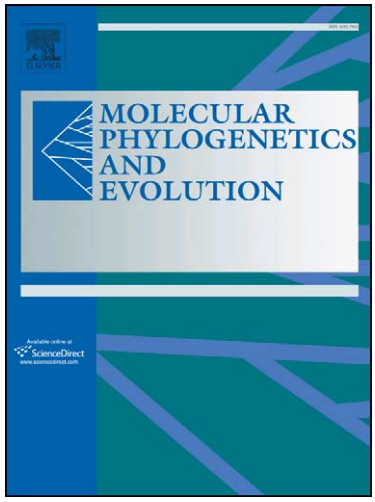


# Extended mitogenomic phylogenetic analyses yield new insight into crocodylian evolution and their survival of the Cretaceous-Tertiary boundary

Jonas Roos, Ramesh K. Aggarwal, Axel Janke

PII: S1055-7903(07)00222-9
DOI: 10.1016/j.ympev.2007.06.018
Reference: YMPEV 2598

To appear in: *Molecular Phylogenetics and Evolution*

Received Date: 20 February 2007
Revised Date: 7 June 2007
Accepted Date: 22 June 2007

Please cite this article as: Roos, J., Aggarwal, R.K., Janke, A., Extended mitogenomic phylogenetic analyses yield new insight into crocodylian evolution and their survival of the Cretaceous-Tertiary boundary, *Molecular Phylogenetics and Evolution* (2007), doi: 10.1016/j.ympev.2007.06.018

# Extended mitogenomic phylogenetic analyses yield new insight into crocodylian evolution and their survival of the Cretaceous-Tertiary boundary.

**Jonas Roos[1], Ramesh K. Aggarwal[2], Axel Janke[1]**

[1] *Department of Cell and Organism Biology, Division of Evolutionary Molecular Systematics, University of Lund, Sölvegatan 29, S-223 62 Lund, Sweden*

[2] *Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad 500 007, India*

Running Head - Crocodylian evolution

Corresponding author: Axel Janke

E-mail: axel.janke@cob.lu.se

Phone: +46 46 2227849

Fax: +46 46 147874

**Abstract**

The mitochondrial genomes of the dwarf crocodile, *Osteolaemus tetraspis*, and two species of dwarf caimans, the smooth-fronted caiman, *Paleosuchus trigonatus*, and Cuvier's dwarf caiman, *Paleosuchus palpebrosus*, were sequenced and included in a mitogenomic phylogenetic study. The analyses also provided molecular estimates of various crocodylian divergences applying recently established crocodylian and outgroup fossil calibration points. The phylogenetic analyses, which included a total of ten crocodylian species, yielded strong support to a basal split between Crocodylidae and Alligatoridae. *Osteolaemus* fell within the Crocodylidae as the sister group to *Crocodylus*. *Gavialis* and *Tomistoma*, which joined on a common branch, constituted a sister group to *Crocodylus*/*Osteolaemus*. Within the Alligatoridae there was a basal split between *Alligator* and a branch that contained *Paleosuchus* and *Caiman*. Molecular estimates based on amino acid data placed the divergence between Crocodylidae and Alligatoridae at 97-103 million years ago and that between Alligator and Caiman/Paleosuchus at 68-72 million years ago. Other crocodilian divergences were placed after the Cretaceous-Tertiary boundary. Thus, according to the molecular estimates, only three lineages among the currently recognized extant crocodylian species have their roots in the Cretaceous. Considering the crocodylian diversification in the Cretaceous the molecular datings suggest that the extinction of the dinosaurs was also to some extent paralleled in the crocodylian evolution. However, for whatever reason, some crocodylian lineages survived into the Tertiary.

*Keywords:* Crocodylia; crocodylian evolution; mass extinction; mitogenomics; *Osteolaemus*; *Paleosuchus*;

## 1. Introduction

Extant crocodylians constitute a small order, Crocodylia, within the class Reptilia. They have a rich fossil record that extends into the late Cretaceous (Brochu, 2001). Crocodylians represent one of only two archosaurian lineages – the other being birds (Aves) – that survived the mass extinction connected to the K/T boundary, at approximately 65 MYA (million years ago). Today, 23 crocodylian species exist. They are, based on current molecular data, divided into two families and eight genera. The family Alligatoridae consists of the genera *Alligator, Caiman, Paleosuchus* and *Melanosuchus*, whereas Crocodylidae consists of *Crocodylus, Osteolaemus, Tomistoma* and *Gavialis*. Previously, the gharial, *Gavialis gangeticus,* was believed to constitute a separate family, Gavialidae. For this reason, some authors (e.g. Willis et al., 2007) still refer to *Gavialis/Tomistoma* as Gavialidae. Because of this, and to avoid similar confusion, we wish to stress that throughout the text, when using the expression "extant crocodylian lineage", we refer only to lineages among, or leading to, the currently recognized, extant crocodylian genera. Currently, fossil and molecular estimates suggest that at least three, and possibly no more than five, extant crocodylian lineages survived the K/T boundary (Brochu, 2003; Janke et al., 2005).

Crocodylian relationships were recently examined in a phylogenetic analysis based on complete mitochondrial (mt) genomes (Janke et al., 2005). The study provided the first conclusive molecular evidence for a position of *Gavialis* within the Crocodylidae. The analyses recognized a sister group relationship between the gharial and the false gharial, *Tomistoma schlegelii*, and a basal Crocodylidae split between *Crocodylus* and *Gavialis*/*Tomistoma*. The findings were inconsistent with the majority of previous phylogenetic proposals based on morphological data (e.g. Salisbury and

Willis, 1996; Brochu, 1997). Both views on crocodilian relationships are illustrated in Fig. 1. Although the result of Janke et al. (2005) could be anticipated from previous molecular studies (Gatesy et al., 1993; Aggarwal et al., 1994; White and Densmore, 2000; Gatesy et al., 2003; Harshman et al., 2003), these studies had either lacked an outgroup (Gatesy et al., 1993; Aggarwal et al., 1994), or were based on much less extensive sequence data (White and Densmore, 2000; Gatesy et al., 2003; Harshman et al., 2003). These previous studies could therefore not reject alternative hypotheses, while the mitogenomic study (Janke et al., 2005) could significantly reject alternative trees, notably those with *Gavialis* in a basal position relative to the remaining crocodylian species.

