Available online at www.sciencedirect.com
SCIENCE DIRECT®
C. R. Palevol 2 (2003) 703–712
COMPTES RENDUS PALEVOL

Evolution (General Phylogenetics and systematic Theory)

Molecular evidence for species level divergence in African Nile Crocodiles *Crocodylus niloticus* (Laurenti, 1786)

Andreas Schmitz [a,*], Patrick Mansfeld [a], Evon Hekkala [b,d], Tara Shine [c], Hemmo Nickel [a], George Amato [d], Wolfgang Böhme [a]

[a] *Zoologisches Forschungsinstitut und Museum Alexander Koenig, Adenauerallee 160, 53113 Bonn, Germany*
[b] *American Museum of Natural History, Division of Invertebrate Zoology, 79th street Central park West, New York, NY 10024, USA*
[c] *42 Meadow Way, Kilkenny, Ireland*
[d] *Wildlife Conservation Society, 2300 Southern Boulevard, Bronx, NY 10460, USA*

Received 6 January 2003; accepted after revision 21 July 2003

Presented by Philippe Taquet

Abstract

Relationships of the newly discovered dwarf crocodiles from Mauritania were inferred from mitochondrial 12S sequences. Specimens from 13 different *Crocodylus niloticus* populations (from East Africa, West Africa and Madagascar) were compared. Additional representatives of the genus *Crocodylus* (one from Africa and one from Australia), the African genus *Osteolaemus* and the South American alligatorid *Paleosuchus palpebrosus* (as outgroup) were included in the analysis. Maximum-likelihood and Bayesian analyses yielded relationships that were strikingly different from currently prevailing phylogenetic hypotheses. Both analyses consistently revealed two groups, one consisting of the monophyletic West- and Central African populations and the other of a paraphyletic group containing the East African and Madagascan populations. High genetic divergence between those groups indicates separation on the species level. Furthermore '*C*' *cataphractus* is clearly shown not to be a member of the genus *Crocodylus*. The resulting nomenclatural changes are discussed. ***To cite this article: A. Schmitz et al., C. R. Palevol 2 (2003).***

Résumé

Preuve moléculaire de divergence au niveau spécifique chez le crocodile du Nil : *Crocodylus niloticus* (Laurenti 1786). Les relations de parenté de deux crocodiles nains de Mauritanie ont été déduites de l'analyse de séquences mitochondriales 12S. Des spécimens de 13 populations différentes de *Crocodylus niloticus* (d'Afrique de l'Est, de l'Ouest et de Madagascar) ont été comparés. Des spécimens supplémentaires du genre *Crocodylus* (un d'Afrique et un d'Australie), du genre africain *Osteolamus* et de l'alligatoridé *Paleosuchus palpebrosus* (comme extra-groupe) ont été inclus dans l'analyse. La probabilité maximale et les analyses bayésiennes ont livré des relations de parenté qui sont remarquablement différentes des hypothèses phylogénétiques classiques. Les deux analyses ont révélé logiquement deux groupes : l'un comprenant les populations

* Corresponding author.
E-mail address: a.schmitz.zfmk@t-online.de (A. Schmitz).

doi:10.1016/j.crpv.2003.07.002

monophylétiques d'Afrique de l'Ouest et centrale et l'autre un groupe paraphylétique comprenant des populations d'Afrique de l'Est et de Madagascar. Une forte divergence génétique entre ces groupes indique une séparation au niveau spécifique. En outre, il est clair que « *C* ». *cataphractus* n'appartient pas au genre *Crocodylus*. Les changements de nomenclature qui en résultent sont discutés. ***Pour citer cet article : A. Schmitz et al., C. R. Palevol 2 (2003).***

Keywords: Crocodylus; *C. niloticus*; *C. suchus*; *Mecistops*; 12S mtDNA; Maximum-likelihood; Bayesian analysis; Taxonomy

Mots clés : Crocodylus ; *C. niloticus* ; *C. suchus* ; *Mecistops* ; 12S mtDNA ; Probabilité maximale ; Analyse bayésienne ; Taxonomie

1. Introduction

There are three extant species of crocodilian currently known to exist in Africa. These are comprised of the endemic and monotypic dwarf crocodile (*Osteolaemus tetraspis)* narrowly confined to forests of West and Central Africa, the slender-snouted crocodile *(Crocodylus cataphractus)* of West and Central Africa and the Nile Crocodile (*C. niloticus*) which has an extensive distribution from Senegambia in the west to Egypt in the east, and southwards to South Africa and Madagascar [35,36]. Historically, several authors have proposed subspecies based on geographically correlated morphological differences. Many of these characters have been used in the reptile skin trade [4]. Nominal subspecies previously recognised [5,34,36] are: *C. n. niloticus* Laurenti, 1768 (restricted type locality: Egypt); *C. n. africanus* Laurenti, 1768 (restricted type locality: Tanzania); *C. n. chamses* Bory, 1824 (restricted type locality: Congo River); *C. n. cowiei* Smith and Hewitt, 1937 (type locality: South Africa); *C. n. madagascariensis* Grandidier, 1872 (type locality: Madagascar); *C. n. pauciscutatus* Deraniyagala, 1948 (type locality: Lake Rudolph); and *C. n. suchus* Geoffroy, 1807 (type locality: Niger River).

