

Phylogenetic Relationships of South American Alligatorids and the *Caiman* of Madeira River

TOMAS HRBEK^{1,2*}, WILLIAM RANGEL VASCONCELOS², GEORGE REBELO³,
AND IZENI PIRES FARIAS²

¹Biology Department, University of Puerto Rico, Rio Piedras, San Juan, Puerto Rico

²Departamento de Ciências Biológicas, Laboratório de Evolução e Genética Animal, Universidade Federal do Amazonas (UFAM), Manaus, AM, Brasil

³Departamento de Ecologia, Instituto Nacional das Pesquisas da Amazonia (INPA), Manaus, AM, Brasil

ABSTRACT We analyzed DNA sequences of the mitochondrial cytochrome *b* gene (*cyt b*), the nuclear Recombination Activating Gene 1 (*RAG1*) and the myelocytomatosis oncogene (*MYC*) to infer the phylogenetic relationship of *Caiman crocodilus* and *Caiman yacare*, and other South American alligatorid crocodylian species. Phylogenetic relationships were robustly supported with *Paleosuchus* sister to *Melanosuchus* and *Caiman*. Phylogenetic relationships of *C. crocodilus* and *C. yacare* were unclear as these two species share mitochondrial and nuclear haplotypes. Specifically this sharing occurs among specimens of *C. yacare* and *C. crocodilus* from the Madeira River drainage. Two potential explanations stand out: secondary contact followed by hybridization, and differentiation along a cline. Current data cannot resolve between these two competing hypotheses. In comparison with *C. yacare* and *C. crocodilus*, *Paleosuchus trigonatus* and *Paleosuchus palpebrosus* are very well differentiated and also show surprising haplotypic diversity in spite of their phenotypic similarity. *J. Exp. Zool.* 309A:588–599, 2008. © 2007 Wiley-Liss, Inc.

How to cite this article: Hrbek T, Vasconcelos WR, Rebelo G, Farias IP. 2008. Phylogenetic relationships of South American Alligatorids and the *Caiman* of Madeira River. *J. Exp. Zool.* 309A:588–599.

The genus *Caiman* is taxonomically one of the most complicated genera of alligatorid crocodylians. Depending on the authority, the genus comprises two or three species. No authors dispute the taxonomic validity of *Caiman latirostris*; however, the species status of *Caiman yacare* is not universally accepted. Some authors accord it species status, whereas others consider it one of the subspecies of *Caiman crocodilus*. Recognized subspecies of *C. crocodilus* are *C. c. apaporiensis* (Medem, 1955); *C. c. chiapasius* (Bocourt, 1976); *C. c. fuscus* (Cope, 1968); the nominal species *C. c. crocodilus* (Linnaeus, 1758); and *C. c. yacare* (Daudin, 1801).

A recent comprehensive analysis of the *C. crocodilus* group by Busack and Pandya (2001) addressed the issue of morphological distinctness of the different subspecies of *C. crocodilus*. The authors analyzed 13 external morphological characters among 360 individuals. On the basis of

discriminant function analysis, Busack and Pandya (2001) concluded that *C. c. yacare* is sufficiently differentiated from all other subspecies of *C. crocodilus*, a result which the authors used to elevate *C. c. yacare* to full species status. Other subspecies were not morphologically different from each other or from the nominal subspecies *C. c. crocodilus*. Despite the overarching scope of this study, the authors excluded six specimens from the upper Madeira River transition zone, an area where *C. crocodilus* and *C. yacare* are thought to overlap (Brazaitis et al., '98). These six

Grant sponsors: CNPq/PPG7/5570902005-9, NIH and NSF.

*Correspondence to: Tomas Hrbek, Biology Department, University of Puerto Rico, Rio Piedras, Ave. Juan Ponce de Leon, San Juan, PR, Puerto Rico 00931. E-mail: hrbe@uprrp.edu

Received 30 July 2007; Revised 20 October 2007; Accepted 25 October 2007

Published online 27 December 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jez.430

specimens were excluded because the authors used area of origin to assign individuals to species and on the grounds that these specimens might also represent hybrids or integrades.

We reexamined the hypothesis of Busack and Pandya (2001) from a molecular perspective. We sampled individuals of *Caiman* from a number of drainages in Brazil, Peru and French Guyana (Fig. 1). We sampled *C. latirostris*, *C. crocodilus* and *C. yacare*, and we also included *C. c. fuscus* from Venezuela. We sampled multiple individuals from the Madeira River drainage, which contains both *C. crocodilus* and *C. yacare*, including the potential transition zone of upper Madeira River (Brazaitis et al., '98). All individuals were assigned to either *C. crocodilus* or *C. yacare*, with animals considered *C. yacare* when they possessed bold dark blotches on the sides of the upper and lower jaws, and marbled dark pattern on the lateral ventral scales, two characteristics that consistently differentiate *C. yacare* and *C. crocodilus* (Brazaitis et al., '98; Busack and

Pandya, 2001). We tested the implication of the study of Busack and Pandya (2001) that *C. crocodilus* and *C. yacare* are evolutionarily distinct entities, but in our analyses we also included multiple individuals from the problematic transition zone. This sampling scheme allowed us to test genetic porousness of the Madeira River transition zone if in fact *C. crocodilus* and *C. yacare* are distinct evolutionary entities.

MATERIALS AND METHODS

Field protocol

The majority of the collections were nocturnal. Each sample was obtained during routine marking of specimens for mark-recapture studies, and consisted of one or more scutes removed from the tail. Scutes were preserved in 95% ethanol, and kept at ambient temperature until delivered to the laboratory. In total 46 individuals were sampled, representing 25 *C. crocodilus*, six *C. yacare*, two *C. latirostris*, five *Melanosuchus niger*, three