Despite the establishment of the basal structure of the crocodylian tree some phylogenetic questions related to less deep divergences still exist. These questions include the position of the dwarf crocodile, *Osteolaemus tetraspis,* and the two dwarf caimans, the smooth-fronted caiman, *Paleosuchus trigonatus*, and Cuvier's dwarf caiman, *Paleosuchus palpebrosus*. *Osteolaemus* has been placed as the sister group to *Crocodylus* in the traditional morphological and molecular trees (Salisbury and Willis, 1996; Brochu, 1997; Harshman et al., 2003). However, with *Gavialis* and *Tomistoma* being the probable sister group to *Crocodylus*, the phylogenetic position of *Osteolaemus* needs to be reexamined. *Osteolaemus* might be basal to all *Crocodylus* and *Gavialis/Tomistoma*, or the sister taxon to one of these. The problems commonly associated with resolving crocodylian relationships were underlined in recent studies (Schmitz et al., 2003, McAliley et al., 2006), which provided molecular evidence against the monophyly of *Crocodylus*. With respect to *Paleosuchus*, morphological and molecular data have suggested a position of the genus as sister group to *Caiman*/*Melanosuchus* within the Alligatoridae (Brochu, 2003; Gatesy et al.,

2003; Harshman et al., 2003). However, on the basis of DNA fingerprinting, *Paleosuchus* has also been identified as sister group to *Alligator* (Aggarwal et al., 1994). The support for a particular position of *Paleosuchus* within the Alligatoridae has been somewhat limited in general, however, indicating the need for more comprehensive data for resolving its position.

Janke et al. (2005) provided molecular estimates of several crocodylian divergences applying three non-crocodylian calibration points and the avian-crocodylian split. These mitogenomic estimates placed several crocodylian divergences unexpectedly early, suggesting *inter alia* that at least five extant crocodylian lineages survived the K/T boundary. The estimates were to some extent inconsistent with the crocodylian fossil record, which suggested more recent divergences (Brochu, 2001, 2003). Müller and Reisz (2005) recently established a new crocodylian calibration point, which resides within the crocodylian tree and constraints the split between *Caiman* and *Alligator* to 66-71 MYA. It has been included in the current study together with the calibration points applied by Janke et al. (2005), which have been revised in accordance with Benton and Donoghue (2007). The increased taxon sampling in this study, which includes ten of 23 crocodylian species and seven of eight crocodylian genera, will allow the estimation of crocodylian divergence times with higher accuracy than previously possible.

The phylogenetic position of turtles among the Reptilia has not yet been conclusively resolved. Using morphological data, Rieppel (1999) suggested that they fall within the Diapsida (lizards, turtles, birds and crocodiles), a hypothesis that was molecularly supported by Hedges and Poling (1999) and Mannen and Li (1999), who placed turtles as a sister group to crocodylians. However, molecular analyses in general have rather argued for a sister group relationship between turtles and

Archosauria (birds plus crocodiles) (Zardoya and Meyer, 1998; Kumazawa and Nishida, 1999; Rest et al., 2003; Iwabe et al., 2005; Janke et al., 2005). In addition to the crocodylian analyses the current taxon sampling has permitted an examination of the phylogenetic position of turtles relative to Crocodylia.

## 2. Materials and methods

### *2.1 PCR amplification*

Takara Ex-Taq polymerase and Fermentas High Fidelity PCR Enzyme Mix were used to PCR amplify mtDNA sequences from whole genomic DNA from *O. tetraspis, P. trigonatus* and *P. palpebrosus* on a RoboCycler® 96 Temperature Cycler. The reactions were made according to the manufacturers' specification and adapted for the specific $T_m$ values and extension times of the primer pairs.

Initial PCR amplification of the mtDNA was done with conserved primers located in the 12S rRNA and 16S rRNA genes and in the anticodon loop of tRNA genes. These primers had either been used before (Janke et al., 2005) or were specifically designed. After amplification, the PCR products were purified twice by precipitation with 2.5 volumes 100% ethanol and 0.1 volumes 3M sodium acetate (pH 5.6) using standard protocol procedures (Sambrook and Russell, 2001). The purified PCR products were sequenced with PCR primers or specific primers by primer walking. Sequencing was done using BIG-DYE version 3.1 cycle sequencing kit on an ABI prism 3100 Genetic Analyser. All newly obtained sequences were aligned by BLAST-search to homologous crocodylian sequences that were available in the NCBI

database (http://www.ncbi.nlm.nih.gov/), for detecting PCR contamination or artifacts. All tRNAs were identified by sequence comparison to published crocodylian mt genomes and they were folded into their putative secondary structure. This allowed identification of the beginning and end of these genes and the neighboring protein coding genes as well as structural aberrations in individual tRNAs.

*2.1 Data alignment*

The phylogenetic analyses were based on the twelve H strand encoded protein-coding sequences in the *O. tetraspis*, *P. palpebrosus* and *P. trigonatus* mt genomes and the 20 other mt genomes given in Table 1. The L strand encoded gene NADH6 was excluded from the analyses because the nucleotide composition of this gene differs significantly from that of the twelve H strand encoded genes. This deviating composition would interfere with the evolutionary models and the assumption of compositional homogeneity of most phylogenetic programs. The tRNA and rRNA genes were excluded from the analyses to avoid problems caused by their deviating mode of evolution, such as nucleotide-nucleotide interactions in stem regions in these molecules.

Alignment of the sequences was done manually in PAUP* v 4.0b10 (Swofford, 2003) and was unproblematic. After alignment, gaps and ambiguous sites adjacent to gaps were removed. For the nucleotide analysis, all third codon positions were removed and first codon positions in the two-fold degenerate leucine (Leu) codons were substituted for Y. This was done to increase the nucleotide homogeneity of the dataset and to avoid analyzing synonymous positions. In total, 6464 first and second

codon nucleotide (nt) positions and 3232 amino acids (aa) were left for the phylogenetic analyses.

*2.3 Phylogenetic reconstruction*

Three different types of phylogenetic analyses were performed: Maximum Parsimony, MP, Neighbor Joining, NJ, and Maximum Likelihood, ML. MP and nt NJ analyses were done in PAUP*, while aa NJ analysis was done on distance tables from TREE-PUZZLE v 5.2, TP, (Schmidt et al., 2002) and the NEIGHBOUR v 3.5 program (Felsenstein, 1993). ML analyses were done in TREEFINDER v May 2007, TF, (Jobb et al., 2004) and TP. For the MP analyses on the nt and aa sequence data, heuristic searches were made with PAUP* using the TBR swapping algorithm and 1000 repetitions.

All nt analyses assumed the GTR (Lanave et al., 1984) model of sequence evolution, as suggested by Modeltest version 3.7 (Posada and Crandall 1998). The aa sequence analyses assumed the mtREV-24 (Adachi and Hasegawa, 1996) model of sequence evolution. ML analyses in TP and TF were done assuming both rate homogeneity and rate heterogeneity; the heterogeneity analyses assuming eight classes of gamma distributed rate categories (Yang, 1994) and one class of invariable sites (8Γ+I).

The robustness of the ML trees was analyzed by evaluating the log-likelihood values (logL) and Shimodaira-Hasegawa probabilities, (pSH; Shimodaira and Hasegawa, 1999) of alternative topologies. In these analyses *Osteolaemus, Paleosuchus*, *Gavialis*, the turtles and the lizards were placed at alternative positions in the tree and the resulting topologies were analyzed by ML methods in TP, using the

same models and parameters as given above.

