Rigorous approaches to species delimitation have significant implications for African crocodilian systematics and conservation

Matthew H. Shirley1,2, Kent A. Vliet3, Amanda N. Carr1 and James D. Austin1

1Department of Wildlife Ecology and Conservation, University of Florida, 110 Newins-Ziegler Hall, Gainesville, FL 32611, USA
2SFM Safari Gabon, BP 1107, Libreville, Gabon
3Department of Biology, University of Florida, 220 Bartram Hall, Gainesville, FL 32611, USA

Accurate species delimitation is a central assumption of biology that, in groups such as the Crocodylia, is often hindered by highly conserved morphology and frequent introgression. In Africa, crocodilian systematics has been hampered by complex regional biogeography and confounded taxonomic history. We used rigorous molecular and morphological species delimitation methods to test the hypothesis that the slender-snouted crocodile (Mecistops cataphractus) is composed of multiple species corresponding to the Congolian and Guinean biogeographic zones. Speciation probability was assessed by using 11 mitochondrial and nuclear genes, and cranial morphology for over 100 specimens, representing the full geographical extent of the species distribution. Molecular Bayesian and phylogenetic species delimitation showed unanimous support for two Mecistops species isolated to the Upper Guinean and Congo (including Lower Guinean) biomes that were supported by 13 cranial characters capable of unambiguously diagnosing each species. Fossil-calibrated phylogenetic reconstruction estimated that the species split $+6.5–7.5$ Ma, which is congruent with intraspecies divergence within the sympatric crocodile genus Osteolaemus and the formation of the Cameroon Volcanic Line. Our results underscore the necessity of comprehensive phylogeographic analyses within currently recognized taxa to detect cryptic species within the Crocodylia. We recommend that the community of crocodilian researchers reconsider the conceptualization of crocodilian species especially in the light of the conservation ramifications for this economically and ecologically important group.

1. Introduction

While numerous concepts have been proposed that emphasize different criteria for delimiting species [1], species themselves may be best conceptualized as population aggregates evolving together as a metapopulation independent of other such aggregates [2–4]. Regardless of definition, accurate species delimitation is critical, because species are a fundamental unit for much of biology [5,6]. Often species delimitation can be relatively trivial owing to allopatry or prezygotic barriers (e.g. different call types in birds and anurans). In reality, species delineation is frequently obfuscated by the presence of cryptic variation [7,8], and the limitations of many species concepts to effectively recognize such [9]. Geographical structuring of lineages resulting from allopatry may be more common in widely distributed taxa subjected to biogeographic or ecogeographic processes at continental or regional scales. As a result, new species are being increasingly detected [10].

The African continent has a long and pronounced geological history of rift formation, volcanic uplift, desertification and ecological heterogeneity, resulting from climatic cycling [11–13]. As many as 30 African biogeographic realms have been recognized [14], three of which are present in sub-Saharan
western Africa: the Congolian, Upper Guinean and Lower Guinean (i.e. Cameroon–Gabon) [15]. The slender-snouted crocodile (*Mecistops cataphractus*) has a convoluted systematic history [16–18] emblematic of widely distributed crocodilians globally [19,20] and other African taxa [21,22]. *Mecistops* ranges throughout western Africa (figure 1), largely sympatric with the distribution of the other western African endemic crocodile genus *Osteolaemus*. In the light of recent evidence for speciation in other previously recognized species of African crocodiles [19,20], our goal was to use rigorous, multi-locus coalescent and phylogenetic methods and analysis of cranial morphology to test the hypothesis that western African biogeographic zones have driven cryptic lineage diversification in *Mecistops*.

### 2. Methods

#### (a) Taxon and molecular character sampling

We sequenced up to 108 wild-caught slender-snouted crocodiles from throughout its range (figure 1; electronic supplementary material, table S1). We examined sequence variation across 11 gene regions for a total 7768 bp for the *Mecistops* in-group (see electronic supplementary material, table S2)—3768 bp from four partial mitochondrial genes (mtDNA—cytb, 12S, COI, ND4) and 4000 bp from seven nuclear genes (nDNA—LDH-A, rag1 and the flanking regions of five anonymous microsatellites). Primers for mtDNA were designed from a complete *Mecistops* mitogenome (GenBank NC_010639), whereas primers for the nuclear genes were taken from a previously published study [19] or designed from available microsatellite clone sequences [23,24].

