

## Mitogenomic Analyses Place the Gharial (*Gavialis gangeticus*) on the Crocodile Tree and Provide Pre-K/T Divergence Times for Most Crocodylians

Axel Janke,<sup>1</sup> Anette Gullberg,<sup>1</sup> Sandrine Hughes,<sup>2,3</sup> Ramesh K. Aggarwal,<sup>4</sup> Ulfur Arnason<sup>1</sup>

<sup>1</sup> Department of Cell and Organism Biology, Division of Evolutionary Molecular Systematics, University of Lund, Sölvegatan 29, S-223 62 Lund, Sweden

<sup>2</sup> Centre de Génétique Moléculaire et Cellulaire, CNRS UMR 5534, Université Claude Bernard Lyon 1, Bât. G. Mendel, 16 rue Raphaël Dubois, 69622 Villeurbanne Cedex, France

<sup>3</sup> Laboratoire d'Anthropologie des Populations du Passé, CNRS UMR 5199 PACEA, Université Bordeaux 1, Bât. B8 avenue des Facultés, 33405 Talence Cedex, France

<sup>4</sup> Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad 500 007, India

Received: 26 November 2004 / Accepted: 14 June 2005 [Reviewing Editor: Dr. Nicolas Galtier]

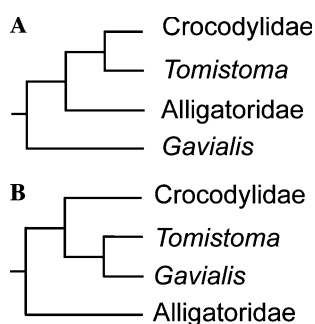
**Abstract.** Based on morphological analyses, extant members of the order Crocodylia are divided into three families, Alligatoridae, Crocodylidae, and Gavialidae. Gavialidae includes one species, the gharial, *Gavialis gangeticus*. In this study we have examined crocodylian relationships in phylogenetic analyses of seven mitochondrial genomes that have been sequenced in their entirety. The analyses did not support the morphologically acknowledged separate position of the gharial in the crocodylian tree. Instead the gharial joined the false gharial (*Tomistoma schlegelii*) on a common branch that was shown to constitute a sister group to traditional Crocodylidae (less *Tomistoma*). Thus, the analyses suggest the recognition of only two Crocodylia families, Alligatoridae and Crocodylidae, with the latter encompassing traditional Crocodylidae plus *Gavialis/Tomistoma*. A molecular dating of the divergence between Alligatoridae and Crocodylidae suggests that this basal split among recent crocodylians took place  $\approx$ 140 million years before present, at the Jurassic/Cretaceous boundary. The results suggest that at least five crocodylian lineages survived the mass extinction at the KT boundary.

**Key words:** Crocodylia — Molecular dating — Mitochondrial DNA — Phylogeny — Reptiles

### Introduction

Recent crocodylians constitute a small order, Crocodylia, within class Reptilia. The order includes only 23 living species. One might be inclined to think that this low number of species would make it easy to reconstruct relationships among living crocodylians. This has not been the case, however, as the morphological evolution of the body shape within this group has been slow. Hence the number of diagnostic morphological characters useful for phylogenetic analysis is limited. Poe (1996) used about 160 characters to examine the crocodylian relationships. However, many are not found in all lineages. Only 12 cranial characters were used when a noncrocodylian outgroup was included in an analysis by Norell (1989). Three primary families are commonly recognized within the Crocodylia: the Alligatoridae (which includes the genera *Alligator*, *Caiman*, *Melanosuchus*, and *Paleosuchus*), the Crocodylidae (*Crocodylus*, *Osteolaemus*, *Tomistoma*), and the Gavialidae, which includes only one species, the gharial, *Gavialis gangeticus*.