*2.4 Divergence time estimates*

Divergence times were estimated with the r8s v 1.70 (Sanderson, 2002) and TF program packages on the aa and nt ML trees provided by TF. The estimates were based on either a penalized likelihood model (r8s) or nonparametric rate smoothing (TF). Errors were estimated on 100 bootstrap replicates of the ML trees and were recorded as the standard deviations (s.d.) of the mean estimates. Divergence times and their s.d. were also estimated on the same nt and aa data by multidivtime, MDT (Thorne and Kishino, 2002), as implemented in the T3 program package v 1.0 2003 (http://abacus.gene.ucl.ac.uk/).

The nt and aa analyses in MDT assumed eight gamma distributed rate categories (8Γ), but no invariable category, because the program had no option for this. The nt parameters were estimated by the baseml package. The observed nucleotide composition (A = 22.8 %, G = 19.3%, T = 33.7 %, C = 24.2%), the gamma distribution parameter ($\alpha = 0.29$) and the relative rates between the eight gamma distribution categories (0.0004, 0.0084, 0.0455, 0.1455, 0.3611, 0.7904, 1.6904, 4.9584) were used to build the model files. For the aa analysis the model files with eight gamma distributed rate categories (8Γ) were built according to Y. Kumazawa's instructions (ftp://statgen.ncsu.edu/pub/thorne/Kumazawa.tgz) on the basis of the aa frequency, tree, substitution matrix and rate heterogeneity of the dataset and the eigen values that were estimated by the correspondingly modified codeml program. The gamma distribution parameter ($\alpha$) was estimated to 0.53 and the eight gamma distributed rate categories to 0.01041, 0.06788, 0.18132, 0.36126, 0.63236, 1.04999,

1.76919 and 3.92759. For both nt and aa analyses the maximum age (bigtime) was constraint to 500 MYA as the oldest imaginable and undisputable time for any divergence in the dataset. The prior expected time between the tip and root (rttm) and the standard deviation (rttmsd) were set to 310 and 10 MYA respectively. From 100,000 Markov chain samples each 100$^{th}$ sample was collected after 100,000 burnin cycles. Remaining parameters were set according to J. Thorne's instructions (multidivtime.readme file). The divergence time estimates and their s.d. given by MDT were recorded.

Five different fossil-based calibration points were used to constrain the divergence time estimates. One of the calibration points, the caiman-alligator split set at 66-71 MYA (Müller and Reisz, 2005), has the advantage to be located within the order Crocodylia. The second crocodylian-related reference point was at the avian-crocodylian split at 235-250 MYA (Benton and Donoghue, 2007). The remaining three calibration points were among the outgroups and included the mouse-cow split at 95-113 MYA, the opossum-kangaroo split at 62-71 MYA and the marsupial-placental (opossum-mouse) split at 125-138 MYA. All outgroup calibration dates followed the recommendations by Benton and Donoghue (2007). The opossum-kangaroo split is notably consistent with the most recent molecular-based divergence time of this group (Nilsson et al., 2004).

The divergence times were also estimated by excluding one reference point at a time in order to examine the consistency among the reference points, notably to the newly established caiman-alligator calibration point.

## 3. Results

### *3.1 Genomes and gene features*

All regions in the crocodylian mt genomes were PCR amplified and sequenced, except for the complete control region (CR) in the two *Paleosuchus* species and the adjacent tRNA-Thr, tRNA-Pro and tRNA-Phe genes in *P. palpebrosus*. The various PCR fragments overlapped with 300-500 bp and neighboring sequences from primer walking overlapped with 100-200 bases. Sequencing was generally performed on only one of the strands, except when sequencing errors occurred or when sequencing artifacts could be suspected. In these cases, the affected regions were sequenced from both strands and from different PCR fragments. A near two-fold sequence coverage was achieved, on average, for each mt genome. The CR in *Paleosuchus* could not be completely sequenced despite several attempts. The three new mt genomes have been deposited in the EMBL database under the accession numbers AM493868 (*O. tetraspis*), AM493869 (*P. trigonatus*) and AM493870 (*P. palpebrosus*).

The organization of the genomes (illustrated for *O. tetraspis* in Fig. 2) is consistent with that of *A. mississippiensis*, the first crocodylian mt genome described (Janke and Arnason 1997). In contrast to the general vertebrate pattern, the crocodylian tRNA-Phe gene is not positioned upstream of the 12S rRNA gene, but instead clusters with the tRNA-Thr and tRNA-Pro genes. Using the one letter code for the aa that the respective tRNAs carry, the three genes form the so-called TPF cluster downstream of the control region. Also the order (SHL) of tRNA-Ser(AGY), tRNA-His and tRNA-Leu(CUN) genes deviates from the general (HSL) vertebrate scheme. The CR in *O. tetraspis* is unique in that it has a long stretch of 51 consecutive adenine and 13

consecutive cytosine sites. These two features have been confirmed by comparison with a partial *O. tetraspis* CR sequence (AF460217). The adenine stretch of AF460217 is slightly shorter (43 bases) and the cytosine stretch is interrupted by a single thymine. Interestingly, a long stretch of 46 uninterrupted adenines has also been observed in the mt genome of *Crocodylus porosus*.

The lengths, the start, and the stop codons of the protein-coding genes in the three new genomes conform to those in previously sequenced crocodylian mt genomes (Janke and Arnason, 1997; Janke et al., 2001, 2005; Wu et al., 2002 unpubl.). However, in *O. tetraspis* a premature stop codon in the sequence of the NADH5 gene has been found, which, if functional, significantly reduces the length of the protein. Both strands of the region (nt positions 13711-13713 in EMBL accession AM493868) were sequenced several times, using different PCR products. The premature stop codon is created by insertion of an extra adenine at nt position 1821 in the coding strand of the presumed gene, changing a GGC (Gly) codon to AGG (stop). A removal of the A extends the open reading frame by 25 nt, yielding a sequence similar in length to that of *Crocodylus*. It is possible that the additional adenine is removed by RNA editing, though this has not been investigated further. The possible RNA-editing in this protein-coding gene has however no effect on the phylogenetic analyses, because the region was not included in the final alignment.

The genes for CO1 and tRNA-Ser are separated by a stretch of 41 nt in *O. tetraspis*. In the closely related *C. niloticus* and *C. porosus* these genes are separated by 12 and 22 nt respectively, while in *Paleosuchus* the CO1 and tRNA-Ser genes are separated by only one nt. The start codon for the NADH4L gene in *O. tetraspis* is unconventional, TTG (Leu). Comparison with the NADH4L gene in other crocodylids

did not identify another potential start codon of the NADH4L gene in *O. tetraspis*. Similarly, the start codon in ATP6 in *O. tetraspis*, ACA (Thr), is also unconventional.