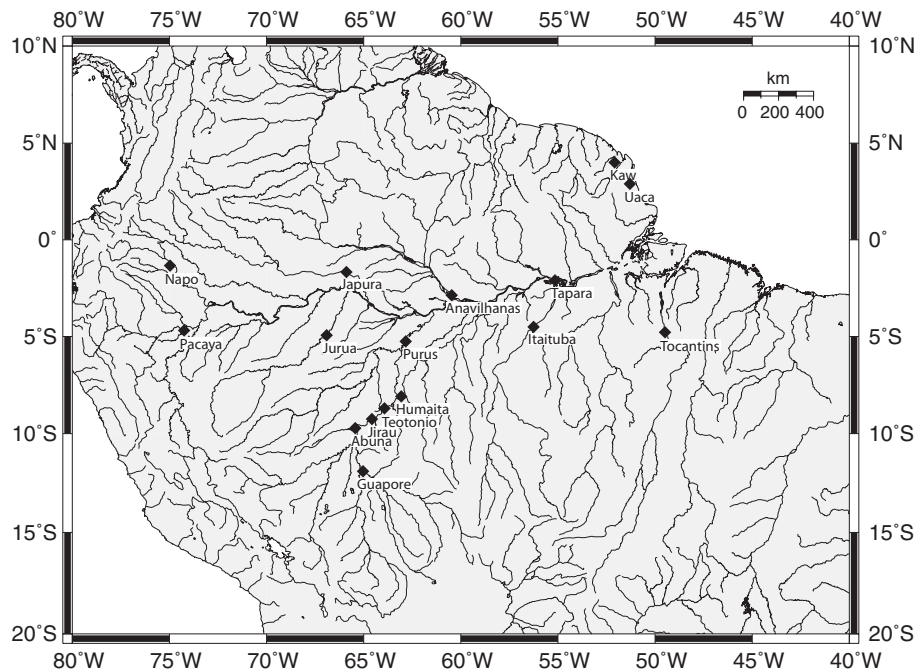


Fig. 1. A map of sampled localities for all species collected for our study. The map does not include the sample of *Caiman latirostris* (Piracicaba, Brazil; 22° 41.15'S, 47° 41.37'W) whose locality lies outside the range of this map. Names refer to closest large geographic feature or a city, and do not necessarily reflect local indigenous name. Geographic coordinates of localities are: Kaw River, French Guiana (4° 9.12'N, 52° 22.93'W); Uaçá River, Brazil (3° 54.47'N, 51° 25.17'W); Tocantins River, Brazil (4° 17.93'S, 49° 43.95'W); Tapará Reserve, Brazil (2° 1.45'S, 54° 49.40'W); Itaituba, Brazil (4° 0.73'S, 55° 39.22'W); Anavilhanas Reserve, Brazil (2° 37.11'S, 60° 48.75'W); Purus River (Piagaçu-Purus Reserve), Brazil (4° 14.43'S, 61° 41.73'W); Japurá River, Brazil (2° 2.00'S, 65° 13.26'W); Jurua River, Brazil (4° 17.93'S, 49° 43.95'W); Napo River (Cuyabeno Reserve), Ecuador (0° 27.15'S, 76° 38.23'W); Pacaya-Saimiria Reserve, Peru, (3° 29.21'S, 75° 25.07'W); Humaitá, Brazil (8° 15.76'S, 63° 22.41'W); Teotonio Rapids, Brazil (8° 51.37'S, 64° 2.76'W); Jirau Rapids, Brazil (9° 18.61'S, 64° 42.09'W); Abunã, Brazil (9° 46.13'S, 65° 43.08'W); Guapore River, Brazil (12° 5.73'S, 64° 46.25'W).

Paleosuchus palpebrosus and five *Paleosuchus trigonatus*. For analyses, additional sequence data obtained from GenBank included *C. c. fuscus*, *Alligators mississippiensis* and *Alligator sinensis*.

Laboratory protocol

Total genomic DNA was extracted from tail scutes using standard phenol chloroform method: samples were digested with a proteinase K/SDS solution, followed by phenol-chloroform extraction, the addition of 5 M NaCl, and a final 70% ethanol precipitation of DNA product (Sambrook et al., '89).

We amplified one mitochondrial and two nuclear genes. The mitochondrial fragment consisted of a complete cytochrome *b* gene (*cyt b*), tRNA threonine and partial tRNA proline and glutamine genes. Nuclear gene regions consisted of exon 2 of Recombination Activating Gene 1 (RAG1) and the myelocytomatosis oncogene (MYC). Nuclear genes were amplified using a nested polymerase chain reaction (PCR) approach (e.g. Hrbek et al., 2007).

All PCR reactions were carried out in a final volume of 25 μ L and contained 4.7 μ L of ddH₂O, 2 μ L of 25 mM MgCl₂, 2.0 μ L of 10 mM dNTPs, 6.7 μ L of 6-M Betaine, 2.5 μ L of unacetylated Bovine Serum Albumin (10 mg/mL), 2.5 μ L of 10 \times PCR buffer (100 mM Tris-HCl, 500 mM KCl), 2 μ L of each 2- μ M primer, 0.7 μ L of 1 U/ μ L Fermentas Taq DNA Polymerase, and 1 μ L of template DNA (between 50 and 100 ng). In case of the second amplification step for nuclear fragments, we used 1 μ L of the first amplification PCR as template.

Cycling conditions were as follows: preheating at 68°C for 60 sec, denaturation at 92°C for 15 sec, annealing at 52–55°C for 35 sec, and extension at 68°C for 90 sec. Denaturation to extension steps were repeated 35 times, followed by a final extension step at 68°C for 7 min. PCR primers and annealing temperatures are listed in Table 1. PCR products were evaluated on a 1% agarose gel, and purified using the Exonuclease I/Shrimp Alkaline Phosphatase method (Werle et al., '94). Between 50 and 100 ng of the PCR product was used as a template for each cycle sequencing reaction containing: 2 μ L of PCR product, 1 μ L of 10 μ M primer, 2.5 μ L of 5 \times sequencing buffer (400 mM Tris-HCl pH 9.0, 10 mM MgCl₂), 0.5 μ L of BigDye v3.1 and 4 μ L of ddH₂O. Cycling conditions were as follows: denaturation at 96°C for 10 sec, annealing at 50°C for 15 sec, and extension at 60°C for 240 sec; these steps were repeated 35 times. Cycle sequencing primers are listed in Table 1. Cycle sequencing reactions were precipitated with ethanol according to manufacturer's recommendations. Precipitated product was resuspended in 10 μ L of deionized formamide (ABI), injected and resolved on the ABI 3730 \times 1 Analysis System.