The placement of the gharial on the crocodylian tree and its relation to the false gharial, *Tomistoma schlegelii*, have attracted considerable interest, as phylogenies based on morphological and molecular data are at odds. Morphological studies (e.g., Norell 1989; Tarsitano et al. 1989; Poe 1996) have favored a placement of the two species in different families, the



**Fig. 1.** Schematic view of crocodylian relationships based on (a) morphological and (b) molecular data.

gharial in the Gavialidae and the false gharial in the Crocodylidae. In this scheme the gharial is placed basal to all other recent crocodiles (Fig. 1a). In contrast to this, most molecular studies have recognized *Tomistoma* and *Gavialis* as sister groups and placed *Gavialis/Tomistoma* as the sister group of the Crocodylidae (Fig. 1b). This molecular relationship was originally identified by Densmore (1983) in an extensive immunological study that included all recent crocodylian species. Other molecular studies, such as RFLP data (Densmore and White 1991), fingerprint analysis (Aggarwal et al. 1994), and analyses of a portion (267 nt including gaps) of the mitochondrial (mt) 12S rRNA gene (Gatesy et al. 1993) have suggested the same relationships. However, some of the molecular studies (e.g., Gatesy et al. 1993; Hass et al. 1992; Poe 1996) did not include an outgroup, were not statistically testable (Densmore and White 1991), or used problematic methods like UPGMA for tree construction (Aggarwal et al. 1994).

Crocodylian relationships were addressed more recently in two molecular studies. The study by Harshman et al. (2003) was based on a part of the *c-myc* nuclear gene, while that by Gatesy et al. (2003) used a combination of the RAG-1 nuclear sequence plus about 800 nt from portions of the mt 12S rRNA, 16S rRNA, and cytochrome *b* genes. The nuclear sequences joined *Gavialis* and *Tomistoma* but the position of the branch was not conclusively settled. An analysis of the mt data set supported the tree shown in Fig. 1b. A “supermatrix” including some additional nuclear gene sequences, but lacking outgroup taxa to the crocodiles for most datasets, found some support for a *Gavialis/Tomistoma* sister-group relationship (Gatesy et al. 2004). However, without an outgroup the phylogenetic position of the gharial could not be conclusively determined.

Poe (1996) and Brochu (1997) examined the discrepancy between the morphological and the molecular findings related to crocodylian relationships. However, this examination did not take into account the generally limited amount of sequence data available at that time. Thus, the use of only  $\approx 250$  nt of the 12S rDNA gene may have affected the statistical

reliability of the analyses. Although the order Crocodylia is a very small group, the amount of sequence data used to examine their internal relationships after including an outgroup has hitherto been limited to partial mitochondrial genes (240–450 nt) or highly conserved nuclear genes (e.g., RAG-1), which may contain insufficient phylogenetic information.

In order to increase the amount of molecular data suitable for phylogenetic analysis and dating of crocodylian relationships, we have sequenced the complete mt genomes from four species: *Crocodylus niloticus* (Nile crocodile), *Crocodylus porosus* (estuarine crocodile), *Gavialis gangeticus* (gharial), and *Tomistoma schlegelii* (false gharial). This sampling together with the previously published alligatorid genomes (Janke and Arnason 1997; Janke et al. 2001; Wu et al. unpublished; accession no. AF511507) allows examination of basal divergences within the order Crocodylia, including the phylogenetic position of *Gavialis*. In addition, we provide molecular estimates of all major divergences within the Crocodylia. This issue was addressed by Brochu (1997), who concluded inter alia that Crocodylidae and Alligatoridae had diverged in Late Cretaceous. This proposal needs further examination, however, as the tree underlying Brochu’s (1997) analysis is inconsistent with the molecular tree found, for example, by Densmore (1983).

## Materials and Methods

Whole genomic crocodylian DNA was used to PCR amplify two large mt fragments (9 and 7 kb, respectively) using conserved primer pairs (Table 1) with TAKARA LA-Taq polymerase. The 9-kb fragment spans from the 12S rRNA gene to tRNA-Gly, while the 7-kb fragment spans from tRNA-Gly to tRNA-Phe. Specific primers were designed to amplify the control region and the tRNA-Gly region using TAKARA Ex-Taq polymerase. The mt genomes were sequenced with the BIG-DYE version 3 cycle sequencing kit on an ABI Prism 3100 Genetic Analyzer using primer walking.