#### (b) Data collection

We used the Qiagen multiplex PCR kit at a volume of 15 μl containing 1.1 μl genomic DNA and primers at 0.2 μM. Cycling conditions for all markers were as follows: 95°C 15 min, 35 cycles of 94°C 30 s/55°C 30 s/60°C 1 min; for nuDNA markers 96°C 3 min, 30 cycles of 96°C 30 s/57°C 30 s/60°C 1 min; for mtDNA markers 96°C 3 min, 30 cycles of 96°C 30 s/57°C 30 s/60°C 100 s. We assembled contigs and aligned individual marker datasets in CLC v. 3.6.2. Nuclear strands were phased, and microsatellite repeats were removed from the alignment prior to all analyses.

#### (c) Molecular species delimitation

Coalescent-based Bayesian species delimitation (BSD) may best account for the uncertainty that arises during speciation [25,26]. The method implemented in BP&P v. 2.0 [25,27] uses reversible-jump Markov chain Monte Carlo (rMCMC) sampling [28] to simultaneously estimate the posterior distribution for different speciation models, mutation-scaled effective population sizes (θ) and divergence times (τ). We established a five-taxon guide tree (figure 2a) incorporating two putative *Mecistops* taxa following results from distance- and character-based methods (see below), and three *Osteolaemus* taxa following Eaton et al. [19]. We repeated the analysis with three different, fully partitioned multi-locus datasets: (i) six loci—concatenated mtDNA plus the five msat flanking sequences partitioned; (ii) seven loci—all seven nuclear markers partitioned; and (iii) eight loci—concatenated mtDNA plus all seven nuclear markers partitioned. We assessed the impact of ancestral effective population size and time of divergence on species delimitation by testing three different prior distributions for θ and τ [21]: (i) LD—large ancestral populations, deep divergences, $\theta = \text{G}(1, 10)$ and $\tau = \text{G}(1, 10)$, both with prior mean = 0.1; (ii) SS—small ancestral populations, shallow divergences, $\theta = \text{G}(2, 2000)$ and $\tau = \text{G}(2, 2000)$, both with prior mean = 0.001; and (iii) LS—large ancestral populations, shallow divergences, $\theta = \text{G}(1, 10)$ and $\tau = \text{G}(2, 2000)$. We repeated all analyses under both the 0 (p = 15) and 1 (α = 3, m = 1) rMCMC algorithms [25,28]. We ran three independent chains of 500 000 steps, sampling every fifth step, with 10 000 burn-in steps, for each analysis to confirm convergence on posterior optima. We estimated speciation probabilities for each node in the guide tree following Leaché & Fujita [21] and considered values greater than or equal to 0.95 strong support for the inference of distinct species. For comparison with previous crocodilian systematics studies [19,20], we used similar distance- and character-based methods. Uncorrected p-distances were calculated in MEGA5 [29] for each

---

**Figure 1.** Map of molecular sampling localities. The base map is shaded to reflect topographic and elevation features of the landscape across the sampling distribution. Sample points are colour-coded by corresponding clade (grey, west; white, central), and labels correspond to localities detailed in the electronic supplementary material, table S1. The black, dashed line delimits the distribution of *Mecistops*. The black arrow indicates the Cameroon Volcanic Line. The white, dashed line delimits the distribution of *Osteolaemus*. In the light of recent evidence for speciation in other previously recognized species of African crocodiles [19,20], our goal was to use rigorous, multi-locus coalescent and phylogenetic methods and analysis of cranial morphology to test the hypothesis that western African biogeographic zones have driven cryptic lineage diversification in *Mecistops*.
Figure 2. Molecular and morphological species delimitation results. (a) Bayesian species delimitation results for Mecistops and Osteolaemus assuming a resolved, five-species guide tree. The marginal probabilities for speciation are displayed at each node for each of the three datasets (six, seven and eight loci) and each combination of priors for $\theta$ and $\tau_0$: LD, prior mean 0.1; LS, prior mean $\theta = 0.1$, prior mean $\tau_0 = 0.001$; SS, prior means 0.001. Results from both $\text{rjMCMC}$ algorithm 0 (left of node) and 1 (right of node) are displayed. There was high speciation probability ($\geq 0.95)$ for all nodes under all combinations of dataset, priors and $\text{rjMCMC}$ algorithm, providing robust support for recognition of two $Mecistops$ and three $Osteolaemus$ species. (b) Haplotype network from CHA analysis of mtDNA. Two distinct haplogroups are evident (boxes), one representing all Central African samples and the other all West African samples, separated by 116 mutational steps. Circle size is representative of the number of individuals with each haplotype. Hash marks on branches represent single base changes; branches without hash marks represent only a single base difference between connected haplotypes. (c) NMDS results of multivariate analysis of cranial morphological characters in $Mecistops$. Individuals are colour-coded by group assignment: Central (black) and West (grey). Ellipses represent 95% concentration limits.