In the case of two individuals of *C. crocodilus* from Uaca (121cc) and Tapar (216cc), and two individuals of *P. trigonatus* from Teotonio (14pt) and Abun (249pt) whose MYC sequence had three, two, two and two heterozygous sites, respectively, we resolved haplotype ambiguity by cloning. Cloning was also used to resolve the haplotypic states RAG1 gene of one individual of *P. trigonatus* from Purus (1pt) also had two

TABLE 1. Primers used in amplification and sequencing of the mitochondrial cytochrome *b* (*cyt b*) and the nuclear myelocytomatosis oncogene (MYC) and Recombination Activating Gene 1 (RAG1) genes

Primer	Sequence 5'–3'	Temp	Use
GluCrf.1	CAACCAAAACCTGAGGYCTGA	52°C	Amp, sequence
CytbCrf.3	CCATACATYGGAGACACCAT	50°C	Sequence
CytbCrr.2	AAGATYAGGTGGGTKATGAG	50°C	Sequence
ProCrr.1	ATTAGAAYGTCGGCTTTGGGG	52°C	Amp
Rag1Crf.1	GAGGAACTTTGCCGCATCTGTGGA	52°C	First amp
Rag1Crr.1	AGTCCTGTACATRTTTRTGGTAYTG	52°C	First amp
Rag1Crf.2	gtaaaacgacggccagtGATCTTTAAGATTGATGTGCGAGG	55°C	Amp, sequence ¹
Rag1Crr.2	aatagactcactatagAAATGTATTGACTCGAATTGC	55°C	Amp, Sequence ²
MYCf.1	GGAGGACGGGAGAGTGCTRTC	52°C	First amp
MYCr.1	CCCAAAGTATTGGTTATGAGGCA	52°C	First amp
MYCf.2	gtaaaacgacggccagtCACATCTGCTGTGCTCAGCAGT	55°C	Amp, sequence ¹
MYCr.2	aatagactcactatagTTGAGGTTATGGCTCAGCTC	55°C	Amp, sequence ²

The lower case letters in the primer sequences are the M13 and T7 tails of the forward and reverse primers, respectively. M13 and T7 primers were used for sequencing.

¹Sequencing reactions were carried out with the M13 (–21) primer at 50°C.

²Sequencing reactions were carried out with the T7 primer at 50°C.

heterozygous sites. Furthermore, we cloned the MYC gene PCR product of *P. palpebrosus* from Teotonio (15pt), whose two haplotypes had different lengths, and *M. niger* from Humaitá (239mn) and Napo (299mn), whose haplotypes differed in the direction of an inverted repeat sequence. We used the pJET cloning kit from Fermentas (Hanover, MD), and transformed the vector-ligated PCR product into electrocompetent JM109 *E. coli* cells from Promega (Fitchburg, WI). After recovery in super optimal catabolite medium, cells were plated on $1 \times$ Luria-Bertani Ampicillin (10 $\mu\text{g/ml}$) agar and allowed to grow overnight. Ten random colonies were picked from each transformation, inserted PCR product was PCR amplified using vector primers provided with the pJET cloning kit, cleaned with Exonuclease I/Shrimp Alkaline Phosphatase mix (Werle et al., '94), and directly sequenced using a T7 primer to resolved the haplotype states for the MYC and RAG1 genes of the eight individuals.

Data analysis

Sequence data were verified by blasting them to GenBank, and by direct comparison with published crocodylian sequences for each gene region. Sequence data were edited, and an alignment was constructed in the program BioEdit 6.0.7 (Hall, '99). Our final dataset consisted of 1147 bp of the complete *cyt b* gene, 886 bp of the exon 2 of RAG 1 gene, and 1043–1047 nucleotides of MYC gene. Differences in lengths of the MYC fragments were due to species-specific indels. For protein coding regions, sequence data were further verified by translation into amino acids; no stop codons or other anomalies were observed. Sequence data generated in this study as well as those obtained from GenBank were aligned using the algorithm Clustal W (Thompson et al., '96) implemented in the program BioEdit (Hall, '99), and confirmed by visual inspection.

For inferences of phylogenetic relationships the GTR+inv+gamma model was suggested as the most appropriate by the software MODELTEST 3.7 (Posada and Crandall, '98). Phylogenetic relationships of the *cyt b* sequences were estimated using a maximum-likelihood algorithm implemented in PAUP* 4b10 (Swofford, 2002) and Bayesian-inference algorithm implemented in MRBAYES 3.01 (Huelsenbeck and Ronquist, 2001) under the site-specific GTR model (Rodriguez et al., '90) of molecular evolution with a portion of sites treated as invariable. Maximum-

likelihood topology was estimated by a heuristic search, with 25 random additions and TBR branch swapping. Parameter values were estimated from the data. Robustness of the maximum-likelihood phylogenetic hypothesis was assessed by 100 bootstrap replicates with one random addition and TBR branch swapping. For Bayesian-inference estimate of phylogenetic relationships, we ran two simultaneous parallel runs of 5,000,000 generations, sampling trees and branch-length every 1,000 generations. Bayesian-inference was carried out an additional two times to access concordance among tree topologies derived from independent runs. Convergence of all estimated parameters was inferred using the Gelman and Rubin ('92) criterion. Log likelihoods stabilized within the first 5% of the run, and we discarded these initial 250,000 trees in the computation of a 50% majority rule consensus tree. Following Huelsenbeck et al. (2001), the frequency with which a clade occurred among the sampled trees was interpreted as the marginal posterior probability of that clade existing. These posterior probabilities are true probabilities under the assumed model of substitution (Rannala and Yang, '96); thus we considered clades to be significantly supported when Bayesian posterior probabilities were $>95\%$. Phylogenetic relationships were also estimated using maximum-parsimony algorithm as implemented in PAUP* 4b10 (Swofford, 2002); all changes were unordered and equally weighted. Bootstrap resampling (Felsenstein, '85) was applied to assess support for individual nodes using 1,000 bootstrap replicates with ten random additions and TBR branch swapping. The sister taxon relationship of the North American and Chinese occurring genus *Alligator* and the South and Central American occurring genera *Caiman*, *Melanosuchus* and *Paleosuchus* is well established by morphological and molecular data (e.g. Densmore and White, '91; Brochu, '97; Gatesy et al., 2003; Harshman et al., 2003). Both species of the genus *Alligator* were therefore used as outgroup taxa in our phylogenetic analyses.

Owing to the small number of mutations observed in the RAG1 and MYC genes, we estimated relationships among haplotypes using the statistical parsimony haplotype network approach (Templeton et al., '92; Clement et al., 2000). As PCR products were directly sequenced, the electrophoregrams represented a consensus of two haplotypes. The sequences of the two nuclear genes were therefore resolved into constituent

haplotypes. In all but few cases, resolution into haplotypes was unambiguous because one or fewer sites were observed to be heterozygous. In the case of two individuals of *C. crocodiles*, three individuals of *P. trigonatus*, one individual of *P. palpebrosus* and two individuals of *M. niger* we resolved haplotype ambiguity by cloning (see above).