The phylogenetic analysis is based on the 12 H-strand encoded protein-coding genes of the newly sequenced mt genomes of *Crocodylus niloticus* (Nile crocodile; accession no. AJ810452), *Crocodylus porosus* (estuarine crocodile; AJ810453), *Gavialis gangeticus* (gharial; AJ810454), and *Tomistoma schlegelii* (false gharial; AJ810455). The complete mt genomes of the Nile crocodile, estuarine crocodile, and gharial were sequenced, while the sequences of a part of the control region of the false gharial remained undetermined. The new genomes were aligned to the 12 H-strand encoded protein-coding genes of *Caiman crocodylus* (caiman; AJ404872), *Alligator mississippiensis* (alligator; Y13113), *Alligator sinensis* (Chinese alligator; AF511507), *Iguana iguana* (common iguana; AJ278511), *Eumeces egregius* (mole skink; AB016606), *Chelonia mydas* (green turtle; AB012104), *Chrysemys picta* (painted turtle; NC\_002073), *Corvus frugilegus* (rook; Y18522), *Falco peregrinus* (falcon; AF090338), *Struthio camelus* (ostrich; Y12025), and *Gallus gallus* (chicken; X52392). The alignment included, in addition, four mammals: *Bos taurus* (cow; V00654), *Didelphis virginiana* (opossum; Z29573), *Macropus robustus* (wallaroo; Y10524), and *Mus musculus* (mouse; J01420). The amniote tree was rooted with the amphibian, *Xenopus laevis*

**Table 1.** Conserved primers that were used to amplify crocodilian mtDNA

Location	Name/gene	Sequence 5'-3'
457	CroL12SRNA	GGGATTAGATACCCCACTAT
4,990	CroLtTrp	AAGCCAAGGGCCTTCAAAG
9,435	CroHtGly	GGGTTTAATGATTGGAAGT
9,443	CroLtGly	AATACAAATGACTTCCAAT
15,620	CroHtPhe	CCATGTAAACATTTTCAG
15,452	CroHtThr	CCAYYTCTGTCTTACAAGG

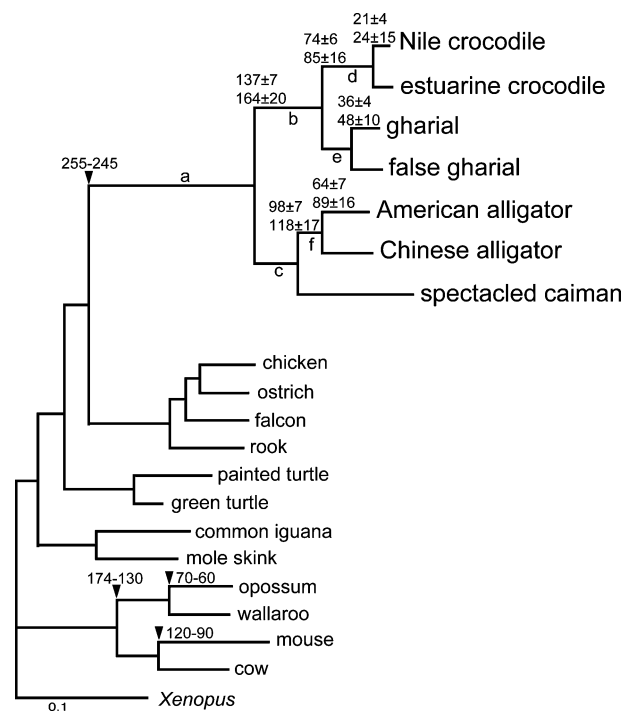
*Note.* Locations refer to the 3' end of the primer in the American alligator mt genome (accession no. Y13113). H and L denote the orientation of the primer according to the H and L strand of the mt genome, and 12S rRNA, tTrp, etc., refer to the 12S rRNA gene or the respective tRNA gene where the primer is located.

(African clawed frog; NC\_001573). The alignment was inspected manually. After excluding gaps and alignment-ambiguous sites around gaps, 9489 nt sites remained for phylogenetic analysis. The third codon positions were excluded from the analysis of nt sequences.