Nile crocodiles have disappeared from many parts of their historic range, particularly in northern Africa and in the Sahara desert [2,14]. However, in one Central Saharan site, viz. the Ennedi Mts. in Chad, and in some places in southern Mauritania, these desert crocodiles have survived up to the present [30,33]. Individuals from these desert populations have been described as significantly smaller than those found in other populations (not exceeding a total length of 2.3 m) raising the possibility of long-term isolates with distinct evolutionary history [14,30].

Presented here are the first mtDNA sequence data comparing the relict populations to other extant Nile crocodile populations. Initial results suggest significant genetic differentiation between relict and Eastern Nile crocodile populations. However, data from additional West African populations suggest an even more complex evolutionary history for *C. niloticus*, as very marked genetic differences could be found between samples from West- and East Africa. We therefore examined additional specimens from several allopatric populations throughout the complete distribution area of *C. niloticus*, and, as a result, propose some taxonomic changes.

2. Material and methods

To examine the genetic variation within *Crocodylus niloticus*, tissue samples of three Mauritanian, three more West African, and eight East African and one Malagasy *niloticus* population(s) were analyzed (Table 1; Fig. 2). To assess genetic differentiation between different species of the genus *Crocodylus*, we included *C. cataphractus* and *C. johnsoni* in the analysis. We further included the second African crocodile genus *Osteolaemus*. For outgroup comparison we further included the South American alligatorid, *Paleosuchus palpebrosus*. The voucher specimens with their collection numbers, localities and Genbank accession numbers are given in Table 1.

2.1. Genetic analysis

DNA was extracted from liver or muscle tissue (either fresh, preserved in 98% ethanol, or dried for museum specimens) using QuiAmp tissue extraction kits (Qiagen). We used the primers 12SA-L (light

Table 1
List of voucher specimens for each species included in the present study, with their respective localities, collection numbers and accession numbers (12S)
Liste des spécimens de chaque espèce inclus dans la présente étude, avec leurs localités respectives, leurs numéros de collection et leurs numéros d'accession (12S)

Species	Locality	Collection number	Accession number
Paleosuchus palpebrosus	South America	ZFMK 73079	AY195960
Osteolaemus tetraspis 1	Edéa, Cameroon	ZFMK 74854	AY195958
Osteolaemus tetraspis 2	Liberia	ZFMK 50692	AY195959
Crocodylus cataphractus	Lambaréné, Gabon	ZFMK 73109	AY195941
Crocodylus johnsoni	Australia	ZFMK 73662	AY195942
Crocodylus niloticus	Lake Nasser, Egypt	ZFMK, uncatalogued	AY195943
Crocodylus niloticus	Gambia	voucher not collected	AY195944
Crocodylus niloticus	Kenya	voucher not collected	AY195945
Crocodylus niloticus	Madagascar	voucher not collected	AY195946
Crocodylus niloticus 1	Aioun el-Atrouss, Mauritania	ZFMK, uncatalogued	AY195947
Crocodylus niloticus 2	Aioun el-Atrouss, Mauritania	ZFMK, uncatalogued	AY195948
Crocodylus niloticus 3	Aioun el-Atrouss, Mauritania	ZFMK, uncatalogued	AY195949
Crocodylus niloticus 1	Natal, South Africa	voucher not collected	AY195950
Crocodylus niloticus 2	Natal, South Africa	voucher not collected	AY195951
Crocodylus niloticus 3	Natal, South Africa	voucher not collected	AY195952
Crocodylus niloticus	Chor Melk en-Nasir, Sudan	ZFMK 50489	AY195953
Crocodylus niloticus 1	Kariba Dam, Zimbabwe	voucher not collected	AY195954
Crocodylus niloticus 2	Kariba Dam, Zimbabwe	voucher not collected	AY195955
Crocodylus niloticus	Senegal	voucher not collected	AY195957
Crocodylus niloticus	Ennedi Mts., Chad	voucher not collected	AY195956

Acronyms: ZFMK for Zoologisches Forschungsinstitut und Museum Alexander Koenig, Bonn

chain; 5′ - AAA CTG GGA TTA GAT ACC CCA CTA T - 3′) and 12SB-H (heavy chain; 5′ - GAG GGT GAC GGG CGG TGT GT - 3′) [15,25] to amplify a section of the mitochondrial 12S ribosomal RNA gene. Cycling procedure was as follows: 35 cycles: denaturation 45 s at 94 °C, primer annealing for 60 s at 50 °C, extension for 120 s at 74 °C.

PCR products were purified using Qiaquick purification kits (Qiagen). Sequences were obtained using an automatic sequencer (ABI 377) and manually corrected using the computer program Sequence Navigator (Applied Biosystems).

Sequences were aligned using the computer program ClustalX [32] (default parameters). Alignment was subsequently adjusted manually using the computer program BioEdit [8]. We explored the quality of our alignment by varying alignment gap opening cost (6, 9, 12) and comparing alignments.