Concordance among results obtained from the analyses of *cyt b*, RAG1 and MYC genes was inferred by visual inspection. Alternate topologies were tested using the Wilcoxon signed ranks test (Templeton, '83) and the Shimodaira–Hasegawa test (Shimodaira and Hasegawa, '99).

RESULTS

Among the 46 individuals analyzed in this study, we observed a total of 23 *cyt b* haplotypes, 20 MYC haplotypes and 11 RAG1 haplotypes. Of these, ten *cyt b*, eight MYC and two RAG1 haplotypes were observed in *C. crocodilus* and three *cyt b*, two MYC and three RAG1 haplotypes were observed in *C. yacare*. The MYC and RAG1 haplotypes observed in this study were also reported by Harshman et al. (2003) and Gatesy et al. (2003) in their specimens, including *Caiman c. fuscus*. Haplotypes observed for other species were also compatible with previously reported haplotypes for the same species (Gatesy et al., 2003; Harshman et al., 2003). All sequence data were deposited in GenBank under the accession numbers EU161660–EU161685 (*cyt b*), EU161686–EU161705 (MYC) and EU161706–EU161716 (RAG1).

Cyt b tree topologies resulting from all three methods and independent runs of phylogenetic analyses were congruent. Within *C. crocodilus* we observed two major clades, one of which also included *C. yacare*. *Caiman crocodilus* and *C. yacare* were sister to *C. latirostris*, and the *Caiman* clade was sister to *Melanosuchus*. The two *Paleosuchus* species were sister, and *Paleosuchus* was sister to the *Caiman* plus *Melanosuchus* clade. All relationships were well supported by maximum-parsimony and maximum-likelihood bootstrap values, and by Bayesian posterior probabilities (Fig. 2). Maximum-parsimony analysis resulted in two equally parsimonious trees with a length of 733 steps. Maximum-likelihood and Bayesian-inference analyses resulted in a single topology with a likelihood $-\ln = 4907.37004$. Enforcing the monophyly of *C. yacare* and *C. crocodilus* resulted in significantly longer (Templeton test, $P < 0.05$) and

significantly less likely (Shimodaira–Hasegawa test, $P < 0.05$) phylogenetic hypotheses.

Statistical parsimony networks of MYC and RAG1 haplotypes showed species-specific clusters with the exception of *C. crocodilus* and *C. yacare*, which shared haplotypes (Figs. 3 and 4); *C. latirostris* also shared haplotypes with *C. crocodilus* and *C. yacare* in the RAG1 network (Fig. 4). Haplotype sharing was observed between *C. crocodilus* and *C. yacare* for all three genes analyzed which otherwise unambiguously differentiated other South American alligatorid species (Figs. 2–4 and Tables 2 and 3). The only exception to this pattern was the RAG1 gene in which the *C. latirostris* haplotype did not differ from the main *C. crocodilus/yacare* haplotype (Fig. 4 and Table 3).

DISCUSSION

Phylogenetic analyses of mitochondrial DNA data and statistical parsimony network analyses of nuclear gene data resulted in concordant results, albeit of different degrees of resolution (Figs. 2–4). Haplotypic variation in all three genes clearly differentiated *P. palpebrosus*, *P. trigonatus* and *M. niger*. *Caiman latirostris* did not share haplotypes of *cyt b* or MYC genes with *C. yacare* and *C. crocodilus*; however, the only haplotype of RAG1 observed in the sample of *C. latirostris* was the most frequent *Caiman* RAG1 haplotype also observed in *C. yacare* and *C. crocodilus*. *Caiman yacare* and *C. crocodilus* shared haplotypes of all three genes analyzed. In the mitochondrial *cyt b* haplotypes where one would expect to observe reciprocal monophyly first (Moore, '95; Palumbi et al., 2001), we observe two well-supported clades, however, one of the clades contains haplotypes of both *C. yacare* and *C. crocodilus*. All *C. crocodilus* individuals in this clade are from the Madeira River drainage. This also includes individuals from near Teotonio on central Madeira, which is several hundred kilometers from the Amazon Rainforest/Pantanal transition where *C. crocodilus* is replaced by *C. yacare* as well as from the Abuña River, which is geographically closer to the transition zone. None of the three gene regions analyzed support the hypothesis that *C. yacare* is evolutionarily distinct from *C. crocodilus*, at least not at the molecular level.

Since 1933 *C. yacare* has been considered a subspecies of *C. crocodilus*, however, King and Burke ('81) questioned the validity of this taxonomic ranking, stating there was little justification for the subspecific rank. Busack and Pandya (2001) used