Phylogenetic relationships were analyzed using the TREE-PUZZLE (Strimmer and von Haeseler 1996), PHYLIP (Felsenstein 1993), MOLPHY (Adachi and Hasegawa 1996), MrBayes (Huelsenbeck and Ronquist 2001), and PAL/VANILLA (Drummond and Strimmer 2001) program packages. The mtREV-24 model of amino acid (aa) sequence evolution (Adachi and Hasegawa 1996) and the TN-93 model of nt evolution (Tamura and Nei 1993) were used for distance and likelihood analyses. Although more parameter-rich models fitted the nt sequence data better, the TN-93 model is the most parameter-rich model that is common to most phylogenetic analysis packages. Parameter estimation was according to the software, using the nt/aa frequencies of the data set. For nt sequences the transition/transversion parameter was estimated to 1.48, and the pyrimidine/purine parameter to 1.42. The analyses were performed under the assumptions of both rate homogeneity and rate heterogeneity among sites, the latter with a  $\Gamma$  model with five classes of variable sites. The rate heterogeneity parameter  $\alpha$  was 0.30 for aa and 0.29 for nt sequences. The SH test (Shimodaira and Hasegawa 1999), likelihood values, standard deviations, and minimum number of substitutions and their standard deviations were used for comparison of alternative trees relative to the best ML tree. A  $\chi^2$  test for compositional homogeneity as implemented in the TREE-PUZZLE program was used to test for the stationarity of the nt/aa composition. Distance tables for the distance analyses were calculated by the TREE-PUZZLE program with the above-mentioned parameters and analyzed by neighbor joining as implemented in the PHYLIP program package.

Divergence times were estimated from distances by taking different evolutionary rates into consideration (Arnason et al. 2000). Divergence times were also estimated on a ML tree based on aa sequences using a penalized likelihood or a nonparametric rate smoothing method as implemented in the r8s program (Sanderson 2002). The ML tree and branch lengths correspond to the tree in shown in Fig 2.

The estimation of divergence times was also performed on first and second codon positions by multidivtime as implemented in the T3 program package (<http://abacus.gene.ucl.ac.uk/>). The dating was done on nt data, because model parameters for mt aa sequence data are not available in the current version. The parameters as estimated by TREE-PUZZLE are A = 24.0%, C = 27.3%, G = 17.3%, and T = 31.4%; the  $\Gamma$  distribution parameter  $\alpha$  = 0.29; and the relative rates ( $r$ ) between five categories of variable sites are 0.0014, 0.0529, 0.3008, 1.0483, and 3.5966, respectively, each with equal probability (0.2). These parameters were used to build the model file (modelinf).



**Fig. 2.** ML tree based on aa sequences showing references (arrowheads) used for estimating crocodilian divergence times (MYBP) using the programs r8s (upper value) and multidivtime (lower value). Support values for branches (a-f) are shown in Table 2.

The divergence times have been based on four well-established vertebrate calibration points. For the origin of the Crocodylia a divergence of 255–245 MYBP was applied. The minimum age of 245 MYBP represents the split between Crocodylotarsi and Ornithosuchia and the maximum age of 255 MYBP represents the origin of Archosauromorpha (Benton 1990). Thus this divergence time has a relatively narrow lower and upper bound. The age of this calibration point is also consistent with other vertebrate divergences (Janke and Arnason 1997).

The other lower and upper bound calibration points used were the split of marsupials and eutherians (174–130 MYBP), the split between mouse and cow (120–90 MYBP), and the split between South American and Australian marsupials (70–60 MYBP). These dates are consistent with previous estimates and the fossil record. The youngest eutherian fossil (*Eomaia*) is about 125 myr old (Ji et al. 2002), suggesting a lower bound of about 130 MYBP for the split between marsupials and eutherians. An upper bound of 170 MYBP for this split is indicated by some molecular estimates (e.g., Kumar and Hedges 1998). The split between South American and Australian marsupials is not expected to be older than 70 MYBP or younger than 60 MYBP (Nilsson et al. 2004). The divergence time between rodents and artiodactyls cannot be older than *Eomaia*, the oldest eutherian (Ji et al. 2002), or younger than 84 MYBP, as eutherians began to diversify around this time (Archibald et al. 2001).

