adachi, J. & Hasegawa, M. 1996*b* Model of amino acid substitution in proteins encoded by mitochondrial DNA. *J. Mol. Evol.* **42**, 459–468.

Anderson, S., de Brujin, M. H. L., Coulson, A. R., Eperon, I. C., Sanger, F. & Young, G. 1982 Complete sequence of the bovine mitochondrial DNA. Conserved features of the mammalian mitochondrial genome. *J. Mol. Biol.* **156**, 683–717.

Arnason, U. & Gullberg, A. 1993 Comparison between the complete mtDNA sequences of the blue and the fin whale, two species that can hybridize in nature. *J. Mol. Evol.* **37**, 312–322.

Arnason, U. & Gullberg, A. 1996 Cytochrome *b* nucleotide sequences and the identification of five primary lineages of extant cetaceans. *Mol. Biol. Evol.* **13**, 407–417.

Arnason, U. & Johnsson, E. 1992 The complete mitochondrial DNA sequence of the harbor seal, *Phoca vitulina*. *J. Mol. Evol.* **34**, 493–505.

Arnason, U., Gullberg, A. & Widegren, B. 1991 The complete nucleotide sequence of the mitochondrial DNA of the fin whale, *Baleonoptera physalus*. *J. Mol. Evol.* **33**, 556–568.

Arnason, U., Gullberg, A., Janke, A. & Xu, X. 1996*a* Pattern and timing of evolutionary divergencies among hominoids based on analysis of complete mtDNAs. *J. Mol. Evol.* **43**, 650–661.

Arnason, U., Gullberg, A. & Xu, X. 1996*b* A complete mitochondrial DNA molecule of the white-handed gibbon, *Hylobates lar*, and comparison among individual mitochondrial genes of all hominoid genera. *Hereditas* **124**, 185–189.

Arnason, U., Xu, X. & Gullberg, A. 1996*c* Comparison between the complete mitochondrial DNA sequences of *Homo* and the common chimpanzee. *J. Mol. Evol.* **42**, 145–152.

Arnason, U., Gullberg, A. & Janke, A. 1997 Phylogenetic analyses of mitchondrial DNA suggest a sister group relationship between Xenarthra (Edentata) and Ferungulates. *Mol. Biol. Evol.* **14**, 762–768.

Arnason, U., Gullberg, A. & Janke, A. 1998 Molecular timing of primate divergences as estimated from two non-primate calibration points. *J. Mol. Evol.* **47**, 718–727.

Bibb, M. J., Van Etten, R. A., Wright, C. T., Walberg, M. W. & Clayton, D. A. 1981 Sequence and gene organization of the mouse mitochondrial DNA. *Cell* **26**, 167–180.

Börner, G. V., Yokobori, S. I., Mörl, M., Dörner, M. & Pääbo, S. 1997 RNA editing in metazoan mitochondria: staying fit without sex. *FEBS Lett.* **409**, 320–324.

Cao, Y., Adachi, J., Janke, A., Pääbo, S. & Hasegawa, M. 1994 Phylogenetic relationships among eutherian orders estimated from inferred protein sequences of mitochondrial proteins: instability of a tree based on a single gene. *J. Mol. Evol.* **39**, 519–527.

Carroll, R. L. 1988 *Vertebrate paleontolgy and evolution*. New York: W. H. Freeman.

de Jong, W. W., Zweers, A. & Goodman, M. 1981 Relationship of aardvark to elephants, hyraxes and sea cows from a-crystallin sequences. *Nature* **292**, 538–540.

de Jong, W. W., Leunissen, J. A. M. & Wistow, G. J. 1993 Eye lens crystallins and the phylogeny of placental orders: evidence for a Marcoscelid–Paenungulate clade? In *Mammal phylogeny, placentals* (ed. F. S. Szalay, M. J. Novacek & M. C. McKenna), pp. 5–12. Berlin: Springer.

D'Erchia, A. M., Gissi, C., Pesole, G., Saccone, C. & Arnason, U. 1996 The guinea-pig is not a rodent. *Nature* **381**, 597–599.

Fearnley, I. M. & Walker, J. E. 1987 Initiation codons in mammalian mitochondria: differences in genetic code in the organelle. *Biochemistry* **26**, 8247–8251.

Felsenstein, J. 1981 Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.* **17**, 368–376.

Felsenstein, J. 1993 PHYLIP (phylogeny inference package) version 3.5p. Distributed by the author. Department of Genetics, University of Washington, Seattle.

Fitch, W. M. 1971 Toward defining the course of evolution: minimum change for a specific tree topology. *Syst. Zool.* **20**, 406–416.

Gadaleta, G., Pepe, G., De Candia, G., Quagliariello, C., Sibisa, E. & Saccone, C. 1989 The complete nucleotide sequence of the *Rattus norvegicus* mitochondrial genome: cryptic signals revealed by comparative analysis between vertebrates. *J. Mol. Evol.* **28**, 497–516.

Gissi, C., Gullberg, A. & Arnason, U. 1998 The complete mitochondrial DNA sequence of the rabbit, *Oryctolagus cuniculus*. *Genomics* **50**, 161–169.

Graur, D. 1993 Molecular phylogeny and higher classification of eutherian mammals. *Trends Ecol. Evol.* **8**, 141–147.

Hasegawa, M., Kishino, H. & Yano, T. 1985 Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22**, 160–174.

Holmes, E. C. 1991 Different rates of substitution may produce different phylogenies of the eutherian mammals. *J. Mol. Evol.* **33**, 209–215.