The putative tRNA-Lys genes of *P. trigonatus* and *P. palpebrosus* contain unusually large TψC loops, 17 and 24 nt, respectively. The loops consist mainly of cytosines (Fig. 3a). These TψC loops are much larger than those of *Crocodylus*, *Osteolaemus, Gavialis and Tomistoma*, which are 9-11 nt long. In *Alligator* and *Caiman* the loops are 14 and 15 nt long respectively. In the putative structure for the tRNA-Arg in *P. trigonatus* only four of the seven base pairs in the acceptor stem form standard Watson-Crick base pairs. Of the remaining pairs, two are A/C and one is an A/A mismatch (Fig. 3b). The tRNA-Arg structure of *P. palpebrosus* is more conventional, although mispairings occur also in this species (Fig. 3c).

A comparison between the protein-coding sequences in *Paleosuchus* showed that three genes are shorter in *P. palpebrosus* than in *P. trigonatus.* In NADH4 and NADH5 this is due to deletion of a single codon corresponding to aa positions 20 (NADH4) and 209 (NADH5) in *P. trigonatus*. Also, the Cyt *b* gene in *P. palpebrosus* is two aa shorter than the Cyt *b* gene in *P. trigonatus*, due to a premature TAA stop codon. The three aa positions all refer to the beginning of the respective putative genes in *P. trigonatus*.

*3.2 Phylogenetic reconstruction*

The nt and aa alignments were examined for compositional homogeneity by a 5% $\chi^2$ test as implemented in the TP program package. None of the crocodylian species deviated significantly from the expected values for compositional homogeneity. Also most of the outgroup species conformed to compositional homogeneity. Examination

of pairwise distances showed that all crocodylians had approximately the same distances to both the chicken and the *Xenopus* outgroup. This indicates that the evolutionary rates among the crocodylians are relatively homogenous, however the crocodylians are by far the fastest evolving tetrapod group in the dataset.

The sequences, both aa and nt, were analyzed by various tree reconstruction methods. The best ML aa tree from TF – the tree with the highest likelihood value under the mtREV-24 8Γ+I model of sequence evolution – is shown in Fig. 4. All ML, NJ and MP analyses were consistent with respect to the crocodylian relationships shown in this tree, however some differences were recorded among the four bird species and in the placement of the turtles. Bootstrap and other support values for the crocodylian branches were significant in all analyses; most crocodylian branches received 100% support with no branches having less than 98% support (Table 2).

The traditional crocodylian tree, with *Gavialis* basal to all other crocodylians (topology 6, Table 3), remained unsupported. The ΔlogL value for this topology in both aa and nt analyses was more than 8 s.d. worse than the best tree (Fig. 4) and the corresponding pSH values were all close to zero. Similarly, alternative placements of *Paleosuchus* within Alligatoridae (topologies 2-3, Table 3) and *Osteolaemus* within Crocodylidae (topologies 4-5, Table 3) were significantly rejected in both nt and aa analyses. Thus, a grouping of *Paleosuchus* with *Alligator* (Aggarwal et al., 1994) received no support in this analysis. The high costs in the likelihood values for altering crocodylian relationships underline the pronounced stability of the mitogenomic crocodylian tree.

Unlike the highly stable crocodylian tree, the placement of the turtles and lizards among the diapsids was less definite (Table 4). While nearly every analytical approach placed turtles as sister group to birds and crocodiles, some distance analyses

grouped turtles with birds. None of the different placements of the turtles (topologies 2-5, Table 4) were rejected by the SH tests. However, with the exception for a grouping of turtles and crocodylians on the same branch, the ΔlogL values for the alternative aa trees were more than two s.d. worse than the best tree. Analyses of 1st plus 2nd codon positions yielded less resolution than aa sequence data (Table 4) and did not allow rejection of a turtle-bird or turtle-crocodile grouping; both of these alternative trees received logL values that were only marginally smaller than that of the best tree.

*3.3 Divergence time estimates*

All divergence time estimates were highly similar, irrespective of what program (r8s, MDT or TF) and data set (aa or nt) that was used (Table 5). The estimates were also only marginally affected by the exclusion of the caiman-alligator calibration point. However, MDT estimated some divergence dates marginally younger, but well within the respective error when the caiman-alligator reference was excluded. When the caiman-alligator reference point was excluded, its age was estimated to 68-72 MYA by r8s and TF applying the remaining four calibration points to the aa dataset. MDT estimated this divergence to 65 ± 14 MYA on the basis of aa sequence, but the difference relative to the paleontologically based divergence time was not significant due to the high standard deviation. When applied to nt sequences MDT placed this divergence at 63 ± 13 MYA, which is also more recent than the suggested fossil calibration date.

Fig. 4 shows the estimated divergence times for various crocodylian divergences based on aa sequences and applying all five (one crocodylian and four non-

crocodylian) calibration points. The estimates suggest that only three of the extant crocodylian lineages survived the K/T boundary at 65 MYA, viz. the genus *Alligator*, the ancestor of the *Crocodylus*, *Osteolaemus*, *Tomistoma* and *Gavialis* clade, and the ancestor of the *Caiman* and *Paleosuchus* clade. Thus, according to these estimates all other divergences of recent crocodylians took place during or after the Paleogene. The estimates placed the divergence between *Caiman* and *Paleosuchus* at 37-41 MYA and that between the two *Paleosuchus* species at 17-19 MYA. Similarly, the estimates placed the divergence between *Crocodylus* and *Osteolaemus* at 28-29 MYA.

## 4. Discussion

### *4.1 Genomes and gene features*

The structure and gene order of the new mt genomes generally conforms to the previously published crocodylian mt genomes (Janke and Arnason 1997; Janke et al., 2001, 2005; Wu et al., 2002 unpubl.), except for some interesting exceptions in tRNA structures and protein-coding sequences. The acceptor stem of mt tRNAs is generally seven nucleotides long. Most of these base-pair with their respective nt of the 3' end of the tRNA gene sequence. However, the putative tRNA-Arg in *P. trigonatus* contains two non-Watson-Crick base pairs (both A/C) and one mismatch (A/A), giving the stem an unstable appearance (Fig. 3b). It has been shown that non-Watson-Crick G-U and A-C pairings may play an important role in aminoacetylation and translation, primarily because of the conformational flexibility they provide (McClain, 2006). Unpaired nt in stem regions may yield additional flexibility of the tRNA

structure. It is possible that tRNA-Arg in *P. trigonatus* might prove to be a new interesting example of mis- and unpaired bp. Other unusual structures in the crocodylian tRNAs involve very large TψC loops in tRNA-Lys in *Paleosuchus* (Fig. 3a). Although the sizes of these loops seem extreme, the large tRNA-Lys TψC loops appear to be common to both *Paleosuchus* and to Alligatoridae in general. However, mt tRNAs are known for their structural flexibility and atypical tRNAs are not uncommon in animal mitochondria (Wolstenholme, 1992; Steinberg and Cedergren, 1994; Qiu et al., 2005).

