# The mitochondrial genomes of the iguana (Iguana iguana) and the caiman (Caiman crocodylus): implications for amniote phylogeny

Janke, Axel; Erpenbeck, Dirk; Nilsson, Malin; Arnason, Ulfur

*Published in:*
Royal Society of London. Proceedings B. Biological Sciences

*DOI:*
10.1098/rspb.2000.1402

2001

THE ROYAL SOCIETY

*doi* 10.1098/rspb.2000.1402

# The mitochondrial genomes of the iguana (*Iguana iguana*) and the caiman (*Caiman crocodylus*): implications for amniote phylogeny

**Axel Janke*, Dirk Erpenbeck†, Malin Nilsson and Ulfur Arnason**

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

The complete mitochondrial genomes of two reptiles, the common iguana (*Iguana iguana*) and the caiman (*Caiman crocodylus*), were sequenced in order to investigate phylogenetic questions of tetrapod evolution. The addition of the two species allows analysis of reptilian relationships using data sets other than those including only fast-evolving species. The crocodilian mitochondrial genomes seem to have evolved generally at a higher rate than those of other vertebrates. Phylogenetic analyses of 2889 amino-acid sites from 35 mitochondrial genomes supported the bird–crocodile relationship, lending no support to the Haematotherma hypothesis (with birds and mammals representing sister groups). The analyses corroborated the view that turtles are at the base of the bird–crocodile branch. This position of the turtles makes Diapsida paraphyletic. The origin of the squamates was estimated at 294 million years (Myr) ago and that of the turtles at 278 Myr ago. Phylogenetic analysis of mammalian relationships using the additional outgroups corroborated the Marsupionta hypothesis, which joins the monotremes and the marsupials to the exclusion of the eutherians.

**Keywords:** amniote evolution; mitochondrial DNA; reptiles; caiman; iguana

## 1. INTRODUCTION

Ever since the first attempts to understand the evolution of the vertebrates were made, relationships between tetrapod classes have attracted considerable attention. On the basis of the presence or absence of specific temporal openings in the skull, the Tetrapoda are divided into three major groups: the Anapsida (today represented only by the turtles), the Diapsida (which include crocodiles, birds and squamates, i.e. lizards and snakes) and the Synapsida (mammals). Although the anapsids, diapsids and synapsids are easily distinguishable by their morphology, the relationship between them remains problematical.

The monophyly of the Diapsida has been questioned (Gardiner 1982, 1993) by revival of the Haematotherma hypothesis which suggests that birds and mammals are sister groups. (We will use the term Haematotherma as Owen (1866) originally coined it. It was reintroduced erroneously as Haemothermia (Gardiner 1982) and then miscorrected to Haematothermia by Eernisse & Kluge (1993).) A strict interpretation of the Haematotherma hypothesis has been rejected on the basis of both fossil data (Gauthier *et al.* 1988) and recent analyses of complete mitochondrial genomes (Janke & Arnason 1997). In some cases, a somewhat modified Haematotherma hypothesis has been supported, according to which mammals, along with birds and crocodiles, are monophyletic. However, this hypothesis has so far not been tested using comprehensive molecular data. A few analyses of single genes have supported a sister-group relationship between birds and mammals (Dene *et al.* 1982; Hedges *et al.* 1990), whereas in other cases they have not (Perutz *et al.* 1981; de Jong *et al.* 1985; Kumazawa & Nishida 1993; Hedges 1994; Janke & Arnason 1997).

The phylogenetic position of the turtles has been a matter of debate in recent years. Traditionally, they have been placed at the base of extant amniotes or, more recently, at the base of the Diapsida. However, molecular studies have placed the turtles in different positions: (i) with the squamates (Hedges 1994), (ii) with the crocodiles (Hedges *et al.* 1990; Kirsch & Mayer 1998; Hedges & Poling 1999), (iii) basal to birds plus crocodiles (Zardoya & Meyer 1998; Kumazawa & Nishida 1999), or (iv) basal to diapsids (Caspers *et al.* 1996). The majority of molecular phylogenetic studies, although otherwise inconsistent in their conclusions, have indicated that the turtles are not a basal lineage among the anapsids which they have traditionally been grouped together with. Molecular studies as well as recent morphological studies have instead grouped the turtles together with the diapsids (Rieppel & DeBraga 1996; Rieppel 1999), thus challenging the monophyly of the Diapsida.

Previous analyses of complete mitochondrial genomes have reconstructed an unconventional relationship between the three main mammalian groups (monotremes, marsupials and eutherians) by joining the monotremes and the marsupials, to the exclusion of the eutherians (Janke *et al.* 1996, 1997). These molecular findings were consistent with the 'Marsupionta' hypothesis (Kühne 1973, 1975). The Marsupionta relationship received further support when tested with the use of various other outgroups (Janke & Arnason 1997; Härlid *et al.* 1998; Kumazawa *et al.* 1998; Zardoya & Meyer 1998). However, support for this relationship was weakened somewhat by inclusion of the alligator (Janke & Arnason 1997). It is possible that the fast evolutionary rate of the alligator (Janke & Arnason 1997) increased the homoplasy in the analysis, thereby obscuring the phylogenetic signal. Additional sequence data could help

*Author for correspondence (axel.janke@gen.lu.se).
†Present address: Department of Coelenterata and Porifera, Institute for Systematics and Ecology (Zoölogisch Museum), University of Amsterdam, PO Box 94766, 1090 GT Amsterdam, The Netherlands.


*Received* 10 July 2000 *Accepted* 15 November 2000

to overcome this obstruction and improve the resolution as improved taxon sampling might allow recognition of heteroplasmic sites in the phylogenetic analysis.

The complete mitochondrial genome sequences (mitochondrial DNAs (mtDNAs)) of some 50 vertebrate species are currently available in databases. However, only a few of the published sequences represent the reptiles, namely those of an alligator (Janke & Arnason 1997), a snake, a mole skink (Kumazawa *et al.* 1998) and three turtles (Zardoya & Meyer 1998; Kumazawa & Nishida 1999; Mindell *et al.* 1999). The fast evolutionary rate observed in some species (the snake, the side-necked turtle and the alligator) complicates the phylogenetic analysis. Furthermore, numerous insertion/deletion events in the mitochondrial genomes of the snake and the turtle along with disrupted reading frames in some of the protein-coding genes of the side-necked turtle drastically reduce the size of any alignment in which these species are included, thus reducing the resolution and stringency of the phylogenetic analysis. In order to analyse the amniote phylogeny in further detail we have increased the number of taxa by sequencing the complete mtDNA sequences of a lizard (*Iguana iguana*) and a crocodile (*Caiman crocodylus*).