Prior to phylogenetic reconstructions, we tested for homogeneity of base frequencies among taxa using the χ^2 test as implemented in PAUP*4.0b10 (which ignores correlation due to phylogenetic structure): (1) over all sites; (2) over parsimony-informative sites only; (3) without constant sites (parsimony-uninformative and constant sites will mislead the χ^2) test [23]. We performed maximum likelihood (ML) and Bayesian reconstructions. All maximum likelihood analyses [3] were performed with PAUP*4.0b10 [31]. In order to compare the results obtained via maximum likelihood analysis and Bayesian inference, the hierarchical likelihood-ratio test was carried out using MRMODELTEST 1.1b [24], a simplified version of MODELTEST [26,27], selecting the best-fit model of nucleotide substitution for our data set. Parameters of the model (substitution parameters, shape of gamma distribution, proportion of invariable sites) were estimated from the data set. The ML tree was calculated with the parameter estimates obtained under the best-fit model. A heuristic search was made with 10 replicates of random stepwise addition and tree bisection-reconnection (TBR) branch-swapping. Because of the extensive computation time used in ML bootstrap calculations, the relative branch support in phylogenetic analysis was evaluated with 100 boot-

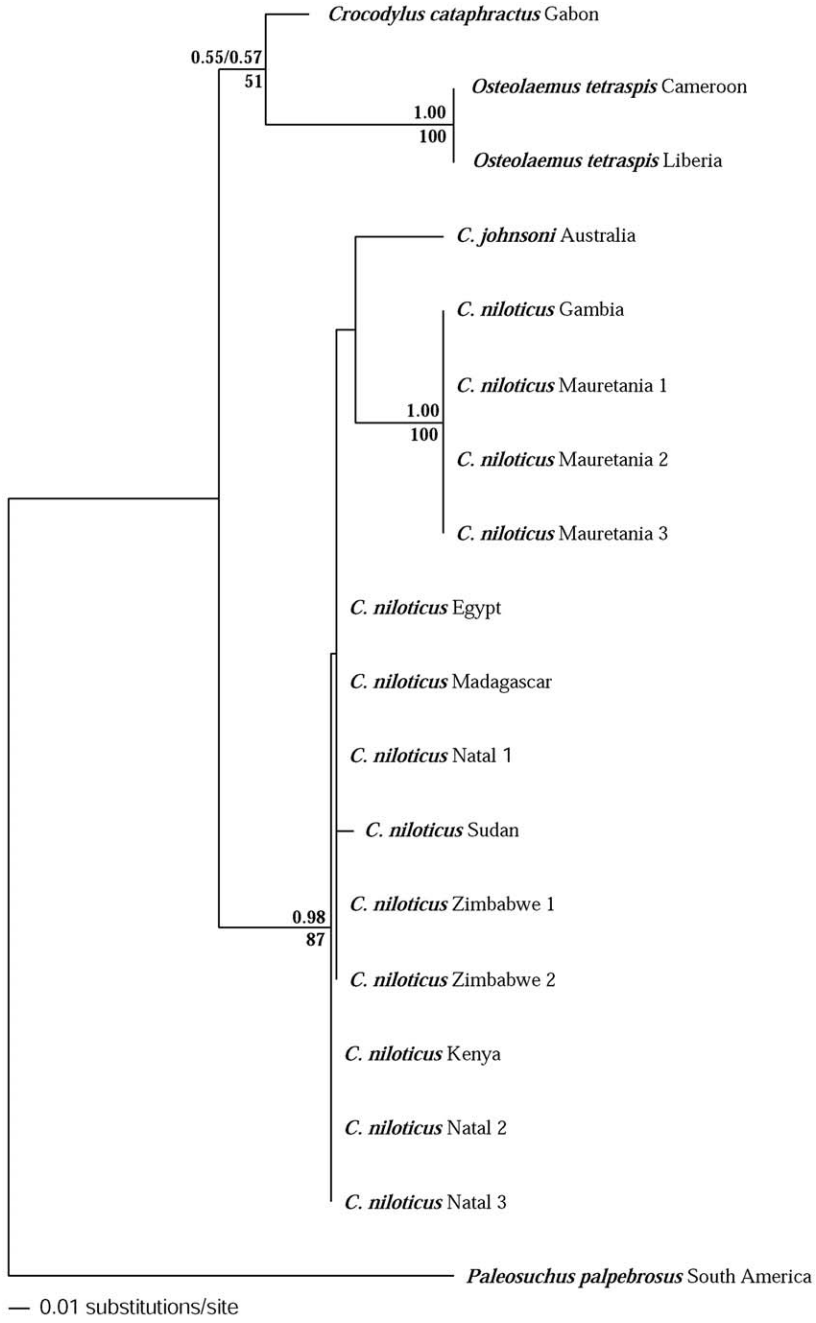

Fig. 1. Phylogram of the maximum likelihood tree (ML) and Bayesian analysis tree obtained from PAUP* and MrBayes searches using *Paleosuchus palpebrosus* as outgroup. Numbers above nodes represent posterior probabilities (PP), numbers below nodes represent bootstrap proportions for 100 pseudoreplicates for the likelihood analysis. Bootstrap proportions less than 50% are not shown.

Fig. 1. Phylogramme de l'arbre de probabilité maximale (ML) et arbre de l'analyse bayésienne obtenus avec PAUP* et avec les recherches de MrBayes en utilisant *Paleosuchus palpebrosus* comme extra-groupe. Les numéros au-dessus des nœuds représentent les probabilités postérieures (PP), les nombres au-dessous des nœuds représentent les proportions des *bootstraps* pour 100 pseudo-répliques de l'analyse de probabilité. Les proportions des *bootstraps* inférieures à 50% ne sont pas indiquées.