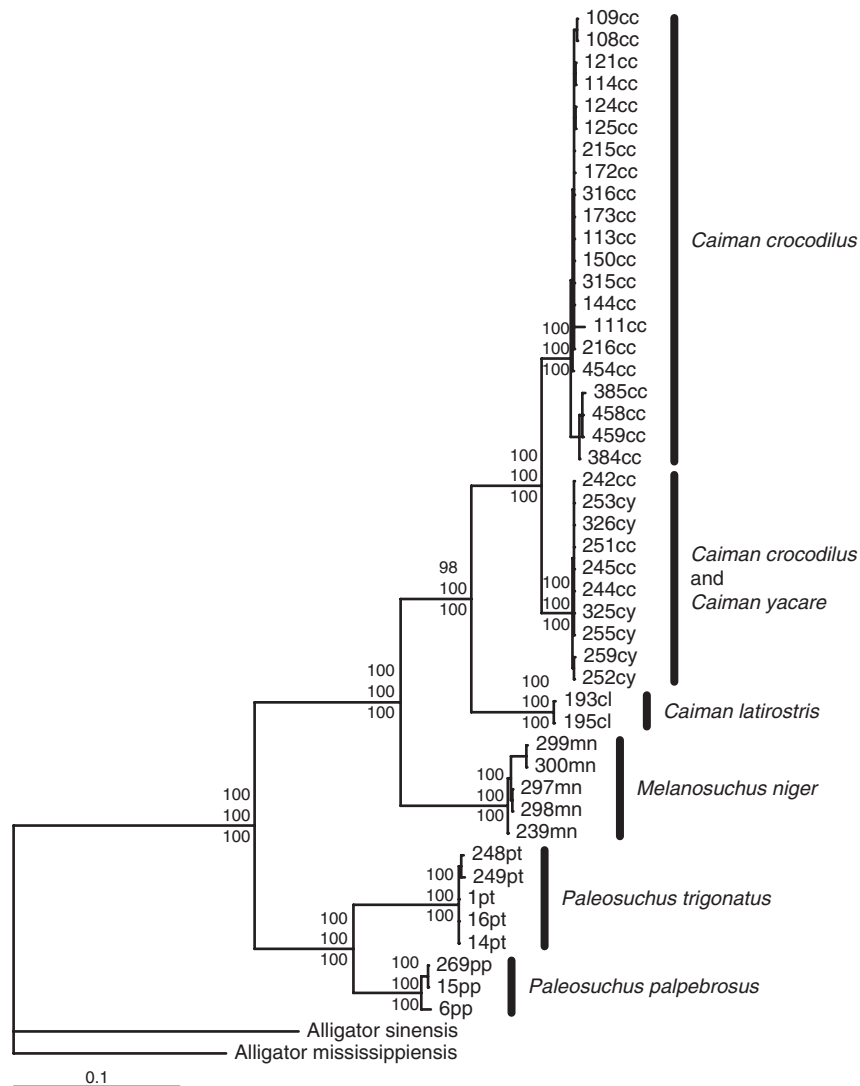


Fig. 2. A maximum-likelihood hypothesis of phylogenetic relationships of the South American alligatorid crocodylians. The parameters for the GTR+I+ Γ model of nucleotide substitution are as follows: Γ shape parameter $\alpha = 2.564744$, $I = 0.492267$ with a $-\ln$ likelihood = 4804.84984. Bayesian-inference hypothesis is topologically identical. Consensus of the 13 equally parsimonious trees is also compatible with the maximum-likelihood hypothesis. The numbers above branches represent Bayesian posterior probabilities and maximum-likelihood bootstrap values, and below branches are maximum-parsimony bootstrap values.

13 external morphological characters to reevaluate the taxonomic rank of *C. yacare*, concluding that there were no reasons to consider *C. yacare* a subspecies of *C. crocodilus* because over 75% of sampled individuals could be assigned to a correct groups using the discriminant function analysis. However, morphological differences between *C. yacare* and *C. crocodilus* may not be as clear as one might be led to believe from the study of Busack and Pandya (2001). *Caiman yacare* and all subspecies of *C. crocodilus* extensively overlapped when the information in the 13 morphological characters was reduced using principal component

analysis, i.e. differences between *C. yacare* and *C. crocodilus* disappear when a priori information of species status is not provided. Busack and Pandya (2001) also excluded six individuals from the upper Madeira River zone, an area of transition between *C. yacare* and *C. crocodilus*, because in their analysis the authors used area of origin to assign individuals to species, thus these specimens could not be assigned unambiguously to either *C. yacare* or *C. crocodilus*. Additionally the authors supposed that these specimens might also represent hybrids or integrades. The inclusion of these individuals would, therefore, obscure

morphological differences between *C. yacare* and *C. crocodilus*, but it also biased the results in the direction of finding differences between *C. yacare* and *C. crocodilus*.

In our study, we included *C. yacare* and *C. crocodilus* from their respective typical areas of natural distribution, but we also included animals of both species from the Madeira River drainage, same region from which Busack and Pandya (2001) excluded the six animals. In the Madeira River region, spanning both the Amazon Rainforest and the Bolivian Savannah biomes, the habitats of *C. crocodilus* and *C. yacare*, respectively, we collected four animals that were identified as *C. crocodilus* and six animals identified as *C. yacare*. Species identifications were based on observations in the field and measurements of external morphological characters. All ten individuals, irrespective of their taxonomic ranking had nuclear gene haplotypes that were found in *C. crocodilus* (Figs. 3 and 4; Tables 2 and 3) and they formed a well-supported mitochondrial haplotype clade sister to all remaining *C. crocodilus* haplotypes (Fig. 2).

There are at least two explanations that can account for this pattern. First, the upper Madeira River may be a zone of secondary contact between *C. yacare* and *C. crocodilus*, and an area where hybridization between these two

evolutionary lineages is geographically and biologically possible. A second possibility is clinal variation between the two species in the upper Madeira River region.

Hybridization between crocodylian species is not rare. A number of instances have been recorded in captivity (Ross, '98); however, potential cases of hybridization have also been observed in nature. For example, Ray et al. (2004) used the mitochondrial control region DNA to study *Crocodilus moreletii* in Belize. In July 2002 the authors sampled three animals from the Banana Bank, Belize that were phenotypically similar to *C. moreletii*; however, two of these three individuals had *Crocodilus acutus* haplotypes rather than *C. moreletii* haplotypes. The authors attributed this observation to potential hybridization. It would, therefore, not be surprising if hybridization could occur at the zone of contact between *C. yacare* and *C. crocodilus*. If hybridization were occurring, one would also likely observe introgression of morphological and genetic characters to differing degrees from one species into another, and to different geographic distances from the zone of contact.

In the region of the upper Madeira River and its tributaries, one could also realistically envision a cline between *C. yacare* and *C. crocodilus*.