Janke, A., Feldmaier-Fuchs, G., Thomas, W. K., von Haeseler, A. & Pääbo, S. 1994 The marsupial mitochondrial genome and the evolution of placental mammals. *Genetics* **137**, 243–256.

Janke, A., Gemmell, N. J., Feldmaier-Fuchs, G., von Haeseler, A. & Pääbo, S. 1996 The mitochondrial genome of a monotreme—the platypus (*Ornithorhynchus anatinus*). *J. Mol. Evol.* **42**, 153–159.

Janke, A., Xu, X. & Arnason, U. 1997 The complete mitochondrial genome of the wallaroo (*Macropus robustus*) and the phylogentic relationship among Monotremata, Marsupialia and Eutheria. *Proc. Natl Acad. Sci. USA* **94**, 1276–1281.

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

Kishino, H., Miyata, T. & Hasegawa, M. 1990 Maximum likelihood inference of protein phylogeny and the origin of chloroplasts. *J. Mol. Evol.* **31**, 153–160.

Krettek, A., Gullberg, A. & Arnason, U. 1995 Sequence analysis of the complete mitochondrial DNA molecule of the hedgehog, *Erinaceus europaeus,* and the phylogenetic position of the Lipotyphla. *J. Mol. Evol.* **41**, 952–957.

Lavergne, A., Douzery, E., Stichler, T., Catzeflis, F. & Springer, M. S. 1996 Interordinal mammalian relationship: evidence for paenungulate monophyly is provided by complete 12S rRNA sequences. *Mol. Phylogenet. Evol.* **6**, 245–258.

Le Gros Clark, W. E. & Sonntag, C. F. 1926 A monograph of *Orycterpus afer*. *Proc. Zool. Soc.* (*Lond.*) **30**, 445–485.

Lopez, J. V., Cevario, S. & O'Brien, S. J. 1996 Complete nucleotide sequences of the domestic cat (*Felis catus*) mitochondrial genome and a transposed mtDNA repeat, Numt, in the nuclear genome. *Genomics* **33**, 229–246.

Lynch, M. 1996 Mutation accumulation in transfer RNAs: molecular evidence for Muller's ratchet in mitochondrial genomes. *Mol. Biol. Evol.* **13**, 209–220.

Madsen, O., Deen, P. M. T., Pesole, G., Saccone, C. & de Jong, W. W. 1997 Molecular evolution of mammalian aquaporin-2: further evidence that elephant shrew and aardvark join the paenungulate clade *Mol. Biol. Evol.* **14**, 363–371.

Miyamoto, M. M. & Goodman, M. 1986 Biomolecular systematics of eutherian mammals: phylogenetic patterns and classification. *Syst. Zool.* **35**, 230–240.

Ojala, D., Montoya, J. & Attardi, G. 1981 tRNA punctuation model of RNA processing in human mitochondria. *Nature* **290**, 470–474.

Penny, D. & Hasegawa, M. 1997 Molecular systematics: the platypus put in its place. *Nature* **387**, 549–550.

Rainey, W. E., Lowenstein, J. M., Sarich, V. M. & Magor, D. M. 1984 Sirenian molecular systematics—including the extinct Steller's sea cow (*Hydrodamalis gigas*). *Naturwissenschaften* **71**, 586–588.

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

Sanger, F. 1981 Determination of nucleotide sequences in DNA. *Science* **214**, 1205–1210.

Springer, M. S., Burk, A., Kavanagh, J. R., Waddell, V. G. & Stanhope, M. J. 1997*a* The interphotoreceptor retinoid binding protein gene in therian mammals: implications for higher level relationships and evidence for loss of function in the marsupial mole. *Proc. Natl Acad. Sci. USA* **94**, 13 754–13 759.

Springer, M. S., Cleven, G. C., Madsen, O., de Jong, W. W., Waddell, V. G., Amrine, H. M. & Stanhope, M. J. 1997*b* Endemic African mammals shake the phylogenetic tree. *Nature* **388**, 61–64.

Stanhope, M. J., Smith, M. R., Waddell, V. G., Porter, C. A., Shivji, M. S. & Goodman, M. 1996 Mammalian evolution and the IRBP gene: convincing evidence for several superordinal clades. *J. Mol. Evol.* **43**, 83–92.

Steinberg, S. & Cedergren, R. 1994 Structural compensation in atypical mitochondrial tRNAs. *Struct. Biol.* **1**, 507–510.

Strimmer, K. & von Haeseler, A. 1996 Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.* **13**, 964–969.

von Haeseler, A., Janke, A. & Pääbo, S. 1993 Molecular phylogenetics. *Verh. Dtsch. Zool. Ges.* **86**, 119–129.

Wilson, A. C., Carlson, S. S. & White, T. J. 1977 Biochemical evolution. *A. Rev. Biochem.* **46**, 573–639.

Xu, X. & Arnason, U. 1994 The complete mitochondrial DNA sequence of the horse, *Equus caballus*; extensive heteroplasmy of the control region. *Gene* **148**, 357–362.

Xu, X. & Arnason, U. 1997 The complete mitochondrial DNA sequence of the white rhinoceros *Ceratotherium simum*, and comparision with the mtDNA of the Indian rhinoceros, *Rhinoceros unicornis*. *Mol. Phylogenet. Evol.* **7**, 189–194.

Xu, X., Janke, A. & Arnason, U. 1996 The complete mitochondrial DNA sequence of the greater Indian rhinoceros, *Rhinoceros unicornis*, and the phylogenetic relationship among Carnivora, Perissodactyla and Artiodactyla (+ Cetacea). *Mol. Biol. Evol.* **13**, 1167–1173.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.