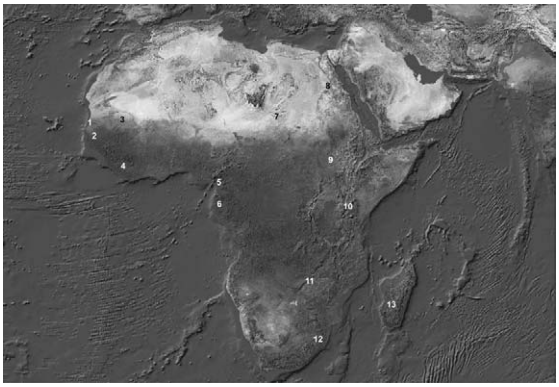

Fig. 2. Map of sample localities for specimens used in this study. Locality names are as follows: (1) Senegal, (2) Gambia, (3) Aioun el-Atrouss, Mauritania, (4) Liberia, (5) Edéa, Cameroon, (6) Lambaréné, Gabon, (7) Ennedi Mts., Chad, (8) Lake Nasser, Egypt, (9) Chor Melk en-Nasir, Sudan, (10) Kenya, (11) Kariba Dam, Zimbabwe, (12) Natal, South Africa, (13) Madagascar. Where the exact locality of the voucher is not known, we placed the corresponding number at a place where crocodiles are known to occur in the given country. The three desert localities are indicated by black numbers.
Fig. 2. Carte des localités ayant fourni des spécimens utilisés dans cette étude. Les noms des localités sont les suivantes : (1) Sénégal, (2) Gambie, (3) Aïoun el-Atrouss, Mauritanie, (4) Libéria, (5) Edéa, Cameroun, (6) Lambaréné, Gabon, (7) massif de l'Ennedi, Tchad, (8) Lac Nasser, (9) Chor Melk en-Nasir, Soudan, (10) Kenya, (11) Barrage de Kariba, Zimbabwe, (12) Natal, Afrique du Sud, (13) Madagascar. Lorsque la localité exacte de la pièce n'est pas connue, nous avons placé le numéro correspondant à l'endroit où les crocodiles sont connus dans la région. Les trois localités du désert sont indiquées par des numéros en noir.

strap pseudoreplicates (heuristic search, random addition of taxa with 10 replicates, TBR branch-swapping). We considered bootstrap values of 80 as giving strong support to the respective node, since this is a more conservative approach than the value suggested in [10].

Confidence in the phylogenetic signal for this molecular data set was assessed by calculating the skewness, or g1 statistic (implemented in PAUP*), which provides a measure of phylogenetic information content [11]. We produced 1000 randomly generated trees for ML (with outgroup excluded in both approaches; settings for ML are identical to the ones described above).

A matrix of pairwise sequence differences for the 12S rRNA genes was calculated using the p-distance (Table 3).

All Bayesian [13,16,17,21,29] analyses were performed with MRBAYES, version 3.0b3 [12], which approximates the posterior probabilities (PP) of trees. The program uses a variant of Markov chain Monte Carlo (MCMC), Metropolis-coupled MCMC [6], which is less prone to entrapment in local optima than is normal MCMC [7,9,22].

To select the best-fit model of nucleotide substitution for our data set, the hierarchical likelihood-ratio test was carried out using MRMODELTEST 1.1b [24]. Consequently, the settings of MRBAYES were speci-

Table 2
Parameter estimates of the substitution model (GTR + G), sampled after the burn-in phase of the chain. The columns indicate the parameter, mean and 95% credible interval for the parameter. The parameters are TL, the tree length; r_{ij}, rate of substitution between nucleotides i and j measured relative to the rate between G and T ($r_{GT} = 1$); π_i, base frequencies; and α, gamma shape parameter for among-site variation. Upper values in each pair correspond to the run No. 1; lower values correspond to the run No. 2
Paramètre estimé du modèle de substitution (GTR+G), échantillonné après la phase de combustion de la chaîne. Les colonnes indiquent le paramètre, la moyenne et l'intervalle de 95% de crédibilité pour le paramètre. Les paramètres sont TL, la longueur de l'arbre ; r_{ij}, rapport de substitution entre les nucléotides i et j mesurés selon le rapport entre G et T ($r_{GT} = 1$) ; π_i, fréquences des bases ; α, forme gamma du paramètre pour la variation de chaque site. Les valeurs supérieures de chaque paire correspondent à l'essai 1 ; les valeurs inférieures correspondent à l'essai 2

Parameter	Mean	95% Credity Interval
TL	0.94	(0.67, 1.37)
	0.94	(0.65, 1.39)
r_{GT}	1.00	
	1.00	
r_{CT}	36.43	(15.09, 49.64)
	30.66	(13.59, 49.12)
r_{CG}	1.05	(0.04, 3.26)
	1.03	(0.03, 4.22)
r_{AT}	6.33	(1.86, 12.82)
	4.98	(1.53, 11.17)
r_{AG}	19.75	(9.36, 35.49)
	19.15	(7.30, 32.72)
r_{AC}	8.93	(3.27, 16.88)
	7.79	(2.80, 15.84)
π_A	0.313	(0.274, 0.351)
	0.313	(0.275, 0.353)
π_C	0.274	(0.238, 0.312)
	0.274	(0.239, 0.311)
π_G	0.196	(0.164, 0.231)
	0.194	(0.162, 0.229)
π_T	0.217	(0.185, 0.252)
	0.218	(0.185, 0.254)
α	0.296	(0.194, 0.450)
	0.295	(0.191, 0.454)