The frame shift caused by an extra nucleotide in the NADH5 gene in the *Osteolaemus* mt genome may be removed RNA editing. Such post-transcriptional processes involve adding, deleting or modifying one or several nts in the primary transcript, thereby enabling the function of the protein. RNA editing was described first in plant mitochondrial genomes (Hiesel et al., 1989) but it is also observed in animal mt genomes (Janke and Pääbo, 1993). The process has been suspected to restore the reading frame in the NADH3 gene in some bird mt genomes (Harlid et al., 1998).

*4.2 Phylogenetic reconstruction*

Crocodylian relationships and estimates of their divergence times based on new mitogenomic datasets were the primary aim of the current study. Regardless of analytical approach and data (aa or nt) all phylogenetic analyses yielded the same crocodylian tree (Fig. 4) with all nodes receiving strong support values. The analyses also found support for a sister group relationship between crocodylians and birds

(Aves), but a tree with the positions of birds and turtles (Chelonia) interchanged could not be refuted.

The crocodylian relationships in the best ML tree (Fig. 4) are consistent with recent molecular studies of the order (Janke et al., 2005; Harshman et al,. 2003; Gatesy et al,. 2003). However, the amount of sequence data included in the current mitogenomic study, and that of Janke et al. (2005), is considerably larger than in any of the previous studies. It is likely that this circumstance has contributed to the throughout strong support for the crocodylian nodes as evident from the values in Tables 3 and 4. Thus, all tested alternative relationships within Crocodylia had logL values more than 4.4 s.d. worse than the best ML tree. The proposed mitogenomic position of *Gavialis* as sister group to *Tomistoma* (Janke et al., 2005) remained strongly supported with the inclusion of the mitogenomic data from *Osteolaemus* and *Paleosuchus*. The position of *Gavialis* in the crocodylian tree has particular phylogenetic interest and implications as the morphological and molecular understandings of its placement have been divergent, with the morphological understanding generally placing it as the sister group to all crocodiles. The non-basal position of *Gavialis* in the crocodylian tree is in concordance with the previous, albeit less comprehensive, molecular studies of Gatesy et al. (2003) and Harshman et al. (2003). A sister group relationship between *Gavialis* and remaining crocodylians is entirely refutable in the current analyses (topology 6, Table 3).

*Osteolaemus* has commonly been considered as the sister group to *Crocodylus* within the family Crocodylidae. This understanding rested, at least for morphologists however, on a basal position of *Gavialis* in the crocodylian tree. With the established position of *Gavialis* and *Tomistoma* within the Crocodylidae, *Osteolaemus* remained the sister group to *Crocodylus*, but now to the exclusion of *Gavialis*/*Tomistoma*.

Within the Alligatoridae, the relationships between *Alligator*, *Caiman* and *Paleosuchus* have previously not been conclusively established. DNA fingerprinting analyses (Aggarwal et al., 1994) have indicated a sister group relationship between *Alligator* and *Caiman*, but the mitogenomic analyses conclusively joined *Caiman* and *Paleosuchus* to the exclusion of *Alligator*. *Melanosuchus* is the only crocodylian genus missing in the current study. Its position as the sister group to *Caiman* within the Alligatoridae has, however, not been questioned in either morphological or molecular studies (Brochu, 1997, 2003; Harshman et al., 2003; Gatesy et al., 2003).

Morphological studies place turtles among the diapsids, but their placement on the diapsid tree has not been established (Rieppel, 1999). The mitogenomic ML trees (both aa and nt) favor a position of turtles as sister group to crocodiles and birds. This relationship is also consistent with most previous molecular studies (Zardoya and Meyer, 1998; Kumazawa and Nishida, 1999; Rest et al., 2003; Iwabe et al., 2005; Janke et al., 2005). However, an aa ML tree with birds as a sister group to crocodiles and turtles (topology 2, Table 4) could not be statistically refuted in our analyses, and has indeed been proposed in some studies (Hedges and Poling, 1999; Mannen and Li, 1999). The aa ML values of a sister group relationship between turtles and birds to the exclusion of crocodiles were considerably worse than those for the other two turtle hypotheses in the aa analyses (topology 3, Table 4). It is nevertheless notable that this topology was the second best in the nt ML analyses. It is thus evident that the conclusive establishment of the phylogenetic position of turtles will need a more extensive sampling with respect to both width and the amount of data representing individual taxa.

*4.3 Divergence time estimates*

In general, the nt and aa data yielded highly similar estimates of the different divergences, both within Crocodylia and among the outgroup taxa (Table 5). The r8s, MDT and TF programs also produced highly similar divergence time estimates irrespective of in- or excluding the caiman-alligator calibration point (Müller and Reisz, 2005), indicating that the rate inferences in these programs are insensitive to this reference. Thus, the recently established caiman-alligator calibration point, 66-71 MYA, appears to be highly consistent with the non-crocodylian calibration points. As an exception, the algorithm used in the MDT software was somewhat sensitive to the inclusion/exclusion of this reference point. When the caiman-alligator calibration point was excluded, MDT estimations of crocodylian divergence times were very marginally more recent for both nt and aa data. The differences were not significant in reference to the fossil record however, due to the increased standard deviations of these estimates (node f/g, Table 5). It appears that the three different algorithms behind the respective dating program are equally effective and any one of these can equally be applied for estimating divergence times.

The general congruence among the crocodylian estimates is probably related to the relatively similar rates of molecular evolution within the group. The current estimates are more recent than those of the mitogenomic of study by Janke et al. (2005). Since the taxon sampling and the evolutionary models used here are reasonably similar to Janke et al. (2005), it is likely that the different divergence time estimates between the studies are related to the calibration points included. The choice of phylogenetically correct and narrowly defined calibration points has been shown to have crucial influence on molecular estimation of divergence times (Arnason et al., 1996, 2000;

Yang and Rannala, 2006). The ages of the non-crocodylian calibration points used in this study are somewhat more recent and more narrowly defined than those used by Janke et al. (2005) due to, and following, the recent revision of these references by Benton and Donoghue (2007). The current estimates are further constrained by the newly established caiman-alligator calibration point (Müller and Reisz, 2005), which was not available for the MDT and r8s estimates in Janke et al. (2005). Though the exclusion of this calibration point did not affect the r8s and TF datings in this study to a greater extent, it did however lead to somewhat younger divergence time estimates when the MDT software was applied.

The divergence time estimates obtained in the current study suggest that only three of the extant crocodylian lineages survived the K/T extinction approximately 65 MYA, viz. Alligatorinae, Caimaninae and Crocodylidae. This result is consistent with the paleontological conclusions of Brochu (2003).