## 2. MATERIAL AND METHODS

Mitochondrial DNA was isolated from frozen liver and heart tissue as described by Arnason *et al.* (1991). The tissue from the iguana was a generous gift from Dr Christine Lendel of the Hellabrunn Zoo, Munich, Germany. The caiman tissue was collected from a deceased pet shop animal imported in accordance with the Convention on International Trade in Endangered Species permit 3006/04189. The mtDNA was cut with different restriction endonucleases, ligated into M13mp18/19 vectors and sequenced using standard procedures (Sambrook *et al.* 1989). The sequencing of the iguana mitochondrial genome was performed manually (Sambrook *et al.* 1989) on single-stranded DNA using the dideoxy termination technique (Sanger 1981) with $^{35}$S-$\alpha$-dATP. The caiman mitochondrial genome was sequenced by cycle sequencing (fluorescent-labelled primer cycle sequencing kit deaza-dGTP, Thermo Sequinase, Amersham, Solna, Sweden) with IRD41/800-labelled primers on an LiCor 4000L (AH Diagnostics, Skärholmen, Sweden or MWG, Munich, Germany). Both universal and numerous specific primers were used in the sequencing process. Regions not covered by natural clones were cycle sequenced after polymerase chain reaction amplification.

In addition to the iguana and caiman, the phylogenetic analyses included the following 33 species: common dogfish *Scyliorhinus canicula* (Delarbre *et al.* 1998), spotted dogfish *Mustelus manazo* (Cao *et al.* 1998), spiny dogfish *Squalus acanthias* (Rasmussen & Arnason 1999*a*), starry skate *Raja radiata* (Rasmussen & Arnason 1999*b*), loach *Crossostoma lacustre* (Tzeng *et al.* 1992), carp *Cyprinus carpio* (Chang *et al.* 1994), cod *Gadus morhua* (Johansen & Bakke 1996), Atlantic salmon *Salmo salar* (Hurst *et al.* 1999), rainbow trout *Onchorynchus mykiss* (Zardoya *et al.* 1995), African clawed frog *Xenopus laevis* (Roe *et al.* 1985), mole skink *Eumeces egregius* and green turtle *Chelonia mydas* (Kumazawa & Nishida 1999), painted turtle *Chrysemys picta* (Mindell *et al.* 1999), alligator *Alligator mississippiensis* (Janke & Arnason 1997), rook *Corvus frugilegus* (Härlid & Arnason 1998), indigo bird *Vidua chalybeata*, broadbill *Smithornis sharpei* and falcon *Falco peregrinus* (Mindell *et al.* 1999), rhea *Rhea americana* (Härlid *et al.* 1998), ostrich *Struthio camelus* (Härlid *et al.* 1998), chicken *Gallus gallus* (Desjardins & Morais 1990), 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), human *Homo sapiens* (Arnason *et al.* 1996), armadillo *Dasypus novemcinctus* (Arnason *et al.* 1997), harbour seal *Phoca vitulina* (Arnason & Johnsson 1992), horse *Equus caballus* (Xu & Arnason 1994), Indian rhinoceros *Rhinoceros unicornis* (Xu *et al.* 1996), cow *Bos taurus* (Anderson *et al.* 1982) and blue whale *Balaenoptera musculus* (Arnason & Gullberg 1993). Thus, the main lineages of different tetrapod groups are represented in the data set. The mitochondrial genome of the snake *Dinodon semicarinatus* (Kumazawa *et al.* 1998) and of the side-necked turtle *Pelomedusa subrufa* (Zardoya & Meyer 1998) were not included in the final analysis since they would have reduced the length of the alignment by 244 amino-acid sites. They also deviate significantly in amino-acid composition as tested by a 5% level $\chi^2$-test as implemented in the PUZZLE program package.

Phylogenetic analyses were performed on the concatenated sequences of 12 protein-coding genes. The data set was analysed using maximum parsimony (Fitch 1971), neighbour joining (Saitou & Nei 1987) and maximum likelihood (Felsenstein 1981) as implemented in the PHYLIP (Felsenstein 1991), MOLPHY (Adachi & Hasegawa 1996) and PUZZLE (Strimmer & Von Haeseler 1996) program packages, respectively. The mtREV24 (Adachi & Hasegawa 1996) matrix of amino-acid evolution was used for distance and maximum-likelihood analysis of the amino-acid sequence data, whereas the TN-93 model (Tamura & Nei 1993) of nucleotide evolution was used for the analysis of nucleotide sequence data. Support values for individual branches were estimated by bootstrapping or by quartet puzzling analysis (Strimmer & Von Haeseler 1996). Bootstrap probabilities for maximum-likelihood trees were calculated by the resampling estimated log-likelihood (RELL) method (Hasegawa & Kishino 1994), with standard errors (s.e.s) of the log-likelihood differences ($\Delta$lnL) being estimated using the Kishino & Hasegawa (1989) formula. The Templeton (1983) test, as implemented in the PHYLIP program package, was used for evaluating differences in the number of substitutions ($\Delta$steps) and their standard deviations (s.d.s) for different topologies.

The complete mtDNA sequences of the iguana and the caiman have been deposited in the EMBL database under accession numbers AJ278511 and AJ404872, respectively.

## 3. RESULTS

The mitochondrial genome of the iguana is 16 633 nucleotides long. The arrangement of the 22 transfer RNAs (tRNAs), the two ribosomal RNAs (rRNAs) and the 13 protein-coding genes conforms to the common vertebrate arrangement. The control region of the iguana contains the typical three conserved sequence blocks, but lacks the repetitive motifs that often characterize the mitochondrial genomes of other vertebrates. ATG was the start codon for all of the protein-coding genes except for cytochrome oxidase subunit I (COI), which starts with GTG. The most common stop codon was TAA. Four genes, COII, ATPase8, ATPase6 and NADH3, have incomplete stop codons, ending with T or TA. These are probably converted to ordinary stop codons by post-transcriptional polyadenylation (Ojala *et al.* 1981). The stop codon of COI is AGA, with that of cytochrome *b* being TAG.