Table 3
Summary of the uncorrected p-distances (sequences excluded from the tree are marked)
Sommaire des distances p, avant correction (des séquences exclues de l'arbre sont indiquées)

	Species	1	2	3	4	5	6	7	8	9	10	11	12
1	*Crocodylus cataphractus*	–											
2	*Crocodylus johnsoni*	0.097	–										
3	*C. niloticus* Egypt	0.074	0.045	–									
4	*C. niloticus* Gambia	0.094	0.073	0.046	–								
5	*C. niloticus* Kenya	0.077	0.048	0.002	0.044	–							
6	*C. niloticus* Madagascar	0.076	0.045	0.000	0.046	0.002	–						
7	*C. niloticus* 1 Mauritania	0.095	0.071	0.045	0.000	0.043	0.045	–					
8	*C. niloticus* 2 Mauritania	0.095	0.071	0.045	0.000	0.043	0.045	0.000	–				
9	*C. niloticus* 3 Mauritania	0.095	0.071	0.045	0.000	0.043	0.045	0.000	0.000	–			
10	*C. niloticus* 1 Natal, S. Africa	0.076	0.045	0.000	0.046	0.002	0.000	0.045	0.045	0.045	–		
11	*C. niloticus* 2 Natal, S. Africa	0.076	0.047	0.002	0.043	0.000	0.002	0.043	0.043	0.043	0.002	–	
12	*C. niloticus* 3 Natal, S. Africa	0.076	0.047	0.002	0.043	0.000	0.002	0.043	0.043	0.043	0.002	0.000	–
13	*C. niloticus* Sudan	0.083	0.053	0.007	0.054	0.010	0.007	0.054	0.054	0.054	0.007	0.010	0.010
14	*C. niloticus* 1 Zimbabwe	0.076	0.045	0.000	0.046	0.002	0.000	0.045	0.045	0.045	0.000	0.002	0.002
15	*C. niloticus* 2 Zimbabwe	0.076	0.045	0.000	0.046	0.002	0.000	0.045	0.045	0.045	0.000	0.002	0.002
16	*Osteolaemus tetraspis* Cameroon	0.078	0.140	0.113	0.131	0.110	0.113	0.133	0.133	0.133	0.113	0.110	0.110
17	*O. tetraspis* Liberia	0.085	0.144	0.117	0.132	0.114	0.117	0.137	0.137	0.137	0.117	0.114	0.114
18	*Paleosuchus palpebrosus*	0.184	0.200	0.186	0.205	0.190	0.188	0.204	0.204	0.204	0.188	0.188	0.188
19	*Crocodylus niloticus* Chad	0.137	0.093	0.055	0.005	0.060	0.055	0.005	0.005	0.005	0.055	0.060	0.060
20	*C. niloticus* Senegal	0.118	0.106	0.044	0.000	0.044	0.044	0.000	0.000	0.000	0.044	0.044	0.044

	Species	13	14	15	16	17	18	19	20
13	*C. niloticus* Sudan	–							
14	*C. niloticus* 1 Zimbabwe	0.007	–						
15	*C. niloticus* 2 Zimbabwe	0.007	0.000	–					
16	*Osteolaemus tetraspis* Cameroon	0.120	0.113	0.113	–				
17	*O. tetraspis* Liberia	0.121	0.117	0.117	0.000	–			
18	*Paleosuchus palpebrosus*	0.195	0.188	0.188	0.200	0.209	–		
19	*Crocodylus niloticus* Chad	0.055	0.055	0.055	0.189	0.189	0.263	–	
20	*C. niloticus* Senegal	0.044	0.044	0.044	0.148	0.148	0.211	0.000	–

fied according to the results of MRMODELTEST. Besides the specific parameters calculated by MRMODELTEST, the default settings of MRBAYES were used. We ran two MCMC analyses for 10^6 generations each. Each chain consisted of one cold and three heated chains and the Markov chains were started from a random tree. The Markov chains were sampled every 100th generation, resulting in 10 000 sampled trees from each chain. The initial 1000 (10%) trees were disregarded as 'burn-in' (the portion of the chain that was sampled before stationarity was reached). Inferences, then, were based on the 9000 trees samples from each chain. The topologies were used to generate a strict-consensus tree, with the percentage of samples recovering any particular clade representing that clade's posterior probability [12]. Unlike the non-parametric bootstrap values of the ML analysis, these are the true probabilities of the clades under the assumed model [29]. Consequently, we consider probabilities of 95% or greater to be significantly supported.