Müller and Reisz (2005) placed the alligator-caiman divergence at 66-71 MYA. The molecular estimates applying the non-crocodylian calibration points placed this divergence at 68-72 MYA, suggesting pronounced coherence between the molecular estimates and the crocodylian fossil record. The molecular estimates placed the *Paleosuchus-Caiman* divergence at 37-41 MYA (late Eocene). The estimate is much more recent than the divergence time, ≈ 58 MYA, deduced from Fig. 5 in Brochu (2003) based on the age of *Orthogenysuchus olseni*. The nature of this discrepancy is unusual, as molecular estimates rather tend to become placed earlier than those suggested by the fossil record. Considering that the molecular datings in this study are congruent with both the intra-crocodylian reference point suggested by Müller and Reisz (2005) and other parts of the crocodylian fossil record, it is possible that the *Caiman-Paleosuchus* split has been misidentified in the fossil record. This is not

entirely unreasonable given the previous problems associated with the morphological identification of the phylogenetic position of *Gavialis*. The split between *P. trigonatus* and *P. palpebrosus* was placed at approximately 17-19 MYA. However, the accuracy of this estimate cannot be paleontologically evaluated due to the lack of *Paleosuchus* fossils.

The molecular estimates placed the *Osteolaemus-Crocodylus* divergence at 28-29 MYA (Oligocene), a dating that is consistent with Brochu (2003, Fig. 7).

In conclusion, the mitogenomic analyses yielded conclusive support to all nodes in the crocodylian tree and the molecular estimates showed a general consistency with the crocodylian fossil record. The phylogenetic tree, in conjunction with the molecular estimates of the different crocodylian divergences, suggests that only three of the extant crocodylian lineages survived the K/T boundary. This may suggest that the crocodylians were severely affected by this transition. The low number of recent crocodylian species and the temporally wide span of their divergences also suggest that their evolution has been at an entirely different scale than that of the mammals as judged from their flourishing evolution after the same transition.

**Acknowledgements**

This study was supported by the Royal Physiographic Society in Lund (Nilsson-Ehle). We are very grateful to Ulfur Arnason, Morgan Kullberg, Björn Hallström and Maria A. Nilsson for comments on the manuscript.

**References**

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

Arnason, U., Gullberg, A., Janke, A. and Xu, X., 1996. Pattern and Timing of evolutionary divergencies among Hominoids based on analysis of complete mtDNAs. J. Mol. Evol. 43, 650-661.

Arnason, U., Gullberg, A., Burguette, A. S., Janke, A., 2000. Molecular estimates of primate divergences and new hypotheses for primate dispersal and the origin of modern humans. Hereditas 133, 217-228.

Aggarwal, R. K., Majumdar, K. C., Lang, J. W., Singh. L., 1994. Generic affinities among crocodilians as revealed by DNA fingerprinting with a Bkm-derived probe. Proc. Natl. Acad. Sci. USA 91, 10601-10605.

Benton, M. J., Donoghue, P.C., 2007. Paleontological evidence to date the tree of life. Mol. Biol. Evol. 24, 26-53.

Brochu, C. A., 2003. Phylogenetic approaches toward crocodylian history. Annu. Rev. Earth Planet. Sci. 31, 357-397.

Brochu, C. A., 2001. Crocodylian snouts in space and time: phylogenetic approaches toward adaptive radiation. Amer. Zool. 41, 564-585.

Brochu, C. A., 1997. Morphology, fossils, divergence timing, and the phylogenetic relationships of *Gavialis*. Syst. Biol. 46(3), 479-522.

Felsenstein, J., 1993. Phylogenetic Inference Programs (PHYLIP). University of Washington, Seattle, and University Herbarium, University of California, Berkeley.

Gatesy, J., Amato, G., Norell, M., DeSalle, R., Hayashi, C., 2003. Combined support for taxic wholesale taxic atavism in gavialine crocodilians. Syst. Biol. 52(3), 403-422.

Gatesy, J., DeSalle, R., Wheeler, W. C., 1993. Alignment-ambiguous nucleotide site and the exclusion of systematic data. Mol. Phyl. Evol. 2, 152–157.

Harlid, A., Janke, A., Arnason, U., 1998. The complete mitochondrial genome of *Rhea americana* and early avian divergences. J. Mol. Evol. 46, 669-679.

Harshman, J., Huddleston, C. J., Bollback, J. P., Parsons, T. J., Braun, M. J., 2003. True and false gharials: a nuclear gene phylogeny of crocodylia. Syst. Biol. 52, 386–402-

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

Hiesel, R., Wissinger, B., Schuster, W., Brennicke, A., 1989. RNA editing in plant mitochondria. Science 246, 1632-1634.

Iwabe, N., Hara, Y., Kumazawa, Y., Shibamoto, K., Saito, Y., Miyata, T., Katoh, K., 2005. Sister group relationship of turtles to the bird-crocodilian clade revealed by nuclear DNA-coded proteins. Mol. Biol. Evol. 22(4), 810-813.

Janke, A., Gullberg, A., Hughes, S., Aggarwal, R. K., Arnason, U., 2005. Mitogenomic analyses place the gharial (*Gavialis gangeticus*) on the crocodile tree and provide pre-K/T divergence times for most crocodilians. J. Mol. Evol. 61(5), 620-626.

Janke, A., Erpenbeck, D., Nilsson, M., Arnason, U., 2001. The mitochondrial genomes of a lizard (*Iguana iguana*) and the caiman (*Caiman crocodylus*): implications for amniote phylogeny. Proc. Roy. Soc. Lond. B 268, 623–631.

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., Pääbo S., 1993. Editing of a tRNA anticodon in marsupial mitochondria changes its codon recognition. Nucleic Acids Res. 21(7):1523-5.

Jobb, G., von Haeseler, A., Strimmer, K., 2004. TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. BMC Evol. Biol. 4, 18.

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(6), 784-792.

Lanave, C., Preparata, G., Saccone, C., Serio, G., 1984. A new method for calculating evolutionary substitution rates. J. Mol. Evol. 20(1), 86-93.

Mannen, H., Li, S. S-L., 1999. Molecular evidence for a clade of turtles. Mol. Phyl. Evol 13, 144-148.

McAliley, L. R., Willis, R. E., Ray, D. A., White, P. S., Brochu, C. A., Densmore, L. D., 2006. Are crocodiles really monophyletic? – Evidence for subdivisions from sequence and morphological data. Mol. Phyl. and Ev. 39, 16-32.

McClain, W. H., 2006. Surprising contribution to aminoacylation and translation of non-Watson-Crick pairs in tRNA. Proc. Natl. Acad. Sci. U S A 103(12), 4570-4775.

Müller, J., Reisz, R. R., 2005. Four well-constrained calibration points from the vertebrate fossil record for molecular clock estimates. BioEssays 27, 1069-1075.

Nilsson, M. A., Arnason, U., Spencer, P. B., Janke, A., 2004. Marsupial relationships and a timeline for marsupial radiation in South Gondwana. Gene 340, 189-196.

Posada, D., Crandall, K. A., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817-818.

Qiu, Y., Song, D., Zhou, K., Sun, H., 2005. The mitochondrial sequences of *Heptathela hangzhouensis* and *Ornithoctonus huwena* reveal unique gene arrangements and atypical tRNAs. J. Mol. Evol. 60(1), 57-71.