Table 1. *Nucleotide composition of the protein-coding sequences*

| | first codon position | | | | second codon position | | | | third codon position | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | T | C | A | G | T | C | A | G | T | C | A | G |
| iguana | 20.1 | 28.0 | 28.9 | 23.1 | 40.4 | 28.0 | 18.8 | 12.8 | 11.3 | 43.0 | 40.8 | 4.9 |
| mole skink | 21.7 | 25.7 | 27.6 | 25.0 | 40.6 | 27.6 | 18.4 | 13.4 | 18.0 | 34.3 | 37.9 | 9.8 |
| caiman | 21.2 | 26.8 | 29.8 | 22.1 | 40.5 | 28.9 | 18.3 | 12.3 | 14.3 | 41.3 | 41.3 | 3.2 |
| alligator | 21.9 | 26.5 | 28.9 | 22.8 | 41.1 | 28.0 | 18.5 | 12.4 | 18.2 | 36.0 | 40.6 | 5.1 |
| turtles[a] | 22.6 ± 0.6 | 25.6 ± 0.9 | 30.0 ± 0.4 | 22.0 ± 0.3 | 41.1 ± 0.0 | 27.2 ± 0.0 | 18.8 ± 0.1 | 13.1 ± 0.2 | 17.5 ± 3.3 | 32.2 ± 3.3 | 48.1 ± 1.9 | 2.2 ± 1.8 |
| birds[a] | 22.0 ± 0.6 | 29.2 ± 0.5 | 27.2 ± 0.9 | 23.5 ± 0.9 | 40.4 ± 0.4 | 28.1 ± 0.4 | 19.3 ± 0.2 | 13.2 ± 0.2 | 14.5 ± 2.8 | 41.0 ± 3.3 | 39.8 ± 3.7 | 4.7 ± 1.4 |
| mammals[a] | 22.8 ± 1.7 | 24.8 ± 1.8 | 30.0 ± 0.7 | 22.4 ± 0.6 | 42.1 ± 0.5 | 26.2 ± 0.6 | 19.0 ± 0.3 | 12.7 ± 0.2 | 21.9 ± 7.1 | 31.1 ± 7.3 | 42.8 ± 3.4 | 4.3 ± 1.1 |
| *Xenopus* | 25.5 | 22.4 | 28.0 | 24.1 | 40.3 | 27.2 | 18.7 | 13.8 | 29.4 | 23.1 | 43.6 | 4.0 |
| bony fishes[a] | 21.2 ± 0.8 | 26.7 ± 0.6 | 25.4 ± 0.5 | 26.8 ± 0.3 | 40.7 ± 0.2 | 26.7 ± 0.2 | 18.5 ± 0.1 | 14.1 ± 0.2 | 23.7 ± 6.5 | 23.1 ± 5.0 | 36.5 ± 4.7 | 6.6 ± 1.5 |
| cartilagenous fishes[a] | 24.3 ± 0.6 | 24.5 ± 0.6 | 26.9 ± 0.3 | 24.2 ± 0.2 | 42.0 ± 0.0 | 25.4 ± 0.1 | 19.0 ± 0.1 | 13.6 ± 0.1 | 29.5 ± 2.5 | 27.7 ± 2.4 | 38.4 ± 1.2 | 4.4 ± 1.1 |

[a] Species as represented in the phylogenetic analysis. The base composition for each codon position is shown in percent followed by the standard deviation.

Table 2. *Relative rates among the Diapsida and Anapsida*

(The values are pairwise differences of the branch lengths given in table 3.)

| | *Xenopus* | squamates | turtles | birds |
|---|---|---|---|---|
| squamates | 1.74 | — | — | — |
| turtles | 1.74 | 1.01 | — | — |
| birds | 2.56 | 1.64 | 1.76 | — |
| crocodiles | 3.60 | 2.46 | 2.77 | 1.66 |

The gene arrangement of the 17 875-nucleotide-long caiman mitochondrial genome conforms to that of the alligator (Janke & Arnason 1997). In other crocodile mitochondrial genomes studied so far, the tRNA for phenylalanine is also located on the 5′-side of the control region (Quinn & Mindell 1996). The caiman mitochondrial genome uses ATT as a start codon in NADH1 and NADH3 and ATC as a start codon in NADH4L. The secondary structures of the reptile tRNAs conform to the general structures observed in other vertebrates. The 5′-end of the *ca*. 2500-nucleotide-long caiman control region is very similar to that of the alligator. A stretch of 454 nucleotides of the caiman and alligator control region can be easily aligned since they differ by only 24.7%. A total of 55 transitions, 51 transversions and five indels were observed in this part of the control region. The control region of the caiman is further characterized by a 288-nucleotide sequence that occurs in four nearly identical repeats plus a 166-nucleotide partial motif of that repeat, making the control region one of the longest so far described in vertebrates.

The alignment of the protein-coding genes did not include the NADH6 gene since this L (light) strand-encoded gene differs significantly in nucleotide composition from that of the 12 H (heavy) strand-encoded genes, as tested by a 5% level $\chi^2$-test. It thus violates the assumptions of character homogeneity used by most phylogenetic programs. The NADH3 gene was included in the analysis despite the uncertain nature of the extra nucleotide (cytosine) at position 174 in the non-passeriform birds. The extra nucleotide found only in these species generates a premature stop codon at position 207 in the gene (Härlid & Arnason 1998). However, it has been suggested that, in the process of translation, this extra nucleotide may be ignored (Mindell *et al.* 1998). This position was therefore eliminated from the alignment. The nucleotide sequences of the 12 protein-coding genes were aligned manually. When the snake and the side-necked turtle were included in the alignment, 2645 amino-acid positions remained for phylogenetic analysis after gaps and ambiguous sites around gaps had been removed. Excluding these two species increased the length of the data set by 9.2% to 2889 amino-acid positions.

The nucleotide compositions of the three different codon positions in the protein-coding genes of the taxa included in the analyses are shown in table 1. The iguana shows a pronounced bias in favour of cytosine (C) over thymine (T) at the third codon positions. The frog (*Xenopus*), on the other hand, has a bias towards T over C. These two species represented the two extremes in the frequency of C and T at this position. The bias in nucleotide usage at the other codon positions is limited, none of the lineages showing an extreme value. A 5% level $\chi^2$-test indicated compositional homogeneity at the second codon position and for the amino-acid sequences in all the species except the hedgehog and the bony fishes. The sequences of the snake and side-necked turtle did not pass the test for compositional homogeneity for any of the character sets.