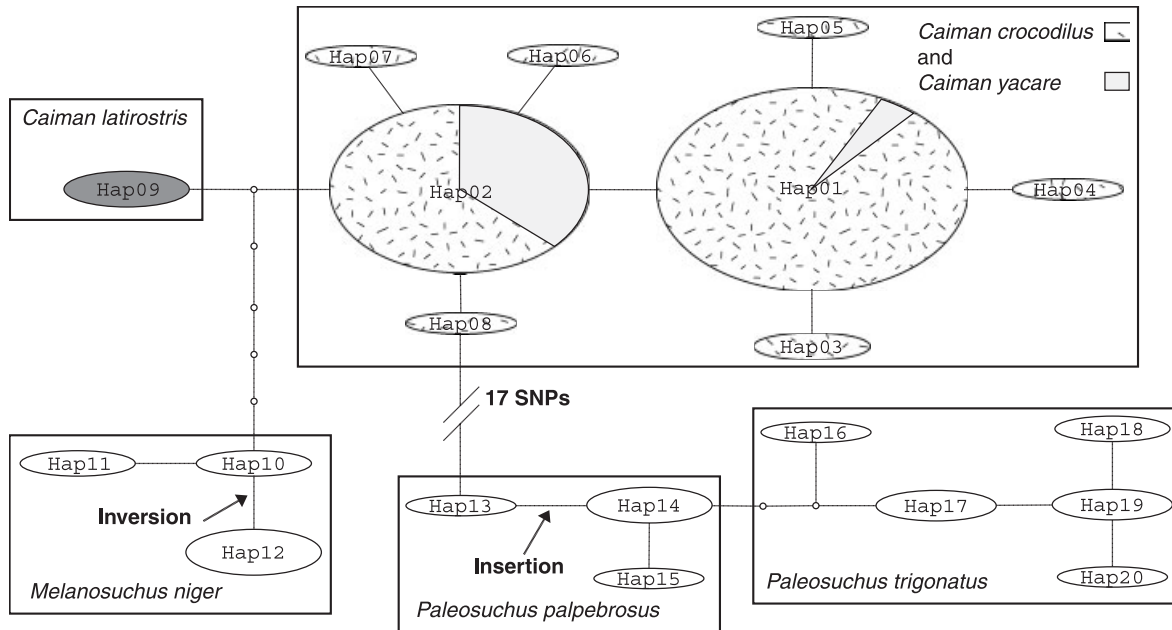


Fig. 3. A statistical parsimony network among the 20 myelocytomatosis oncogene (MYC) haplotypes. Each line in the network represents a mutational change. The small circles represent inferred but not detected haplotypes. Inversion represents the inversion first reported in Harshman et al. (2003), and insertion represents the additional GAA tri-nucleotide coding for the amino acid phenylalanine also reported by Harshman et al. (2003).

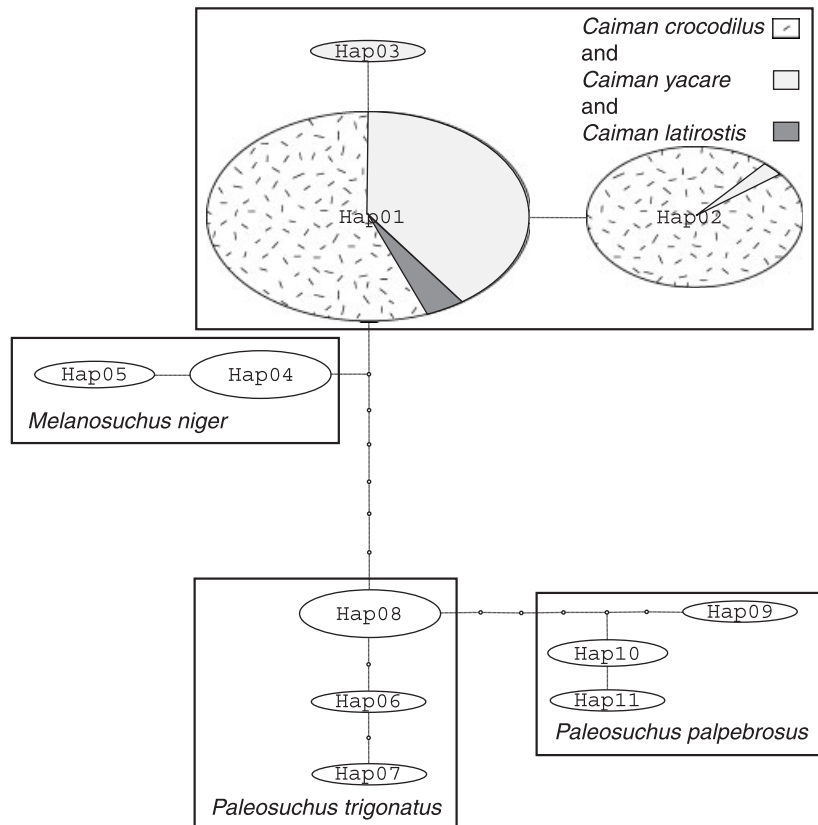


Fig. 4. A statistical parsimony network among the 11 Recombination Activating Gene 1 (RAG1). Each line in the network represents a mutational change. The small circles represent inferred but not detected haplotypes.

Diversification along a cline is one process that has received theoretical (Endler, '77) if not extensive empirical support (Haffer, '97) as a major process driving diversification in the tropics. This hypothesis predicts a parapatric distribution of sister taxa across a sharp ecological gradient. The ecological gradient is paralleled by a sharp gradient or a step cline in a morphological and genetic characters differentiating the parapatric sister taxa. The pattern of association of ecological parameters and species or groups of evolutionary lineages alone, however, does not allow one to differentiate between causal mechanisms of initial diversification, and mechanisms that are maintaining or reinforcing these differences, if these mechanisms are different. The region of the upper Madeira River forms a transition zone between the Pantanal plus Bolivian Savannah and Amazon rainforest biomes. Five of the 13 morphological characters studied by Busack and Pandya (2001) show a north/south cline, albeit an apparently gradual cline at a larger geographic scale, along a transect from *C. crocodilus* to *C. yacare*. The remaining seven characters showed no geographic

pattern, and one character, the blotching of the mandible, show a north/south decrease within *C. crocodilus*, and then a sudden increase in *C. yacare*; blotching of the mandible is the only character consistent with diversification along a sharp ecological cline. Mandible blotching is also one of the only two characters that show significant difference between *C. yacare* and *C. crocodilus*, but not among the remaining three subspecies of *C. crocodilus*. Clinal variation appears not uncommon, and in all but one instance shows a continuous transition from *C. crocodilus* to *C. yacare*. Differentiation along a cline is also compatible with haplotype sharing between *C. crocodilus* to *C. yacare* as a result of the retention of an ancestral polymorphism. Both *C. crocodilus* to *C. yacare* have large population sizes estimated at over a million individuals of *C. crocodilus* and at least 200,000 individuals of *C. yacare* (Ross, '98), and the time necessary to reach reciprocal monophyly is four times the inbreeding effective population size (e.g. Avise, 2004). Although on average one would expect *C. yacare*, the species with smaller population size,

TABLE 2. Distribution of myelocytomatosis oncogene (MYC) alleles among different species and individuals analyzed