Rest, J. S., Ast, C. A., Austin, C. C., Wadell, P. J., Tibbetts, E. A., Hay, J. M., Mindell, D. P., 2003. Molecular systematics of primary reptilian lineages and the tuatara mitochondrial genome. Mol. Phyl. Evol. 29, 289-297.

Rieppel, O., 1999. Turtles as diapsid reptiles, Zoologica Scripta 29, 199-212.

Salisbury, S. W., Willis, P. M. A., 1996. A new crocodylian from the Early Eocene of south-eastern Queensland and a preliminary investigation of the phylogenetic relationships of crocodyloids. Alcheringa 20, 179–227.

Sambrook, J., Russell, D. W., 2001. Molecular cloning, a laboratory manual. New York: Cold Spring Harbor Press.

Sanderson, M. J., 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Mol. Biol. Evol. 19, 101-109.

Schmidt, H. A., Strimmer, K., Vingron, M., von Haeseler, A., 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics.* 18:502-504.

Schmitz, A., Mansfeld, P., Hekkala, E., Shine, T., Nickel, H., Amato, G., Bohme, W., 2003. Molecular evidence for species level divergence in African Nile crocodiles *Crocodylus niloticus* (Laurenti, 1786). Comp. Rendus Palevol. 2, 703Ð712.

Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16, 1114-1116.

Steinberg, S., Cedergren, R., 1994. Structural compensation in atypical mitochondrial tRNAs. Structural Biology 1, 507-510.

Swofford, D. L. 2003. PAUP* Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

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.

Thorne, J. L., Kishino, H., 2002. Divergence time and evolutionary rate estimation with multilocus data. Syst. Biol. 51, 689-702.

White, P. S., Densmore, L. D., 2000. DNA sequence alignments and data analysis methods: their effect on the recovery of crocodylian relationships. In: Grigg, G., Seebacher, F., Franklin, C. E. (Eds.), Crocodilian Biology and Evolution. Surrey Beatty and Sons, Chipping Norton, Sydney, pp. 29-37.

Willis, R. E., McAliley, L. R., Neeley, E. D., Densmore III, L. D., 2007. Evidence for placing the False Gharial (*Tomistoma schlegelii*) into the family Gavialidae; Inferences from Nuclear Gene Sequences, Mol. Phyl. Evol. (2007), doi: 10.1016/j.ympev.2007.02.005

Wolstenholme, D. R., 1992. Animal mitochondrial DNA structure and evolution. In: Wolstenholme, D. R., Jeon, K. W. (Eds.), Int. Rev. Cytology Vol. 141. Academic Press, San Diego, USA, pp. 173-215.

Wu, X., Wang, Y., Zhou, K., Zhu, W., Nie, J., Wan, C., Xie, W., 2002. The complete mitochondrial genome sequence of Chinese alligator, AF511507 (unpubl.).

Yang, Z., 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. J. Mol. Evol. 39, 306-314.

Yang, Z., Rannala, B., 2006. Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. Mol. Biol. Evol. 23, 212-226.

Zardoya, R., Meyer, A., 1998. Complete mitochondrial genome suggests diapsid affinities of turtles. Proc. Natl. Acad. Sci. U S A. 95, 14226-14231.

**Figure texts**

Fig. 1. Schematic view over crocodylian relationships. In the traditional morphological tree (a) *Gavialis* is basal to all other extant crocodylians, whereas in the molecular tree (b) *Gavialis* is the sister taxon to *Tomistoma*.

Fig. 2. The genetic map of the *Osteolaemus tetraspis* mitochondrial genome.

Fig. 3. Putative tRNA structures in *P. palpebrosus* (a and c) and *P. trigonatus* (b).

Fig. 4. Maximum likelihood tree based on aa sequence data and the mtREV24 8Γ+I model of sequence evolution. The estimated divergences (MYA) indicate the range of the values from r8s, MDT and TF (open triangles). Black triangles indicate the range of calibration dates (MYA).

**Tables**

Table 1

Accession numbers for the mt genomes that were included in the phylogenetic analysis.

| **Common name** | **Scientific name** | **Accession no.** |
|---|---|---|
| False gharial | *Tomistoma schlegelii* | AJ810455 |
| Gharial | *Gavialis gangeticus* | AJ810454 |
| Estuarine crocodile | *Crocodylus porosus* | AJ810453 |
| Nile crocodile | *Crocodylus niloticus* | AJ810452 |
| American alligator | *Alligator mississippiensis* | Y13113 |
| Dwarf crocodile | *Osteolaemus tetraspis* | AM493868 |
| Chinese alligator | *Alligator sinensis* | AF511507 |
| Spectacled caiman | *Caiman crocodylus* | AJ404872 |
| Common iguana | *Iguana iguana* | AJ278511 |
| Smooth-fronted caiman | *Paleosuchus trigonatus* | AM493869 |
| Cuvier's dwarf caiman | *Paleosuchus palpebrosus* | AM493870 |
| Mole skink | *Eumeces egregius* | AB016606 |
| Green turtle | *Chelonia mydas* | AB012104 |
| Painted turtle | *Chrysemys picta* | AF069423 |
| Rook | *Corvus frugilegus* | Y18522 |
| Falcon | *Falco peregrinus* | AF090338 |
| Rockhopper penguin | *Eudyptes chrysocome* | AP009189 |

| | | |
|---|---|---|
| Chicken | *Gallus gallus* | X52392 |
| Cow | *Bos taurus* | V00654 |
| Opossum | *Didelphis virginiana* | Z29573 |
| Wallaroo | *Macropus robustus* | Y10524 |
| Mouse | *Mus musculus* | J01420 |
| African clawed frog | *Xenopus laevis* | M10217 |

Table 2

Bootstrap and support values for branches within Crocodylia.

| | MP | | NJ | | ML (TF) | | ML (TP) | |
|---|---|---|---|---|---|---|---|---|
| **Branch** | **nt** | **aa** | **nt** | **aa** | **nt** | **aa** | **nt** | **aa** |
| a | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| b | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| c | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| d | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| e | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| f | 100 | 100 | 100 | 100 | 98 | 100 | 100 | 100 |
| g | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| h | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| i | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

NOTE – Branches (a-i) refer to those in Fig. 4.