The relative differences in evolutionary rates between the iguana and caiman and the other tetrapod lineages were calculated for each diapsid–anapsid pair (table 2) on the basis of the maximum-likelihood branch lengths of the amino-acid tree (table 3). The most pronounced rate differences among diapsids and anapsids can be observed between squamates and crocodiles. A relative-rate test (Wilson *et al.* 1977) of the amino-acid differences shows furthermore that the crocodiles have a significantly higher evolutionary rate than the remaining reptiles. However, both crocodilian sequences were included in the analysis since there is currently no alternative to the use of these data. The snake and side-necked turtle were excluded on the basis of this test since data from more slowly evolving species are available.

Two different data sets, second codon nucleotide position and amino-acid sequences, were analysed using the three most commonly employed methods for tree

Table 3. *Maximum-likelihood branch lengths of the tree shown in figure 1*

| branch | length | s.e. | branch | length | s.e. |
|---|---|---|---|---|---|
| platypus | 0.1130 | 0.0074 | a | 0.0164 | 0.0037 |
| opossum | 0.0639 | 0.0055 | b | 0.0457 | 0.0051 |
| wallaroo | 0.0495 | 0.0049 | c | 0.1099 | 0.0077 |
| hedgehog | 0.1610 | 0.0087 | d | 0.0133 | 0.0032 |
| armadillo | 0.0778 | 0.0060 | e | 0.0100 | 0.0025 |
| mouse | 0.1123 | 0.0072 | f | 0.0350 | 0.0047 |
| harbour seal | 0.0503 | 0.0046 | g | 0.0116 | 0.0026 |
| horse | 0.0292 | 0.0035 | h | 0.0259 | 0.0038 |
| Indian rhino | 0.0291 | 0.0035 | i | 0.0159 | 0.0034 |
| cow | 0.0359 | 0.0040 | j | 0.0142 | 0.0029 |
| blue whale | 0.0636 | 0.0052 | k | 0.0174 | 0.0038 |
| *Homo* | 0.1400 | 0.0081 | l | 0.0308 | 0.0046 |
| chicken | 0.0493 | 0.0047 | m | 0.0153 | 0.0031 |
| ostrich | 0.0286 | 0.0036 | n | 0.0255 | 0.0036 |
| rhea | 0.0310 | 0.0037 | o | 0.0152 | 0.0032 |
| falcon | 0.0601 | 0.0052 | p | 0.0114 | 0.0030 |
| broadbill | 0.0723 | 0.0058 | q | 0.0836 | 0.0069 |
| indigo bird | 0.0538 | 0.0049 | r | 0.0403 | 0.0047 |
| rook | 0.0363 | 0.0042 | s | 0.0325 | 0.0048 |
| caiman | 0.1160 | 0.0077 | t | 0.0452 | 0.0052 |
| alligator | 0.0622 | 0.0061 | u | 0.1942 | 0.0103 |
| iguana | 0.0758 | 0.0061 | v | 0.0223 | 0.0040 |
| mole skink | 0.0796 | 0.0063 | w | 0.0675 | 0.0059 |
| painted turtle | 0.0664 | 0.0054 | x | 0.0213 | 0.0040 |
| green turtle | 0.0256 | 0.0037 | y | 0.0594 | 0.0058 |
| *Xenopus* | 0.1081 | 0.0073 | z | 0.0537 | 0.0054 |
| trout | 0.0126 | 0.0024 | A | 0.0299 | 0.0038 |
| salmon | 0.0116 | 0.0023 | B | 0.0165 | 0.0031 |
| cod | 0.0769 | 0.0059 | C | 0.0183 | 0.0034 |
| carp | 0.0321 | 0.0039 | D | 0.0331 | 0.0045 |
| loach | 0.0471 | 0.0046 | E | 0.0210 | 0.0033 |
| starry skate | 0.0654 | 0.0055 | F | 0.0202 | 0.0039 |
| spiny dogfish | 0.0392 | 0.0041 | — | — | — |
| star spotted dogfish | 0.0261 | 0.0035 | — | — | — |
| common dogfish | 0.0327 | 0.0040 | — | — | — |

reconstruction, maximum parsimony, neighbour joining and maximum likelihood/quartet puzzling. The rationale of this approach is that the three methods differ in the assumptions upon which they are based. Congruent results obtained using these methods would increase the confidence one could place in the reconstructed phylogeny since the methods differ in their strengths and limitations.

A heuristic maximum-likelihood analysis of the amino-acid data set reconstructed the tree shown in figure 1. Since heuristic searches can become trapped in local minima an exhaustive search followed, during which the major tetrapod lineages with undisputed relationships were constrained to the following nine taxonomical units: fishes, amphibians, turtles, lizards, crocodiles, birds, monotremes, marsupials and eutherians. The same tree as shown in figure 1 was reconstructed with the highest log likelihood among all possible 135 135 topologies. The corresponding maximum-likelihood branch lengths are given in table 3. In order to investigate the relationships between lizards, crocodiles, birds and turtles further, extended maximum-likelihood and maximum-parsimony analyses were carried out. The log likelihoods of all 15 rooted trees that can be constructed from these taxa were calculated under the assumption of rate homogeneity within sites or of rate heterogeneity between sites (table 4). The gamma distribution parameter $\alpha$ was estimated as 0.77 by the PUZZLE program by assuming four classes of gamma-distributed rates among sites and one class of invariable sites. Maximum-parsimony analysis and a Templeton test of the 15 topologies complemented the analysis. The most strongly supported tree obtained for each of these tests (number 1 in table 4) conformed to the maximum-likelihood tree that is shown in figure 1. In the second- and third-best topologies, the turtles appeared as a sister group to either the crocodiles or the lizards. A turtle–crocodile relationship could not be significantly excluded at the 5% level by all of the analytical approaches and received the second-best support. This may be an artefact due to the rapid evolutionary rate of the crocodilian lineage. The existence of any other relationship between the four groups other than that shown in figure 1 can be significantly rejected by at least one method.