## Results

It has been shown previously (Quinn and Mindell 1996; Janke and Arnason 1997; Janke et al. 2001) that the organization of crocodilian mt genomes differs slightly from that characterizing most gnat-

**Table 2.** Support for crocodilian branches

Branch	MP	NJ	FM	LBP	QP	QP <sub>rh</sub>	BP	BPrh
a	100	100	100	100	100	100	1.0	1.0
b	100	100	100	100	100	100	1.0	1.0
c	100	100	99	100	100	100	1.0	1.0
d	100	100	100	100	100	100	1.0	1.0
e	100	100	100	100	100	100	1.0	1.0
f	100	100	100	100	100	100	1.0	1.0

*Note.* MP, maximum parsimony; NJ, neighbor joining; FM, Fitch–Margoliash with least squares; LBP, local bootstrap probability; QP, quartet puzzling; rh, rate heterogeneity; BP, Bayesian probability. The values show the respective bootstrap support for the MP, NJ, FM, and LBP methods, the quartet puzzle support for the QP and QP<sub>rh</sub> methods, and the probability for the BP and BPrh analysis for branches a–f in Fig. 2.

hostomes. Consistent with the previous findings, the new mt genomes show a rearranged gene order for some of their tRNA genes. Thus the tRNA-Phe gene is located in 5' of the control region, forming a cluster with the tRNA-Pro and tRNA-Thr genes. Also, the tRNA-Ser(AGY) and tRNA-His genes have a different position compared to that in other vertebrates, forming the cluster tRNA-Ser(AGY), tRNA-His, tRNA-Leu(CUN) instead of the more common arrangement among vertebrates, tRNA-His, tRNA-Ser(AGY), tRNA-Leu(CUN).

The mitochondrial control region contains large regions that are surprisingly conserved among crocodiles compared to the control region of other vertebrates. This characteristic needs to be examined further. Some crocodilian mt tRNA and protein-coding genes are separated by noncoding regions that are larger than in any other vertebrate group studied to date. This is particularly the case where tRNA genes have moved to new positions. Thus, the noncoding regions between genes for the tRNA-Thr and tRNA-Pro, and the tRNA-Pro and tRNA-Phe, are 11–20 nt long. As the rearrangement of tRNA genes is common to all crocodiles, this event must have taken place in the ancestor of all recent crocodiles, conceivably some time in the Jurassic (208–146 MYBP) or earlier. The presence of large noncoding sequences contradicts the idea that such sequences should erode quickly in order to maintain the mt genome as compact as possible.

The alignment of the crocodilian mt protein coding sequences was unproblematic, except for a small sequence region at the 3' region of the cytochrome oxidase subunit 1 (COI) gene of the Chinese alligator (Wu et al., unpublished). Additional single nt at positions 1187, 1195, and 1237 lead to shifts in the reading frame of this gene, although these sites are well conserved in all vertebrates. As the extra nt may suggest simple sequencing artifacts, the affected 51-nt region was removed from the aligned data set.

After excluding gaps and ambiguous sites around gaps, 9489 nt sites remained for phylogenetic analysis. A  $\chi^2$  test as implemented in the TREE-PUZZLE program package did not indicate a significant deviation

of the aa composition or of the nt composition at first or second codon positions, except for the opossum. Inspection of distance values indicated that all crocodilian lineages evolve much faster than other amniote groups. A  $\chi^2$  test showed that this increase in the evolutionary rate is significant. Among the crocodiles the evolutionary rates are not significantly different, except for the caiman.

Irrespective of evolutionary model or method, all phylogenetic analyses of both aa and nt sequences resulted in the same crocodilian relationships as shown in Fig. 2. All crocodilian branches received bootstrap support or probability values exceeding 95%. The support values for aa sequences are given in Table 2. The mitogenomic analyses split Crocodylia into three main lineages. The basal split is between Alligatoridae (genera *Alligator* and *Caiman*) and a branch that includes Crocodylidae (genus *Crocodylus*) and *Gavialis/Tomistoma*. The sister-group relationship between *Gavialis* and *Tomistoma* is well supported, and alternative positions for either species are significantly rejected by all analysis (Table 3). Similarly, the grouping of *Crocodylus* and *Gavialis/Tomistoma* receives significant support. A position of *Gavialis* as the sister group to all other crocodilians as in the traditional tree is 8–11 standard deviations worse than the best tree and would require  $245 \pm 18$  additional aa substitutions (Table 3).