Table 3

ML analysis of alternative crocodylian relationships based on aa and nt sequence data.

| | Topology | Amino acid data | | | | Nucleotide data | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Rate heterogeneity | | Rate homogeneity | | Rate heterogeneity | | Rate homogeneity | |
| | | ΔlogL / s.d. | pSH | ΔlogL / s.d. | pSH | ΔlogL / s.d. | pSH | ΔlogL / s.d. | pSH |
| 1 | Fig. 4 | <-51,674> | 1 | <-54,675> | 1 | <- 44,494> | 1 | <-47,633> | 1 |
| 2 | OG,(((All,Cai),Pal),((Gav,Tom),(Ost,Cro))) | -197 / ±27.2 | <0.001 | -247 / ±31.6 | <0.001 | -169 / ±25.0 | <0.001 | -228 / ±31.1 | <0.001 |
| 3 | OG,(((All,Pal),Cai),((Gav,Tom),(Ost,Cro))) | -197 / ±27.4 | <0.001 | -242 / ±32.3 | <0.001 | -168 / ±25.2 | <0.001 | -218 / ±32.2 | <0.001 |
| 4 | OG,((All,(Cai,Pal)),(Ost,((Gav,Tom),Cro))) | -117 / ±21.7 | <0.001 | -150 / ±26.5 | <0.001 | -84.6 / ±19.1 | 0.012 | -117 / ±25.8 | 0.024 |
| 5 | OG,(((All,Pal),Cai),(((Gav,Tom),Ost),Cro)) | -119 / ±21.4 | <0.001 | -155 / ±26.1 | <0.001 | -89.1 / ±18.6 | 0.009 | -130 / ±24.6 | 0.010 |
| 6 | OG,(Gav,((All,(Cai,Pal)),(Tom,(Ost,Cro)))) | -326 / ±34.2 | <0.001 | -501 / ±44.7 | <0.001 | -266 / ±30.0 | <0.001 | -482 / ±44.2 | <0.001 |

NOTE - All: *Alligator*, Cai: *Caiman*, Cro: *Crocodylus*, Pal: *Paleosuchus*, Ost: *Osteolaemus*, Gav: *Gavialis*, Tom: *Tomistoma*, OG: outgroup, ΔlogL: difference in log-likelihood value relative to the best log-likelihood value shown in angle brackets, s.d.: standard deviation.

Table 4

ML analysis of alternate positions of turtles and lizards, based on aa and nt sequence data.

| | | Amino acid data | | | | Nucleotide data | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Topology** | **Rate heterogeneity** | | **Rate homogeneity** | | **Rate heterogeneity** | | **Rate homogeneity** | |
| | | **ΔlogL / s.d.** | **pSH** | **ΔlogL / s.d.** | **pSH** | **ΔlogL / s.d.** | **pSH** | **ΔlogL / s.d.** | **pSH** |
| 1 | Fig. 4 | <-51,674> | 1 | <-54,675> | 1 | <- 44,494> | 1 | <-47,633> | 1 |
| 2 | OG,((bir,(tur,cro)),liz) | -13.4 / ±10.6 | 0.700 | -12.0 / ± 17.2 | 0.769 | -2.70 / ±7.99 | 0.947 | -7.66 / ±17.5 | 0.899 |
| 3 | OG,((cro,(tur,bir)),liz) | -22.0 / ±9.00 | 0.494 | -35.2 / ±14.1 | 0.375 | -1.22 / ±8.39 | 0.956 | -2.03 / ±18.3 | 0.939 |
| 4 | OG,((bir,cro),(tur,liz)) | -42.4 / ±14.1 | 0.157 | -64.3 / ±19.2 | 0.089 | -37.3 / ±12.6 | 0.248 | -58.4 / ±18.8 | 0.241 |
| 5 | OG,(tur,(liz,(bir,cro))) | -45.6 / ±13.5 | 0.124 | -69.4 / ±18.5 | 0.056 | -39.0 / ±12.2 | 0.223 | -63.2 / ±18.1 | 0.184 |

NOTE - tur: turtles, cro: Crocodylia, bir: birds, liz: lizards, OG: outgroup, ΔlogL: difference in log-likelihood value relative to the best log-likelihood value shown in angle brackets, s.d.: standard deviation.

Table 5

Estimated crocodylian divergence times (MYA) and their standard deviations.

| | r8s | | | | MDT | | | | TF | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | All points | | W/o caim-allig | | All points | | W/o caim-allig | | All points | | W/o caim-allig | |
| **Node** | **aa** | **Nt** | **aa** | **nt** | **aa** | **nt** | **aa** | **nt** | **Aa** | **nt** | **aa** | **nt** |
| a/bird | 250±0.59 | 250±0.0 | 250±0 | 250±0.0 | 240±3.8 | 242±4.2 | 240±3.8 | 241±4.1 | 250±2.2 | 250±3.1 | 250±0.0 | 250±0.0 |
| b/c | 101±3.0 | 102±2.8 | 101±3.6 | 102±3.8 | 97±4.8 | 96±4.8 | 93±15 | 90±14 | 103±3.2 | 104±3.5 | 105±3.5 | 106±3.7 |
| d/e | 47±2.4 | 48±2.3 | 47±2.4 | 48±2.3 | 49±5.7 | 50±5.9 | 47±13 | 46±12 | 49±3.4 | 50±3.8 | 50±2.9 | 51±3.4 |
| f/g | 68±2.0 | 69±1.6 | 68±3.7 | 69±3.7 | 68±1.4 | 68±1.4 | 65±14 | 63±13 | 71±1.4 | 71±1.3 | 72±3.7 | 73±4.1 |
| h/Ote | 28±1.6 | 29±1.6 | 28±1.7 | 29±1.6 | 28±4.9 | 30±5.2 | 27±8.6 | 28±8.4 | 28±2.5 | 30±2.8 | 28±2.1 | 31±2.7 |
| i/Ccr | 40±1.8 | 40±1.7 | 40±2.5 | 40±2.3 | 37±3.7 | 37±3.7 | 36±9.6 | 34±8.8 | 41±1.7 | 40±1.9 | 41±2.5 | 41±2.9 |
| Gga/Tsc | 22±1.7 | 24±1.9 | 22±1.7 | 24±1.8 | 26±4.9 | 28±5.1 | 25±8.5 | 26±8.2 | 25±2.8 | 27±3.1 | 25±2.8 | 27±3.1 |
| Ami/Asi | 47±2.7 | 50±3.0 | 47±3.3 | 50±3.8 | 52±3.3 | 55±3.2 | 50±12 | 50±12 | 53±3.0 | 55±3.0 | 54±4.3 | 56±4.1 |
| Cni/Cpo | 11±1.0 | 13±1.2 | 11±1.0 | 13±1.2 | 12±2.9 | 15±3.5 | 12±4.6 | 14±5.0 | 12±1.5 | 14±2.1 | 12±1.4 | 15±2.1 |
| Ptr/Ppa | 17±1.3 | 18±1.7 | 17±1.5 | 18±1.2 | 18±3.1 | 18±3.1 | 17±5.9 | 16±5.3 | 19±1.7 | 18±1.7 | 19±2.1 | 19±2.2 |

NOTE – Divergences estimates were based on both inclusion ("All points") and exclusion ("W/o caim-allig") of the caiman-alligator calibration point. Ote: *O. tetraspis*, Ccr: *C. crocodylus*, Gga: *G. gangeticus*, Tsc: *T. schlegelii*, Ami: *A. mississippiensis*, Asi: *A. sinensis*, Cni: *C. niloticus*, Cpo: *C. porosus*, Ptr: *P. trigonatus*, Ppa: *P. palpebrosus*. The node letters (a-i) refer to the branches shown in Fig. 4.