Analyses of the second codon nucleotide positions did not resolve the relationship of the turtles to other tetrapods conclusively, but the findings were consistent with the results based on amino-acid sequences. The 12S and 16S rRNA genes were also aligned and analysed. The total length of the combined alignment including gaps was 2950

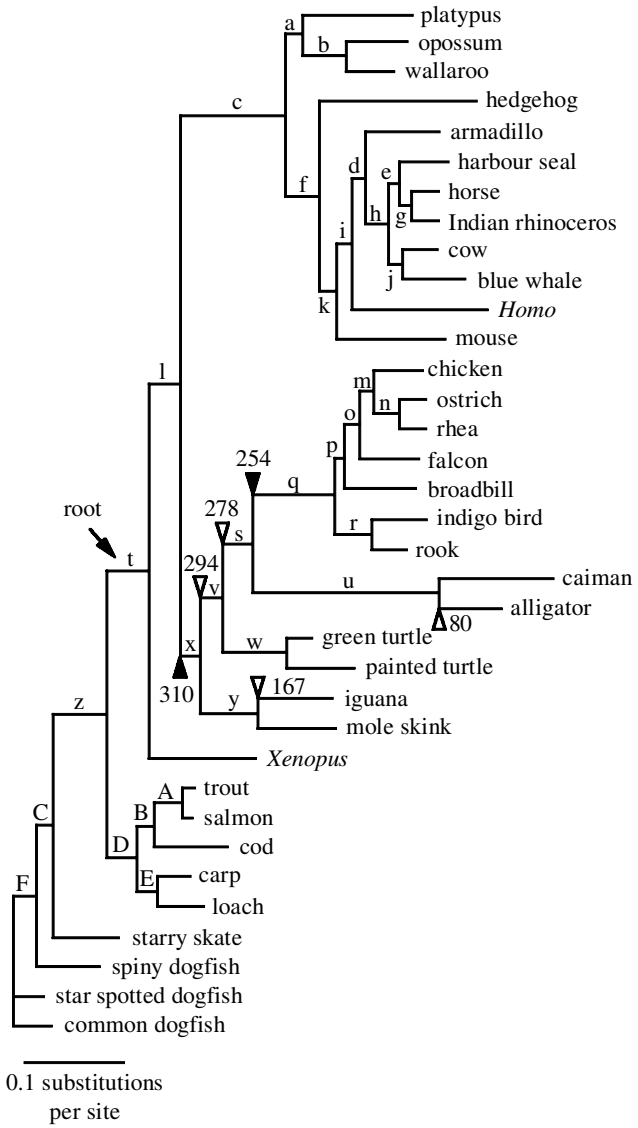


Figure 1. Maximum-likelihood tree based on analysis of amino acid sequence data. The lengths of branches a–F are given in table 3. The solid arrowheads indicate reference points used for estimates of divergence times (indicated by the open arrowheads).

positions. After the exclusion of gaps and of ambiguous sites adjacent to gaps, only 1175 (40%) positions remained for phylogenetic analysis. Although the carp has the most deviant nucleotide composition, it was retained in the data set so as to make the results comparable to those of the amino-acid analysis. A phylogenetic analysis of the rDNA alignment failed to resolve the relationship between turtles, squamates, birds, crocodiles and mammals conclusively and did not reveal larger log-likelihood differences than found in the amino-acid maximum-likelihood analysis. The quartet-puzzling and neighbour-joining analyses placed the iguana basal to the birds, the crocodile and the mammals, yet the bootstrap and quartet-puzzling support values for this relationship were below 50%. Maximum parsimony identified three equally parsimonious trees that were identical except for the position of the iguana. Maximum likelihood yielded only limited support ($\Delta\ln = 1.6$) for a topology in which the iguana represents a sister group to the birds and the crocodiles.

The divergence times for the basic splits within the diapsids, as estimated from maximum-likelihood branch lengths (table 3), are included in figure 1. However, due to the large differences in evolutionary rates in different lineages, the calculations only allow rough estimates of the divergence times. The divergence between the Synapsida and the Diapsida was set at 310 million years (Myr) ago by Benton (1990) and that between the Crocodylidae and the Aves at 254 Myr ago by Janke & Arnason (1997). The latter dating is consistent with the fossil record which allows a reasonably narrow estimate of this split (Benton 1990). The addition of the crocodilian sequences did not change the previous estimate of 254 Myr ago. The pronounced differences in evolutionary rates between the Diapsida and Anapsida make it difficult to assign specific rates to the internal branches. In order to obtain a rough estimate of divergence times, the simplest assumption, that of a constant evolutionary rate along the reference points (solid arrowheads in figure 1), was made. The origin of the Squamata was then calculated as 294 Myr ago and that of the turtles as 278 Myr ago. On the basis of the origin of crocodiles being at 254 Myr ago and assuming a constant rate along the crocodilian lineage, the divergence of the alligator and caiman was estimated to be *ca*. 80 Myr ago. Similarly, based on Squamata origin 294 Myr ago, the divergence between the mole skink and the iguana was placed at 167 Myr ago.

Both a sister-group relationship between birds and mammals, in line with the Haematotherma hypothesis and an archosaurian–mammalian relationship were investigated by means of maximum-likelihood analysis of the amino-acid sequences. Both relationships could be rejected above the 5% level of significance (s.e.s of 2.6 and 2.1, respectively).

The Marsupionta relationship remains the best-supported hypothesis regarding the relationship between monotremes, marsupials and eutherians. Support for this relationship in the neighbour-joining, FITCH (as implemented in the PHYLIP program package) and quartet-puzzling analyses was strong (96, 93 and 99% support, respectively), but no significant support was obtained for any particular arrangement in the maximum-parsimony and maximum-likelihood analyses.

## 4. DISCUSSION

The analyses showed that the evolutionary rate of the iguana mitochondrial genome is slower than that of the birds and mammals, whereas the high evolutionary rates of the alligator and the caiman mitochondrial genomes suggest that this may be a general feature of the crocodilian lineage. This and the drastically different rates in turtles and squamates contradict the hypothesis of a correlation between the rate of molecular evolution and generation time (Thomas & Beckenbach 1989; Martin & Palumbi 1993). The amino-acid composition of the caiman conforms to that assumed by the maximum-likelihood model despite the evolutionary rate of its mitochondrial genome being high. This is in contrast to the compositional bias of the rapidly evolving mitochondrial genomes of the snake and of the African side-necked turtle. Including these two reptilian taxa reduces the amino-acid alignment by 8% and also decreases the

Table 4. *Maximum-likelihood and maximum-parsimony analyses of all possible arrangements of the four main groups of Diapsida*