Turtles show a strong affinity to the Archosauromorpha (birds and crocodiles) and are most likely basal to this group. Any other position of turtles except for a (bird (turtle, crocodile)) relationship can be significantly rejected (not shown).

Figure 2 shows also the estimated divergence times and their two standard deviations among the crocodilian lineages under study. The divergence times have been estimated by two different approaches. The divergence times based on “multidivtime” are somewhat older than those based on the penalized likelihood method of the “r8s” programs, however, the dates are overlapping. Other dating methods based on distance data (Arnason et al. 2000) and non-parametric rate smoothing (Sanderson 2002) yielded values closer to that based on the penalized likelihood

**Table 3.** ML analysis of crocodilian relationships based on amino acid (aa) sequences

Topology	$\Delta\ln L/SE$	$p_{\text{Boot}}$	$p_{\text{SH}}$	$\Delta\ln L/SE_{\text{rh8}}$	$p_{\text{SH}_{\text{rh8}}}$	Steps/SD
allig.(croc,(Gav,Tom))	< -49,012 >	1.00	1.00	< -46,455 >	1.00	< 8,753 >
(Gav,Tom),(allig,croc)	-256/± 31.3	0.00	0.00	-176/± 25.2	0.00	+ 144/± 14
Gav,(allig,(croc,Tom))	-463/± 43.2	0.00	0.00	-295/± 32.6	0.00	+ 245/± 18

*Note.* allig—Alligatoridae (alligator, caiman); croc—Crocodylidae (Nile and estuarine crocodile); Gav—*Gavialis gangeticus*; Tom—*Tomistoma schlegelii*.  $\Delta\ln L$ : differences in log-likelihood ( $\ln L$ ) values and standard error (SE) relative to the best  $\ln L$  values (shown in brackets).  $p_{\text{Boot}}$ : bootstrap probability.  $p_{\text{SH}}$ : probability of a topology according to the SH test. rh8: assuming rate heterogeneity with eight classes of variable sites. Steps: number of additional nt substitutions relative to the number of substitutions of the MP tree (shown in brackets).

method but the dates did not differ significantly from those given in Fig. 2 (not shown). It may be argued that the reference points may be slightly older or younger. It should be kept in mind, however, that the use of even 10% older or younger divergence time will affect other datings only by approximately that proportion, i.e., less than the error based on the method and sequence data.

The most basal divergence among recent crocodiles was estimated to have taken place in the Jurassic at  $\approx 150$  MYBP. The splits within the family Alligatoridae are notably deep with the divergence between the alligators and the caiman in the mid-Cretaceous at  $\approx 110$  MYBP. The Chinese and American alligator were estimated to have diverged during the late Cretaceous at  $\approx 76$  MYBP, close to the K/T boundary. The split between *Tomistoma*/*Gavialis* and genus *Crocodylus* was placed in the late Cretaceous at  $\approx 80$  MYBP. The split between the two *Crocodylus* species was placed close to the Oligocene/Miocene border, at  $\approx 23$  MYBP, and that between *Tomistoma* and *Gavialis* in the Eocene at  $\approx 42$  MYBP.

## Discussion

The current results have provided conclusive statistical support for earlier molecular findings that identified *Gavialis* and *Tomistoma* as sister groups. Likewise, the mitogenomic findings place *Gavialis*/*Tomistoma* unequivocally as the sister group of genus *Crocodylus*. Hence these relationships can no longer be ignored or disregarded as being the result of homoplasy, an insignificant number of data, poor taxon sampling, or long branch attraction. The analyses have further stressed the discrepancy between the molecular and the morphological understanding of the position of *Gavialis* in the crocodilian tree. The disagreement between the morphological and the molecular results can, however, in essence be reduced to the placement of the root of the crocodilian tree. Morphological data have in general placed it on *Gavialis* (Norell 1989; Brochu 1997; Poe 1996), while the mitogenomic analysis and other molecular data (Densmore 1983; Gatesy et al. 1993, 2003; Hass et al. 1992; Harshman et al. 2003) place it on the branch that separates the Alligatoridae.

Previously available molecular data have not been comprehensive enough for allowing resolution and molecular dating of crocodilian divergences. Consequently, estimates of the divergences among recent crocodiles were largely based on the fossil record and the assumption that the gharial had a basal position among extant crocodiles.