Individual	Geographic location	Allele (MYC oncogene)																			
		01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20
<i>Caiman crocodilus</i>																					
108cc	Uaçá	X	X																		
109cc	Uaçá	X	X																		
111cc	Anavilhanas	XX																			
113cc	Uaçá	XX																			
114cc	Uaçá	XX																			
121cc	Uaçá				X			X													
124cc	Japurá	X	X																		
125cc	Japurá			XX																	
144cc	Japurá	XX																			
150cc	Anavilhanas	XX																			
172cc	Pacaya		XX																		
173cc	Pacaya	XX																			
215cc	Tapará	XX																			
216cc	Tapará		X			X															
315cc	Purus	XX																			
316cc	Purus	X	X																		
384cc	Tocantins		XX																		
385cc	Anavilhanas		XX																		
454cc	Juruá	XX																			
458cc	Itaituba		XX																		
459cc	Itaituba		X					X													
242cc	Abunã		XX																		
244cc	Abunã		XX																		
245cc	Abunã	X	X																		
251cc	Teotonio		X			X															
<i>Caiman yacare</i>																					
252cy	Jirau		XX																		
253cy	Jirau		XX																		
255cy	Jirau		XX																		
259cy	Teotonio		XX																		
325cy	Guapore	X	X																		
326cy	Guapore	X	X																		
<i>Caiman latirostris</i>																					
193cl	Piracicaba									XX											
195cl	Piracicaba									XX											
<i>Melanosuchus niger</i>																					
239mn	Humaitá										X	X									
297mn	Kaw											XX									
298mn	Kaw											XX									
299mn	Napo										X	X									
300mn	Napo									XX											
<i>Paleosuchus trigonatus</i>																					
1pt	Purus																		XX		
14pt	Teotonio																	X			X
16pt	Teotonio																			XX	
248pt	Abunã																	X			X
249pt	Abunã																X	X			
<i>Paleosuchus palpebrosus</i>																					
6pp	Tapará													XX							
15pp	Teotonio												X	X							
269pp	Abunã													X	X						

TABLE 3. Distribution of exon 2 Recombination Activating Gene 1 (RAG1) alleles among different species and individuals analyzed

Individual	Geographic location	Allele (RAG1)										
		01	02	03	04	05	06	07	08	09	10	11
<i>Caiman crocodilus</i>												
108cc	Uaçá	X	X									
109cc	Uaçá	X	X									
111cc	Anavilhanas	XX										
113cc	Uaçá	XX										
114cc	Uaçá	XX										
121cc	Uaçá	XX										
124cc	Japurá	X	X									
125cc	Japurá	X	X									
144cc	Japurá	X	X									
150cc	Anavilhanas		XX									
172cc	Pacaya	X	X									
173cc	Pacaya	XX										
215cc	Tapará	XX										
216cc	Tapará		XX									
315cc	Purus	XX										
316cc	Purus		XX									
384cc	Tocantins	X	X									
385cc	Anavilhanas		XX									
454cc	Juruá		XX									
458cc	Itaituba		XX									
459cc	Itaituba		XX									
242cc	Abunã	X	X									
244cc	Abunã	X	X									
245cc	Abunã	X	X									
251cc	Teotonio	XX										
<i>Caiman yacare</i>												
252cy	Jirau	XX										
253cy	Jirau	X		X								
255cy	Jirau	XX										
259cy	Teotonio	XX										
325cy	Guapore	XX										
326cy	Guapore	X	X									
<i>Caiman latirostris</i>												
193cl	Piracicaba	XX										
195cl	Piracicaba	XX										
<i>Melanosuchus niger</i>												
239mn	Humaitá				X	X						
297mn	Kaw				XX							
298mn	Kaw				XX							
299mn	Napo				XX							
300mn	Napo				X	X						
<i>Paleosuchus trigonatus</i>												
1pt	Purus						X	X				
14pt	Teotonio								XX			
16pt	Teotonio								XX			
248pt	Abunã								XX			
249pt	Abunã								XX			
<i>Paleosuchus palpebrosus</i>												
6pp	Tapará									XX		
15pp	Teotonio										XX	
269pp	Abunã										X	X

to become monophyletic first, genetic drift is a stochastic process, and thus it is possible that *C. crocodilus* will have become monophyletic before *C. yacare* reaches monophyly.

Both secondary contact and diversification along a cline give different predictions with respect of distribution and covariation of morphological and genetic diversity. No matter whether *C. yacare* and *C. crocodilus* speciated allopatrically and now are in secondary contact, or if they diverged or are diverging along a cline, the divergence between these two entities is most likely recent. Haplotypes of the RAG1 and MYC genes observed in *C. yacare* and *C. crocodilus* are not exclusive to either species. Furthermore, greater haplotypic diversity is observed in *C. crocodilus*, and *C. yacare* is nested within the most common RAG1 haplotype of *C. crocodilus*, and the two most common haplotypes of MYC gene. Similarly, our analysis of the mitochondrial *cyt b* gene reveals little divergence within the *C. yacare*/Madeira *C. crocodilus* clade compared with the remaining *C. crocodilus*. All data point to a recent divergence between *C. crocodilus* and *C. yacare*, with *C. crocodilus* potentially being a phylogenetic ancestor to *C. yacare*. Close relationship and recent divergence has previously also been suggested by Hass et al. ('92) based on immunological data. Although possible, it seems unlikely that divergence was ancient and was subsequently erased by hybridization. This explanation requires a second, recent cycle of divergence following the hybridization event that erased species differences to produce the currently observed genetic patterns. Accepting that divergence between *C. crocodilus* and *C. yacare* is recent, there are, however, insufficient data to discriminate between different competing diversification processes responsible for the observed molecular as well as morphological patterns. Resolution will need to come from a detailed sampling transect along the Madeira River that will examine morphological characters as well as mitochondrial and numerous independently assorting nuclear genes, and the concordance of patterns of distribution of these morphological, molecular and ecological characters. The region of the upper Madeira River represents a major biogeographical division for a number of taxa (e.g. Hamilton et al., 2001; Hubert and Renno, 2006), and therefore the process driving the differentiation in *Caiman* is likely to represent a general mechanism likely to affect many taxa in addition to *Caiman*.

Genomic evolution in *Paleosuchus* and *Melanosuchus*

Considering the small number of individuals sampled, it is worth noting a surprisingly large number of haplotypes observed within both species of *Paleosuchus*. Majority of individuals were heterozygous for both the RAG1 and MYC genes, and haplotypes differed at large number of sites. In the case of the RAG1 gene, intraspecific haplotype diversity and interspecific haplotype divergence for both *Paleosuchus* species was greater than observed within and among all three species of the genus *Caiman*. Given the shyness and inferred sedentary nature of both species of *Paleosuchus*, the observed haplotype diversity potentially is a signal of strong population structuring.