(The values in angled brackets show the log-likelihood values of the best tree or the number of substitutions in the maximum-parsimony tree. ΔlnL indicates the log-likelihood difference as compared to the best tree followed by the standard error and the bootstrap probability (pboot) for this particular topology. Likelihood values were calculated under the mtREV-24 model of amino-acid evolution with ($lnL_{RH}$) and without the assumption of rate heterogeneity among sites. The number of additional substitutions (Δsteps) and their standard deviations (s.d.s) are shown for the alternative topologies for the maximum-parsimony analysis. OG, outgroup; S, squamates; T, turtles; B, birds; C, crocodiles.)

| tree | ΔlnL | s.e. | pboot | $\Delta lnL_{RH}$ | s.e. | Δsteps | s.d. |
|---|---|---|---|---|---|---|---|
| 1 (OG,(S,(T,(B,C))) | <−59996.4 > | | 0.8800 | <−47290.2 > | | < 9065.0 > | |
| 2 (OG,(T,(S,(B,C))) | −55.4 | 16.6 | 0.0000 | −28.3 | 10.5 | 20.0 | 7.0 |
| 3 (OG,((B,C),(S,T)) | −43.7 | 18.4 | 0.0330 | −20.2 | 12.4 | 4.0 | 8.1 |
| 4 (OG,(B,(C,(S,T))) | −98.9 | 25.9 | 0.0000 | −42.6 | 15.8 | 40.0 | 12.4 |
| 5 (OG,(C,(B,(S,T))) | −90.5 | 26.8 | 0.0000 | −41.2 | 16.3 | 28.0 | 13.0 |
| 6 (OG,(S,(C,(B,T))) | −42.3 | 15.1 | 0.0000 | −18.1 | 17.5 | 28.0 | 10.2 |
| 7 (OG,(C,(S,(B,T))) | −73.5 | 23.1 | 0.0000 | −38.4 | 12.6 | 39.0 | 13.0 |
| 8 (OG,((B,T),(S,C)) | −79.7 | 22.1 | 0.0000 | −40.5 | 12.0 | 43.0 | 12.6 |
| 9 (OG,(B,(T,(C,S))) | −106.6 | 25.8 | 0.0000 | −49.7 | 15.1 | 53.0 | 13.1 |
| 10 (OG,(T,(B,(C,S))) | −113.2 | 24.4 | 0.0000 | −52.7 | 14.4 | 56.0 | 11.7 |
| 11 (OG,(T,(C,(B,S))) | −111.8 | 24.6 | 0.0000 | −52.2 | 14.5 | 56.0 | 11.7 |
| 12 (OG,(C,(T,(B,S))) | −104.2 | 26.4 | 0.0000 | −50.3 | 15.2 | 44.0 | 13.5 |
| 13 (OG,((B,S),(C,T)) | −75.2 | 25.5 | 0.0000 | −37.8 | 14.4 | 35.0 | 13.0 |
| 14 (OG,(B,(S,(C,T))) | −69.9 | 26.0 | 0.0030 | −35.3 | 14.9 | 31.0 | 13.3 |
| 15 (OG,(S,(B,(C,T))) | −23.4 | 17.5 | 0.0840 | −10.6 | 19.1 | 20.0 | 10.0 |

general resolution along the diapsid branches. This is probably a consequence of (i) the data set being smaller, (ii) the introduction of homoplastic sites, and/or (iii) the significantly biased amino-acid composition of these two species. However, the increasing number of reptile mitochondrial genomes available now makes it possible to select less biased and more slowly evolving lineages and to eliminate the rapid ones without representation of any major lineages being reduced to just one species. However, this is currently not possible for the crocodiles since the rapid evolutionary rate of their mitochondrial genome seems to be a general feature of this lineage.

Since maximum-likelihood analyses produce unrooted trees, the polarity of the basal vertebrate relationships (fishes) cannot be automatically inferred. The results of other studies (Rasmussen & Arnason 1999*a*,*b*) in which non-gnathostomous outgroups were included in the analysis have suggested that cartilaginous fishes (neoselachians) are not basal to other vertebrates. The split most probably lies between the tetrapods and all of the gnathostomous (jawed) fishes, as indicated by the position of the root in figure 1.

The phylogenetic analyses of the diapsid–anapsid relationships using different analytical methods resulted in a stable position of the turtles at the base of the archaeosaurian (bird + crocodiles) branch. Any alternative positions of the turtles were significantly rejected by at least one analytical approach in all cases. As a consequence of the placement of the turtles among the diapsids, the anapsid condition in the turtles appears to have evolved secondarily from a diapsid state. This makes the character of the temporal openings found on the skull appear highly problematical despite its popularity since it appears to be subject to convergent evolution. A recent morphological study supported this view and placed the turtles among the diapsids without an alliance to any specific lineage (Rieppel 1999) or, alternatively, as being more closely related to the lepidosaurian than to the acheosaurian lineage. Interestingly, Løvtrup (1977) suggested a sister-group relationship between the turtles and the birds plus crocodiles, basing his conclusions on three morphological characters. Although others have criticized this classification, it exemplifies the difficulties in defining the relationships of the turtles to other animals. The first convincing proposal of a turtle–archaeosaurian relationship came from comprehensive molecular data on completely sequenced mitochondrial genomes (Kumazawa & Nishida 1999).

Studies of extinct eurapsids (pleisosaurs and ichthyosaurs) that are characterized by an upper temporal skull opening have shown the condition of temporal openings to be variable. The eurapsid state in pleisosaurs can be derived from the diapsid condition on the basis of a closing of the lower temporal opening (Carroll 1988). It is probable that mechanical stress leads to a thickening of the affected bone and that the absence of such stress leads to reduced bone thickness or to a complete absence of bone tissue in the area in which no mechanical stress is present. In eurapsids, therefore, differential degrees of mechanical stress and of muscle attachment appear to have modified the initially diapsid condition to produce the eurapsid state (Carroll 1988). It is surprising that, despite this long being known, systematics have continued to denote the differing states of the temporal skull openings as being a major differentiating character within the amniotes, with respect to the turtles in particular. By the same mechanism of reduction of the lower temporal openings in the eurapsids, the turtles appear to have reversed to the ancestral character state by having no temporal skull openings at all.