There is solid fossil evidence that crocodiles originated in the late Permian/early Triassic,  $\approx 250$  MYBP. The oldest fossil leading to Crocodylia (*Crocodylotarsi*), *Stagonosuchus*, has been dated to 240 MYBP, while the age of *Protorosaurus*, the ancestor of birds and crocodiles (Archosauromorpha), has been dated to 255 MYBP (Benton 1990). Thus the origin of the crocodilian lineage is quite narrowly defined and can be used as a local reference point for estimating crocodilian divergences. The use of this crocodile-specific reference point is important, because the mt genomes of all crocodilian lineages evolve significantly faster than those of any other vertebrate group. This makes it difficult to use reference points placed outside the crocodilians themselves, as assumptions must be made about evolutionary rates along the branches leading to the crocodiles. Nevertheless, additional reference points that aid the divergence times estimates can be put on the mammalian branches. These calibration points are relatively deep in the vertebrate tree and they are relatively narrowly defined.

Although this approach has provided dating of crocodilian origin that is roughly consistent with the fossil-based dates (Janke and Arnason 1997; Janke et al. 2001), it can nevertheless be criticized as being vulnerable to error. Some molecular datings based on noncrocodilian reference points and ignoring their particular evolutionary rates have placed the time of crocodilian origin at 222 MYBP, thereby making their origin appreciably younger than suggested by the fossil record (Kumar and Hedges 1998). Another factor that may have contributed to this low estimate is the limited amount of data used.

The current mitogenomic estimates that are based on four reference points, two different dating methods and two different data sets (aa and nt), yield divergence times among recent crocodiles that are about two times older than previously assumed from

molecular data (Densmore 1983, Hass et al. 1992). Yet, the mitogenomic divergence times are not contradicted by the fossil record and are consistent with the appearance of, e.g., *Gavialis* in the Oligocene. Previous estimates of divergence dates may be biased by the use of very limited molecular data or by directly taking the appearance of a crocodile lineage in the fossil record as their molecular divergence time (Brochu 2004). This approach may, however, severely underestimate the time of genetic separation.

The molecularly based estimates of divergence times indicate that at least five crocodylian lineages (the caimans, the American alligator, the Chinese alligator, the crocodiles, and the gharial/false gharial clade) survived the K/T extinction 65 MYBP. In addition to several lineages of mammals (Janke et al. 1994) and birds (Cooper and Penny 1997), several lineages of crocodiles have survived the K/T extinction.

The current analyses and those by Gatesy et al. (2003) have yielded statistically significant support to the molecular view of crocodile relationships, originally proposed by Densmore (1983). In particular, and in contradiction to the traditional morphological interpretation, the molecular findings stress the non-basal position of *Gavialis* and the sister-group relationship between *Gavialis* and *Tomistoma*. On the assumption that the molecular findings are correct, the molecular results suggest that the evolutionary polarity of at least some morphological characters have been incorrectly interpreted in the morphological studies.

The traditional taxonomy groups crocodiles into three families, Alligatoridae, Crocodylidae, and Gavialidae. The molecular findings suggest that this system needs to be revised to recognize only two families of recent crocodylians, Alligatoridae and Crocodylidae, with the latter including *Crocodylus*, *Ostocolaemus*, *Gavialis*, and *Tomistoma*.

*Acknowledgments.* The study was supported by the Swedish Science Council, the Erik Philip-Sörensen Foundation, and the Nilsson-Ehle Foundation.