In a recent study by Harshman et al. (2003), the authors observed an insertion of a fourth GAA codon in a string of three GAA codons in *P. palpebrosus*. This represented an addition of a fourth Phenylalanine in *P. palpebrosus* when compared with other crocodylians. We confirm this observation; however, we note there is population level variation in the number of Phenylalanines in *P. palpebrosus*. Of the three *P. palpebrosus* analyzed in this study, two were homozygous for four Phenylalanine codons, and the third individual (15pt) was heterozygous for three/four Phenylalanine codons. All six individuals *P. trigonatus* had four Phenylalanine codons.

We also observe population level variation for the presence of the 29 base-pair inverted repeat reported by Harshman et al. (2003) in a MYC intron of *M. niger*. Of the five individuals analyzed in this study, one individual (300mn) was homozygous for non-inverted repeat, two individuals (297nm and 298nm) were homozygous for the inverted repeat, and two (239nm and 299nm) were heterozygous for the inverted repeat. The genomes of crocodylians are clearly dynamic, and the previously reported insertions and inversions are not fixed character states that can be used for species identification.

ACKNOWLEDGMENTS

We thank the following individuals for contributing samples to this study: Priscilla Marqui Schmidt Villela for *C. latirostris* samples, Ronis Da Silveira and Augusto Ruffeil for *P. trigonatus* samples, Benoit de Thoisy for samples from French Guiana and Ecuador, Boris Marioni for samples from Peru, and Adriana Kulaif Terra for

assistance in the field. We acknowledge the use of the UPR Sequencing and Genotyping facility. We also thank Bill Magnusson for discussing the Madeira River *Caiman*. Financial support was provided by CNPq/PPG7/ 5570902005-9, NIH and NSF, and RAN/IBAMA provided research permits.

LITERATURE CITED

- Avise JC. 2004. Molecular markers, natural history and evolution, 2nd edition. Sunderland, MA: Sinauer Associates. 541p.
- Brazaitis P, Rebêlo GH, Yamashita C. 1998. The distribution of *Caiman crocodilus crocodilus* and *Caiman yacare* populations in Brazil. *Amph-Rept* 19:193–201.
- Brochu CA. 1997. Morphology, fossils, divergence timing, and the phylogenetic relationships of *Gavialis*. *Syst Biol* 46:479–522.
- Busack SD, Pandya S. 2001. Geographic variation in *Caiman crocodilus* and *Caiman yacare* (Crocodylia: Alligatoridae): systematics and legal implications. *Herpetologica* 57: 294–312.
- Clement M, Posada D, Crandall KA. 2000. TCS: a computer program to estimate gene genealogies. *Mol Ecol* 9: 1657–1659.
- Densmore LD III, 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.
- Endler JA. 1977. Geographic variation, speciation, and clines. Princeton, NJ: Princeton University Press.
- Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783–791.
- 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.
- Gelman A, Rubin DB. 1992. Inference from iterative simulation using multiple sequences. *Stat Sci* 7:457–472.
- Haffer J. 1997. Alternative models of vertebrate speciation in Amazonia: an overview. *Biodivers Conserv* 6:451–476.
- Hall T. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41:95–98.
- Hamilton H, Caballero S, Collins AG, Brownell RL Jr. 2001. Evolution of river dolphins. *Proc R Soc Lond B* 268:549–556.
- 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 III, Maxson LR. 1992. Crocodylian evolution: insights from immunological data. *Mol Phylogenet Evol* 1:193–201.
- Hrbek T, Seckinger J, Meyer A. 2007. A phylogenetic and biogeographic perspective on the evolution of poeciliid fishes. *Mol Phylogenet Evol* 43:986–998.
- Hubert N, Renno J-F. 2006. Historical biogeography of South American freshwater fishes. *J Biogeogr* 33:1414–1436.
- Huelsenbeck JP, Ronquist FR. 2001. MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* 17:754–755.
- Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP. 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* 294:2310–2314.
- King FW, Burke RL. 1981. Crocodylian, tuatara and turtle species of the world. A taxonomic and geographic reference. Washington, DC: Association of Systematic Collections. 216p.
- Moore WS. 1995. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. *Evolution* 49:718–726.
- Palumbi SR, Cipriano F, Hare MP. 2001. Predicting nuclear gene coalescence from mitochondrial data: the three-times rule. *Evolution* 55:859–868.
- Posada D, Crandall KA. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14:817–818.
- Rannala B, Yang Z. 1996. Probability distribution of molecular evolutionary trees: a new method of phylogenetic inference. *J Mol Evol* 43:304–311.
- Ray DA, Dever JA, Platt SG, Rainwater TR, Finger AG, McMurry ST, Batzer MA, Barr B, Stafford PJ, McKnight J, Densmore LD III. 2004. Low levels of nucleotide diversity in *Crocodylus moreletii* and evidence of hybridization with *C. acutus*. *Conserv Gen* 5:449–462.
- Rodríguez FJ, Oliver JL, Marín A, Medina JR. 1990. The general stochastic model of nucleotide substitution. *J Theor Biol* 142:485–501.
- Ross JP. 1998. Crocodiles. Status survey and conservation action plan. 2nd edition. Gland, Switzerland: IUCN-SSC. Crocodile Specialist Group. 167p.
- Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: a laboratory manual. Cold Springs Harbor, NY: Cold Springs Harbor Laboratory Press.
- Shimodaira H, Hasegawa M. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol Biol Evol* 16:1114–1116.
- Swofford DL. 2002. PAUP*. Phylogenetic Analysis Using Parsimony (* and Other Methods), Beta Version 4b10. Sunderland, MA: Sinauer Associates.
- Templeton AR. 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. *Evolution* 37:221–244.
- Templeton AR, Crandall KA, Sing CF. 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data: III. Cladogram estimation. *Genetics* 132:619–633.
- Thompson JD, Higgins DG, Gibson TJ. 1996. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680.
- Werle E, Schneider C, Renner M, Völker M, Fiehn W. 1994. Convenient single-step, one tube purification of PCR products for direct sequencing. *Nucleic Acids Res* 22:4354–4355.