The estimated divergence time for turtles is 278 Myr ago. This molecular-based estimate is congruent with the palaeontological record, despite turtles not being found to be basal to diapsids. The age of the oldest anapsid fossil,

*Protocaptorhinus*, is given as 272 Myr ago (Benton 1990). Since the anapsids seem to be paraphyletic themselves (Carroll 1988; Rieppel 1999) this fossil may be of limited value because it defines the earliest possible appearance of an anapsid which may even not be an ancestor to the turtles. The oldest turtle fossil from the upper Triassic period (*Triassochyles*) is considerably younger than the molecular estimate of its origin places it. The same discrepancy of molecular and fossil data is observed for the age of the Lepidosaurian lineage (squamates and sphenodontids). On the basis of molecular data their origin is estimated at 294 Myr ago, which predates the age of the oldest fossil of this lineage, *Youngina*, which has been dated to 249 Myr ago (Benton 1990). This discrepancy is not surprising given the unequal amount of time required for morphological and molecular differences to arise.

The results of the phylogenetic analysis provided further support for the generally accepted sister-group relationship of birds and crocodiles, thus contradicting the Haematotherma hypothesis. Previous analysis of single genes or small data sets, such as that of combined 12S and 16S rDNA sequences or a limited representation of taxa, gave contradictory answers to this question. This is not unexpected due to the effect of the stochastic fluctuations that occur in small data sets (Cao *et al.* 1994). It has recently been shown that 12S rDNA sequences, despite their popularity, have a somewhat limited ability to resolve ordinal relationships, even those as recent as divergences among mammals (Emerson *et al.* 1999).

The phylogenetic analysis presented here corroborates earlier studies of tetrapods, such as those of the non-basal position of palaeognath birds (Mindell *et al.* 1997; Härlid *et al.* 1998) and of mammalian relationships. In the first molecular study based on the analysis of a complete monotreme mitochondrial genome (Janke *et al.* 1996), the traditionally accepted relationship of the three main mammalian groups (Monotremata, Marsupialia and Eutheria) was challenged. The analysis failed to support the commonly accepted marsupial–eutherian clade, but indicated that monotremes are a sister group to the marsupials, which is in line with the Marsupionta hypothesis (Kühne 1973, 1975). This relationship was the best supported one of the three alternatives, even after the inclusion of further ingroup and outgroup species (Janke & Arnason 1997; Janke *et al.* 1997; Kumazawa & Nishida 1999). The current addition of the iguana and caiman mitochondrial genomes to the phylogenetic analyses yielded additional support for the Marsupionta hypothesis.

We would like to thank Dr Christine Lendel for tissue samples. The Nilsson–Ehle and Carl Trygger foundations granted financial support for this work.

## REFERENCES

Adachi, J. & Hasegawa, M. 1996 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. & 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., Xu, X. & Gullberg, A. 1996 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 mitochondrial DNA suggest a sister group relationship between Xenarthra (Edentata) and Ferungulates. *Mol. Biol. Evol.* **14**, 762–768.

Benton, M. J. 1990 Phylogeny of the major tetrapod groups: morphological data and divergence dates. *J. Mol. Evol.* **30**, 409–424.

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.

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.

Cao, Y., Waddell, P. J., Okada, N. & Hasegawa, M. 1998 The complete mitochondrial DNA sequence of the shark *Mustelus manazo*: evaluating rooting contradictions to living bony vertebrates. *Mol. Biol. Evol.* **15**, 1637–1646.

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

Caspers, G.-J., Reinders, G.-J., Leunissen, J. A. M., Wattel, J. & De Jong, W. W. 1996 Protein sequences indicate that turtles branched off from the amniote tree after mammals. *J. Mol. Evol.* **42**, 580–586.

Chang, Y. S., Huang, F. L. & Lo, T. B. 1994 The complete nucleotide sequence and gene organization of carp (*Cyprinus carpio*) mitochondrial genome. *J. Mol. Evol.* **38**, 138–155.

de Jong, W. W., Zweers, A., Versteeg, M., Dessauer, H. C. & Goodman, M. 1985 α-Crystallin A sequences of *Alligator mississippiensis* and lizard *Tupinambis teguixin*: molecular evolution and reptilian phylogeny. *Mol. Biol. Evol.* **2**, 484–493.

Delarbre, C., Spruyt, N., Delmarre, C., Gallut, C., Barriel, V., Janvier, P., Laudet, V. & Gachelin, G. 1998 The complete nucleotide sequence of the mitochondrial DNA of the dogfish, *Scyliorhinus canicula*. *Genetics* **150**, 331–344.

Dene, H., Sazy, J., Goodman, M. & Romero-Herrera, A. E. 1982 The amino acid sequence of alligator (*Alligator mississippiensis*) myoglobin. *Biochim. Biophys. Acta* **624**, 397–408.

Desjardins, P. & Morais, R. 1990 Sequence and gene organization of the chicken mitochondrial gene. A novel gene order in higher vertebrates. *J. Mol. Biol.* **212**, 599–634.

Eernisse, D. J. & Kluge, A. G. 1993 Taxonomic congruence versus total evidence, and amniote phylogeny inferred from fossils, molecules, and morphology. *Mol. Biol. Evol.* **10**, 1170–1195.

Emerson, G. L., Kilpatrick, C. W., McNiff, B. E., Ottenwalder, J. & Allard, M. W. 1999 Phylogenetic relationships of the order Insectivora based on complete 12S rRNA sequences from mitochondria. *Cladistics* **15**, 221–230.

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

Felsenstein, J. 1991 *Phylogenetic inference programs* (*PHYLIP*). University of Washington and Berkeley, CA: University Herbarium and University of California.

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

Gardiner, B. G. 1982 Tetrapod classification. *Zool. J. Linn. Soc.* **74**, 207–232.

Gardiner, B. G. 1993 Haematothermia: warm-blooded amniotes. *Cladistics* **9**, 369–395.

Gauthier, J. A., Kluge, A. G. & Rowe, T. 1988 Amniote phylogeny and the importance of fossils. *Cladistics* **4**, 104–209.

Härlid, A. & Arnason, U. 1998 Analysis of mitochondrial DNA nest ratite bird within the neognathae—supporting a neotenous origin of ratite morphological characters. *Proc. R. Soc. Lond.* B **266**, 1–5.

Härlid, A., Janke, A. & Arnason, U. 1998 The complete mitochondrial genome of *Rhea americana* and early avian divergencies. *J. Mol. Evol.* **46**, 669–679.

Hasegawa, M. & Kishino, H. 1994 Accuracies of the simple methods for estimating the bootstrap probability of a maximum likelihood tree. *Mol. Biol. Evol.* **11**, 142–145.

Hedges, S. B. 1994 Molecular evidence for the origin of birds. *Proc. Natl Acad. Sci. USA* **91**, 2621–2624.