3. Results

The obtained 12S sequences (lengths referring to the aligned sequences including gaps) comprised 433 bp, with the exception of the tissue samples taken from populations from Senegal and Chad, which had strongly degenerated, making it impossible to get complete sequences from these samples. Nonetheless, we managed to get shorter fragments of the populations

sequenced (195 bp for the Chad sample and 78 bp for the Senegal sample) (Table 1). As the latter two sequences are much shorter than the other obtained sequences they were excluded from the calculations. Nonetheless, as these two sequences are part of a hypervariable region of the 12S gene, they could be used to classify the respective specimens (see below). In our 12S data set no ambiguous sites could be detected. The complete alignment is available from the authors on request. The matrix for the uncorrected p-distances for all nucleotide sites is presented in Table 3.

In the data set a phylogenetic signal is clearly present (g1 = -1.8185, p = 0.01). When all characters were included, we found no significant deviation from the homogeneity of base frequencies among taxa (χ^2 = 11.5660, p = 1.0000, df = 51). The same was true for the parsimony-informative sites only (χ^2 = 46.8861, p = 0.6377, df = 51) and without constant sites (χ^2 = 32.2954, p = 0.9739, df = 51).

Likelihood scores for models examined under MRMODELTEST showed that the GTR + G model [37] was determined to be the appropriate model for our data set. This model incorporates unequal base frequencies [$\pi_{(A)}$ = 0.3138, $\pi_{(T)}$ = 0.2117, $\pi_{(C)}$ = 0.2721, $\pi_{(G)}$ = 0.2023], and a gamma distribution shape parameter (α = 0.2931).

Both the ML and the Bayesian approaches produced identical topologies. Fig. 1 shows the ML tree (with lnL = -1292.25), with the posterior probabilities (if not identical, for the first and the second run) above the nodes and the ML bootstrap values below the nodes. Table 2 provides the estimates of the substitution parameters calculated by MRBAYES. The two independent MCMC runs converged on similar log-likelihood scores and reached stationarity no later than 100 000 generations. The posterior probability (PP) values supporting congruent nodes between the two runs were highly correlated (Fig. 1), further indicating that the analyses converged.

Three major clades are evident. The first clade comprises the two included *Osteolaemus tetraspis* and *Crocodylus cataphractus*. This clade, placing *tetraspis* and *cataphractus* as sister species, is only very weakly supported by a ML bootstrap value of 51 and posterior probability values of 0.55 and 0.57 for the first and for the second run, respectively. Nonetheless, that both species are very clearly placed outside the actual *Crocodylus* clade receives very high support in both the ML (87) and the Bayesian trees (0.98 for both runs). The second clade comprises all other *Crocodylus* species, with all East African *Crocodylus* specimens forming a large polytomy. This second clade is strongly supported by both the ML and the Bayesian analysis (ML: 87 / PP: 0.98). Within this monophyletic *Crocodylus* group our analyses revealed a group consisting of *Crocodylus johnsoni* and a maximally supported monophyletic group (ML: 100 / PP: 1.00), comprising all included West African *Crocodylus niloticus* specimens. However, the sister relationship of *C. johnsoni* towards the monophyletic West African Crocodylus group does not receive any bootstrap support.

On the genus level we find genetic differentiations (Table 3) between *Crocodylus* and *Osteolaemus* of at least 7.8% between *C. cataphractus* and *Osteolaemus*, between *Crocodylus* and *Paleosuchus* of at least 18.4%, between *C. cataphractus* and *Paleosuchus*, and between *Osteolaemus* and *Paleosuchus* of at least 20.0%. Genetic variation within *Crocodylus* ranged from 0.0% to 9.7%. The West African *Crocodylus* specimens show a genetic variation of 0.0%. The East African Crocodylus populations show a genetic variation of 0.0–0.7%. The West African *Crocodylus* specimens differ from *C. johnsoni* in 7.1–7.3%, and from their East African congeners in at least 4.3%. *Crocodylus johnsoni* shows a genetic differentiation of at least 4.5% from the East African *Crocodylus* specimens. *C. cataphractus* differs from the East African *C. niloticus* specimens in at least 7.4%. It differs from the West African specimens in 9.4–9.5%, and differs from *Osteolaemus* in at least 7.8%.

4. Discussion

Our results demonstrate that the African crocodiles are genetically much more diverse than previously thought. Both the ML and Bayesian analyses revealed the existence of two independent clades of *C. niloticus*: One monophyletic lineage comprising all West African samples studied, the genetic variation within this clade being 0.0% (Table 3); and a second independent clade containing all East African samples studied including the Madagascan one, but showing, nonetheless, only a very slight genetic variation throughout the entire range. This result does not rule out further genetic

subdivision within these lineages, as the 12S gene is rather conservative and may not resolve smaller differences to be found between conspecific populations. Therefore we will not try to address the question of intraspecific evolution in regard to the nominal subspecies described (see Introduction) without additional data. Such a study, employing broader regional sampling and additional mitochondrial and nuclear gene regions is underway (Hekkala et al., unpublished data).

The inclusion of *C. johnsoni* within the general *Crocodylus* clade shows this Australian species to be congeneric with the African *Crocodylus* species. However, the sister relationship of *C. johnsoni* towards the monophyletic West African *Crocodylus* group does not receive any bootstrap support, and together with the high genetic differences of *johnsoni* towards either *niloticus* clade (4.5–7.3) clearly shows that *C. johnsoni* can be regarded as a full independent species within the genus *Crocodylus*.