Figure 2. Molecular and morphological species delimitation results. (a) Bayesian species delimitation results for $Mecistops$ and $Osteolaemus$ assuming a resolved, five-species guide tree. The marginal probabilities for speciation are displayed at each node for each of the three datasets (six, seven and eight loci) and each combination of priors for $\theta$ and $\tau_0$: LD, prior mean 0.1; LS, prior mean $\theta = 0.1$, prior mean $\tau_0 = 0.001$; SS, prior means 0.001. Results from both $\text{rjMCMC}$ algorithm 0 (left of node) and 1 (right of node) are displayed. There was high speciation probability ($\geq 0.95)$ for all nodes under all combinations of dataset, priors and $\text{rjMCMC}$ algorithm, providing robust support for recognition of two $Mecistops$ and three $Osteolaemus$ species. (b) Haplotype network from CHA analysis of mtDNA. Two distinct haplogroups are evident (boxes), one representing all Central African samples and the other all West African samples, separated by 116 mutational steps. Circle size is representative of the number of individuals with each haplotype. Hash marks on branches represent single base changes; branches without hash marks represent only a single base difference between connected haplotypes. (c) NMDS results of multivariate analysis of cranial morphological characters in $Mecistops$. Individuals are colour-coded by group assignment: Central (black) and West (grey). Ellipses represent 95% concentration limits.

3. Results

(a) Molecular species delimitation

Bayesian species delimitation resulted in unequivocal support (i.e. speciation probability 1.0) for two $Mecistops$
species and highly robust support (i.e. speciation probability greater than 0.95) for three Osteolaemus species (figure 2a). The different prior distributions on $\theta$ and $\tau_0$, different marker datasets, and different rjMCMC algorithms resulted in nearly identical posterior support for speciation events in these taxa (figure 2a). Distance analysis supported the BSD results by revealing two genetic clusters within Mecistops corresponding geographically to Central and West Africa. Mean intragroup divergences were estimated to be 0.0–0.0027 ($\pm 0.0001$), whereas the intergroup distances were 0.0194 ($\pm 0.0046$)–0.0789 ($\pm 0.0083$) depending on the gene region(s) analysed (see electronic supplementary material, table S4). The COI barcoding gap was 4.9% (see electronic supplementary material, figure S1). PAA offered unambiguous support for phylogenetic divisions between these two geographical regions with no shared haplotypes in nine of the genes and 2–77 segregating sites per gene between regions (see electronic supplementary material, table S2). CHA produced networks that clustered all intraregional samples in contiguous regions of the network separated by a single branch substantially longer than any internal branch (figure 2b).

(b) Morphological species delimitation
Comparison of 91 non-juvenile skulls revealed 13 discrete morphological characters distinguishing the two genetically defined West and Central African lineages (figure 3 and the electronic supplementary material provide character descriptions). Of these characters, two showed fixed variation segregating the lineages. Of the remaining characters, one had a fixed state in the Central group (20% frequency in the West group), two had a fixed state in the West group (5–8% frequency in the Central group), eight had the predominantly West trait at a frequency less than 10% in the Central group and six had the predominantly Central trait at a frequency less than 10% in the West group. NMDS detected two discrete clusters of skulls that corresponded to West and Central Africa (figure 2c). The NMDS stress value of 0.17 indicated good fit of the data. All individuals of unknown or questionable provenance were grouped within the 95% confidence ellipses of one of the geographical regions with outliers entirely explained by the degree of missing data and not by locality. The COI fragments from the two skulls confirmed molecular and morphological congruence, and matched geographical expectations.