Hedges, S. B. & Poling, L. L. 1999 A molecular phylogeny of reptiles. *Science* **283**, 998–1001.

Hedges, S. B., Moberg, K. D. & Maxson, L. R. 1990 Tetrapod phylogeny inferred from 18S and 28S ribosomal RNA sequences and a review of the evidence for amniote relationships. *Mol. Biol. Evol.* **7**, 607–633.

Hurst, C. D., Bartlett, S. E., Davidson, W. S. & Bruce, I. J. 1999 The complete mitochondrial DNA sequence of the Atlantic salmon, *Salmo salar*. *Gene* **239**, 237–242.

Janke, A. & Arnason, U. 1997 The complete mitochondrial genome of *Alligator mississippiensis* and the separation between recent archosauria (birds and crocodiles). *Mol. Biol. Evol.* **14**, 1266–1272.

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 phylogenetic relationship among Monotremata, Marsupialia and Eutheria. *Proc. Natl Acad. Sci. USA* **94**, 1276–1281.

Johansen, S. & Bakke, I. 1996 The complete mitochondrial DNA sequence of Atlantic cod (*Gadus morhua*): relevance to taxonomic studies among codfishes. *Mol. Mar. Biol. Biotech.* **5**, 203–214.

Kirsch, J. A. W. & Mayer, G. C. 1998 The platypus is not a rodent: DNA hybridization, amniote phylogeny and the palimpsest theory. *Phil. Trans. R. Soc. Lond.* B **353**, 1221–1237.

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.

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.

Kühne, W. G. 1973 The systematic position of monotremes reconsidered (Mammalia). *Z. Morph. Tiere* **75**, 59–64.

Kühne, W. G. 1975 Marsupium and marsupial bone in Mesozoic mammals and in the marsupionta. *Coll. Int. CNRS* **218**, 585–590.

Kumazawa, Y. & Nishida, M. 1993 Sequence evolution of mitochondrial tRNA genes and deep-branch animal phylogenetics. *J. Mol. Evol.* **37**, 380–398.

Kumazawa, Y. & Nishida, M. 1999 Complete mitochondrial DNA sequences of the green turtle and blue-tailed mole skink: statistical evidence for archosaurian affinity of turtles. *Mol. Biol. Evol.* **16**, 784–792.

Kumazawa, Y., Ota, H., Nishida, M. & Ozawa, T. 1998 The complete nucleotide sequence of a snake (*Dinodon semicarinatus*) mitochondrial genome with two identical control regions. *Genetics* **150**, 313–329.

Løvtrup, S. 1977 *The phylogeny of vertebrata*. London: Wiley.

Martin, A. P. & Palumbi, S. R. 1993 Body size, metabolic rate, generation time, and the molecular clock. *Proc. Natl Acad. Sci. USA* **90**, 4087–4091.

Mindell, D. P., Sorenson, M. D., Huddleston, C. J., Miranda Jr, H. C., Knight, A., Sawchuk, S. J. & Yuri, T. 1997 Phylogenetic relationships among and within select avian orders based on mitochondrial DNA. In *Avian molecular evolution and systematics* (ed. D. P. Mindell), pp. 213–247. New York: Academic Press.

Mindell, D. P., Sorenson, M. D. & Dimcheff, D. E. 1998 An extra nucleotide is not translated in mitochondrial ND3 of some birds and turtles. *Mol. Biol. Evol.* **11**, 1568–1571.

Mindell, D. P., Sorenson, M. D., Dimcheff, D. E., Hasegawa, M., Ast, J. C. & Yuri, T. 1999 Interordinal relationships of birds and other reptiles based on whole mitochondrial genomes. *Syst. Biol.* **48**, 138–152.

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

Owen, R. 1866 *On the anatomy of vertebrates*, vols I and II. London: Longmans, Green.

Perutz, M. F., Bauer, C., Gros, G., Leclercq, F., Vandecasserie, C., Schnek, A. G., Braunitzer, G., Friday, A. E. & Joysey, K. A. 1981 Allosteric regulation of crocodilian haemoglobin. *Nature* **291**, 682–684.

Quinn, T. W. & Mindell, D. P. 1996 Mitochondrial gene order adjacent to the control region in crocodile, turtle and tuatara. *Mol. Phylogenet. Evol.* **5**, 344–351.

Rasmussen, A. S. & Arnason, U. 1999*a* Phylogenetic studies of complete mitochondrial DNA molecules place cartilaginous fishes within the tree of bony fishes. *J. Mol. Evol.* **48**, 118–123.

Rasmussen, A. S. & Arnason, U. 1999*b* Molecular studies suggest that cartilaginous fishes have a terminal position in the piscine tree. *Proc. Natl Acad. Sci. USA* **96**, 2177–2182.

Rieppel, O. 1999 Turtle origins. *Science* **283**, 945–946.

Rieppel, O. & DeBraga, M. 1996 Turtles as diapsid reptiles. *Nature* **384**, 453–455.

Roe, B. A., Ma, D.-P., Wilson, R. K. & Wong, J. F.-H. 1985 The complete nucleotide sequence of the *Xenopus laevis* mitochondrial genome. *J. Biol. Chem.* **260**, 9759–9774.

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

Sambrook, J., Fitsch, E. F. & Maniatis, T. 1989 *Molecular cloning, a laboratory manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

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

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

Tamura, K. & Nei, M. 1993 Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10**, 512–526.

Templeton, A. 1983 Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to humans and apes. *Evolution* **37**, 221–244.

Thomas, W. K. & Beckenbach, A. T. 1989 Variation in salmonid mitochondrial DNA: evolutionary constraints and mechanisms of substitution. *J. Mol. Evol.* **29**, 233–245.

Tzeng, C.-S., Hui, C.-F., Shen, S.-C. & Huang, P. C. 1992 The complete nucleotide sequence of the *Crossostoma lacustre* mitochondrial genome: conservation and variation among vertebrates. *Nucl. Acids. Res.* **20**, 4853–4858.

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., 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.

Zardoya, R. & Meyer, A. 1998 Cloning and characterization of a microsatellite in the mitochondrial control region of the African side-necked turtle, *Pelomedusa subrufa*. *Gene* **216**, 149–153.

Zardoya, R., Garrido-Pertierra, A. & Bautista, J. M. 1995 The complete nucleotide sequence of the mitochondrial DNA genome of the rainbow trout, *Oncorhynchys mykiss*. *J. Mol. Evol.* **41**, 942–951.

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