## References

- Adachi J, Hasegawa M (1996) MOLPHY version 2.3: Programs for molecular phylogenetics based on maximum likelihood. Inst Stat Math, Tokyo. Comput Sci Monogr 28:1–150
- Aggarwal RK, Majumdar KC, Lang JW, Singh L (1994) Generic affinities among crocodylians as revealed by DNA fingerprinting with a Bkm-derived probe. Proc Natl Acad Sci USA 91:10601–10605
- Archibald JD, Averianov AO, Ekdale EG (2001) Late Cretaceous relatives of rabbits, rodents, and other extant eutherian mammals. Nature 414:62–65
- Arnason U, Gullberg A, Gretarsdottir S, Ursing B, Janke A (2000) The complete mitochondrial genome of the sperm whale and the establishment of a new molecular reference for estimating eutherian divergence dates. J Mol Evol 50:569–578
- Benton MJ (1990) Phylogeny of the major tetrapod groups: morphological data and divergence dates. J Mol Evol 30:409–424
- Brochu CA (1997) Morphology, fossils, divergence timing, and the phylogenetic relationships of *Gavialis*. Syst Biol 46:479–522
- Brochu CA (2004) Patterns of calibration age sensitivity with quartet dating methods. J Paleont 78:7–30
- Cooper A, Penny D (1997) Mass survival of birds across the Cretaceous–Tertiary boundary: molecular evidence. Science 275:1109–1113
- Densmore LD (1983) Biochemical and immunological systematics of the order Crocodylia. In: Hecht MK, Wallace B, Prance GH (eds) Evolutionary biology, Vol 16. Plenum, New York, pp 397–465
- Densmore LD, White PS (1991) The systematics and evolution of the Crocodylia as suggested by restriction endonuclease analysis of mitochondrial and nuclear ribosomal DNA. Copeia 1991:602–615
- Drummond A, Strimmer K (2001) PAL: an object-oriented programming library for molecular evolution and phylogenetics. Bioinformatics 17:662–663
- Felsenstein J (1993) Phylogenetic Inference Programs (PHYLIP). University of Washington, Seattle, and University Herbarium, University of California, Berkeley
- Gatesy J, DeSalle R, Wheeler WC (1993) Alignment-ambiguous nucleotide site and the exclusion of systematic data. Mol Phylogenet Evol 2:152–157
- Gatesy J, Amato G, Norell M, DeSalle R, Hayashi C (2003) Combined support for wholesale taxic atavism in gavialine crocodylians. Syst Biol 52:403–422
- Gatesy J, Baker RH, Hayashi C (2004) Inconsistencies in arguments for the supertree approach: supermatrices versus supertrees of Crocodylia. Syst Biol 53:342–355
- Harshman J, Huddleston CJ, Bollback JP, Parsons TJ, Braun MJ (2003) True and false gharials: a nuclear gene phylogeny of crocodylia. Syst Biol 52:386–402
- Hass CA, Hoffman MA, Densmore LD, Maxson LR (1992) Crocodylian evolution: insights from immunological data. Mol Phylogenet Evol 1:193–201
- Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755
- 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 WK, von Haeseler A, Pääbo S (1994) The marsupial mitochondrial genome and the evolution of placental mammals. Genetics 137:243–256
- Janke A, Erpenbeck D, Nielsson M, Arnason U (2001) The mitochondrial genomes of a lizard, *Iguana iguana*, and the caiman, *Caiman crocodylus*: implications for amniotes phylogeny. Proc Roy Soc Lond B 268:623–631
- Ji Q, Luo ZX, Yuan CX, Wible JR, Zhang JP, Georgi JA (2002) The earliest known eutherian mammal. Nature 416:816–822
- Kumar S, Hedges SB (1998) A molecular timescale for vertebrate evolution. Nature 392:917–920
- Nilsson AM, Gullberg A, Spencer P, Arnason U, Janke A (2004) Marsupial relationships and a timeline for marsupial radiation in South Gondwana. Gene 340:189–196
- Norell MA (1989) The higher level relationships of extant crocodylia. J Herpetol 23:325–335
- Poe S (1996) Data set incongruence and the phylogeny of crocodylians. Syst Biol 45:393–414
- Quinn TW, Mindell DP (1996) Mitochondrial gene order adjacent to the control region in crocodile, turtle and tuatara. Mol Phylogenet Evol 5:344–351

- Sanderson MJ (2002) Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Mol Biol Evol* 19:101–109
- Shimodaira H, Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol Biol Evol* 16:1114–1116
- 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
- Tarsitano SF, Frey E, Riess J (1989) The evolution of the Crocodylia: a conflict between morphological and biochemical data. *Am Zool* 29:843–856
- Wu X, Wang Y, Zhou K, Zhu W, Nie J, Wan C, Xie W (2002) The complete mitochondrial genome sequence of Chinese alligator. AF511507 (unpublished)