The very low internal genetic variation of 0.0–0.7% found within the eastern *niloticus* clade contrasts strongly with the high genetic divergence towards the western *niloticus* clade (at least 4.5%). As described in the results section above, the two short sequence fragments obtained from the Chad and the Senegal specimens are part of a hypervariable region within the 12S gene. Therefore, even these short fragments have accumulated a rather large number of substitutions and can consequently be safely assigned to belong to either of the described clades within *Crocodylus*. It becomes clear that the described western *niloticus* clade also includes the Central African sample (Chad) (only one substitution against the other members of the western clade, but 12 against the east African populations), and that the same is true for the sample from Senegal. Here is clear evidence that the West African *Crocodylus* clade reaches at least to Central Africa, even though due to the short sequence lengths the Chad and Senegal samples were excluded from the general phylogenetic analyses.

The level of distinctness indicated by these results suggests that taxonomic revision of *C. niloticus* is warranted. This includes resurrection of *Crocodylus suchus* Geoffroy, 1807, as the type locality of *Crocodylus niloticus* Laurenti, 1768 has been restricted to Egypt [see 35, 36].

The geographic split between the lineages runs roughly through an area that was strongly influenced by the repeated extensions and regressions of the Sahara desert, particularly during the Pliocene and Pleistocene [1,18–20]. Geographic barriers formed by rain forests may have repeatedly been effective enough to further an allopatric speciation process. The eastern species (C. niloticus s. str.) had always a hydrographic connection along the Nile from Egypt down to Ethiopia, Kenya and southwards, and the so-called ‘arid corridor’ plays an important biogeographic role as a link for many savannah-adapted plant and animal species [e.g. 28]. The Saharan relict crocodile populations, however, seem to be remnants of just the last humid phase of the desert, thus being isolated from the range of their big-growing southern conspecifics only for a few thousand years at best, by far not sufficiently long for establishing taxonomically relevant genetic differences. Clearly more work is needed to properly reconstruct the evolutionary history of Africa’s largest predators. What was once thought of as the common Nile crocodile, is perhaps not so common after all.

A final aspect important to note is that both ML and Bayesian inference consistently placed *cataphractus* outside the monophyletic *Crocodylus* clade. Confidence that *cataphractus* may not be congeneric with *niloticus* s.l. and *johnsoni* is indicated by the strong support of the corresponding node (ML: 87; PP: 0.98), which places *cataphractus* well outside the combined *Crocodylus* clade.

“*C.*” *cataphractus* clustered with *Osteolaemus*, but with only low support (ML: 51; PP: 0.55, 0.57), showing that it is also not congeneric with the *O. tetraspis*. High genetic differentiation to *Osteolaemus* (> 7.8%) shows that “*C.*” *cataphractus* is not a member of the genus *Osteolaemus* but probably deserves generic rank. *Mecistops* Gray 1844 would be available. Still, since we only included a single specimen and only one mitochondrial gene, it is clear that additional specimens and further genetic data (e.g. from nuclear genes) are required to clarify the taxonomic situation of *cataphractus*. This will be done in a subsequent study (Hekkala, unpublished data).

Acknowledgements

We thank Nicolà Lutzmann (ZFMK), Hans-Werner Hermann (Zoological Garden Cologne), Thomas Wilms (SMN, Karlsruhe), Rudolf Wicker (Zoological

Garden Frankfurt), Frank Glaw (ZSM, Munich), Miguel Vences (University of Amsterdam), the "Ferme aux crocodiles" (Pierrelatte), Michael Klemens (Wildlife Conservation Society), The American Museum of Natural History (AMNH), and The Field Museum of Natural History (FMNH), who provided samples for this study. The second author is indebted to John P. Huelsenbeck, Fredrik Ronquist, Jonathan P. Bollback and Johan Nylander for their introduction to the MrBayes program.

References

[1] W. Böhme, Die wechselvolle Geschichte der Sahara: Untersuchung von Reptilienzönosen entlang eines westsaharischen Transektes, mit einem überaschenden Fund im Südosten Mauretaniens, Tier u. Mus. 7 (1) (2000) 11–21.

[2] K. de Smet, Status of the Nile crocodile in the Sahara desert, Hydrobiologia 391 (1999) 81–86.

[3] J. Felsenstein, Evolutionary trees from DNA sequences: a maximum likelihood approach, J. Mol. Evol. 17 (1981) 368–376.

[4] K. Fuchs, Die Krokodilhaut, E. Roether, Darmstadt, 1977 183 p.

[5] K. Fuchs, R. Mertens, H. Wermuth, Die Unterarten des Nilkrokodils, *Crocodylus niloticus*, Salamandra 10 (3/4) (1974) 107–114.

[6] C.J. Geyer, Markov chain Monte Carlo maximum likelihood, in: E.M. Keramidas (Ed.), Computing Science and Statistics: Proc. 23rd Symposium on the Interface, Interface Foundation, Fairfax Station, 1991, pp. 156–163.

[7] P.J. Green, Reversible jump Markov chain Monte Carlo computation and Bayesian model determination, Biometrika 82 (1995) 711–732.

[8] T.A. Hall, BIOEDIT: A user-friendly biological sequence alignment editor and analysis program for WindowS95/98/NT, Nucl. Acids. Symp. Ser. 41 (1999) 95–98.

[9] W.K. Hastings, Monte Carlo sampling methods using Markov chains and their applications, Biometrika 57 (1970) 97–109.

[10] D.M. Hillis, J.J. Bull, An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis, Syst. Biol. 42 (1993) 182–192.

[11] D.M. Hillis, J.P. Huelsenbeck, Signal, noise, and reliability in molecular phylogenetic analyses, J. Hered. 83 (1992) 189–195.

[12] J.P. Huelsenbeck, F. Ronquist, MRBAYES: Bayesian inference of phylogeny, Bioinformatics 17 (2001) 754–755.

[13] J.P. Huelsenbeck, F. Ronquist, R. Nielsen, J.P. Bollback, Bayesian inference of phylogeny and its impact on evolutionary biology, Science 294 (2001) 2310–2314.

[14] L. Joleaud, Études de Géographie zoologiques sur la Berbérie, Bull. Soc. Zool. France 58 (1933) 397–404.

[15] T.D. Kocher, W.K. Thomas, A. Meyer, S.V. Edwards, S. Pääbo, F.X. Villablanca, A.C. Wilson, Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers, Proc. Natl Acad. Sci. USA 86 (1989) 6196–6200.

[16] B. Larget, D. Simon, Markov chain Monte Carlo for the Bayesian analysis of phylogenetic trees, Mol. Biol. Evol. 16 (1999) 750–759.

[17] S. Li, D.K. Pearl, H. Doss, Phylogenetic tree reconstruction using Markov chain Monte Carlo, J. Am. Stat. Assoc. 95 (2000) 493–508.

[18] K. Littmann, War die Sahara in der Vorzeit grün? Klimaschwankungen und ökologische Veränderungen im größten Trockengebiet der Erde – ein Überblick, Natur u. Mus. 117 (1987) 386–394.

[19] J. Maley, The African rain forest – main characteristics of changes in vegetation and climate from the Upper Cretaceous to the Quaternary, Proc. R. Soc. Edinb. 104B (1996) 31–73.

[20] C. Martin, Die Regenwälder Westafrikas. Ökologie, Bedrohung, Schutz, Birkhäuser, Basel Boston Berlin, 1989, 235 p.

[21] B. Mau, M. Newton, B. Larget, Bayesian phylogenetic inference via Markov chain Monte Carlo methods, Biometrics 55 (1999) 1–12.

[22] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, Equations of state calculations by fast computing machines, J. Chem. Phys. 21 (1953) 1087–1091.

[23] B. Misof, A.M. Rickert, T.R. Buckley, G. Fleck, K.P. Sauer, Phylogenetic signal and its decay in mitochondrial SSU and LSU rRNA gene fragments of Anisoptera, Mol. Biol. Evol. 18 (2001) 27–37.

[24] J.A.A. Nylander, MrModeltest 1.1b, Uppsala, Sweden, 2002.

[25] S.R. Palumbi, A. Martin, S. Romano, W.O. McMillan, L. Stice, G. Grabowski, The Simple Fool's Guide to PCR, Version 2.0, Department of Zoology, University of Hawaii, Honolulu, HI, 96822, 1991, Privately published document compiled by S. Palumbi 47 p.

[26] D. Posada, K.A. Crandall, MODELTEST: testing the model of DNA substitution, Bioinformatics 14 (1998) 817–818.

[27] D. Posada, K.A. Crandall, Selecting the best-fit model of nucleotide substitution, Syst. Biol. 50 (2001) 580–601.

[28] J.C. Poynton, The 'arid corridor' distribution in Africa: a search for instances among amphibians, Madoqua 19 (1) (1995) 45–48.

[29] B. Rannala, Z. Yang, Probability distribution of evolutionary trees: a new method of phylogenetic inference, J. Mol. Evol. 43 (1996) 304–311.

[30] T. Shine, W. Böhme, H. Nickel, D.F.G. Thies, T. Wilms, Rediscovery of relict populations of the Nile crocodile *Crocodylus niloticus* in south-east Mauritania, with observations on their natural history, Oryx 35 (3) (2001) 260–262.

[31] D.L. Swofford, PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Ver. 4, Sinauer Associates, Sunderland, Massachusetts, 2002.

[32] J.D. Thompson, T.J. Gibson, F. Plewniak, F. Jeanmougin, D.G. Higgins, The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools, Nucleic Acids Res. 24 (1997) 4876–4882.

[33] J. Tubiana, Quelques aberrations sahariennes: les crocodiles d'Archaï, Courr. Nat. 153 (1995) 27–29.
[34] H. Wermuth, K. Fuchs, Bestimmen von Krokodilen und ihrer Häute, G. Fischer, Stuttgart New York, 1978 100 pp.
[35] H. Wermuth, R. Mertens, Schildkröten, Krokodile, Brückenechsen, G. Fischer (reprinted 1996), Jena, 1961 506 pp.
[36] H. Wermuth, R. Mertens, Liste der rezenten Amphibien und Reptilien, Testudines, Crocodylia, Rhynchocephalia, Das Tierreich 100 (1977) 1–174.
[37] Z. Yang, Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods, J. Mol. Evol. 39 (1994) 306–314.