(c) Fossil-calibrated divergence dating
The topologically unconstrained Bayesian analysis of the combined, partitioned by genome and fully partitioned datasets recovered all expected relationships within the Crocodylia at the species, genus and family levels with most nodes receiving 100% posterior probability support (figure 4; electronic supplementary material, S2–S4). Our results supported three distinct Osteolaemus species, as well as a sister group relationship of Osteolaemus and Mecistops to the ‘true crocodiles’ of Crocodylus. Analysis of the partitioned nuDNA-only dataset recovered all family groupings, as well as monophyletic genera, though it placed Mecistops as sister to a clade consisting of Osteolaemus and Crocodylus, and had difficulty resolving interspecies relationships within Crocodylus (see electronic supplementary material, figure S4). In all analyses, Mecistops consisted of two highly divergent, reciprocally monophyletic groups—one composed entirely of individuals from West Africa and the other of individuals from Central Africa.

Estimated divergence times from Bayesian and maximum-likelihood methods were highly congruent as evidenced by the maximum-likelihood estimation falling within the 95% HPD of the Bayesian analyses (see electronic supplementary material, table S5). The different datasets produced slightly different estimates for the MRCA of most clades (electronic supplementary material, table S5), though this is likely to be best explained by the degree of missing outgroup data in the microsatellite
flanking sequences. We estimated the MRCA of the *Mecistops + Osteolaemus + Crocodylus* clade in the Late Oligocene (± 26 Ma), the *Osteolaemus + Mecistops* sister clade in the Early Miocene (± 21 Ma) and the radiation of extant *Crocodylus* species originating in the Mid-Miocene (± 14 Ma). The split between the two *Mecistops* lineages was estimated at 6.5–7.5 Ma, slightly after the mean estimate for the split between *O. tetraspis* and *O. sp. nov.* in West Africa (7.5–8.5 Ma), though the 95% HPD values for these two estimates overlapped substantially (figure 4).

4. Discussion

(a) Multiple-criteria species delimitation

It has been recommended that multiple criteria (e.g. genetic, morphological, ecological, biogeographic) are used to delimit species [25,50–52]. In the case of *Mecistops*, we present results from three molecular and two morphological species diagnostic approaches that unambiguously support two highly divergent taxa. Further, our speciation model is well predicted by African biogeography with the two taxa isolated in the Congolian (including Lower Guinean) and Upper Guinean biomes. These combined molecular and morphological findings clarify historic taxonomic uncertainties, and the strength of the evidence presented does not warrant subspecies designation or previously suggested subspecies [18].

The Bayesian species delimitation model of Yang & Rannala [25] uses coalescent theory to predict that increasing $\theta$ and decreasing $\tau_0$ will favour fewer species. In the cases of *Mecistops* and *Osteolaemus*, varying the means of the prior distributions by over two orders of magnitude did not impact the speciation inference, and we found unambiguous support for complete isolation and allopatric speciation between the proposed taxa at all $\theta$ sizes and $\tau_0$ ages. Aside from prior assumptions on $\theta$ and $\tau_0$ [52], the primary limitation of BSD is the use of an inappropriate guide tree [21]. The five-taxon guide tree used in our analyses was not susceptible
to this issue as it was well supported by external evidence (i.e. previous studies [19]. African biogeography and results from alternative analyses in this paper), and the seven speciation models explored by the rjMCMC analysis represented all biologically plausible alternative species trees.

BSD results were supported by the more traditional, phylogenetic species delimitation methods. The observed molecular divergence was congruent with that between other recognized pairs of extant crocodilian species (see electronic supplementary material, table S4), and fixed molecular character differences in both genomes can be used to unambiguously diagnose the two Mecistops species [32,33,53]. Partitioned phylogenetic analysis resulted in reciprocally monophyletic lineages that shared an MRCA in the Late Miocene. This estimated divergence timing is congruent not only with sympatric Osteolaemus crocodiles, but also with expectations derived for other crocodilian sister species pairs globally (see figure 4; electronic supplementary material, table S5 and figures S2–S4).

The standard in morphological taxonomy has traditionally been that a single fixed character difference is necessary to differentiate species [6]. Here, we found 13 characters that segregate Central from West Africa, at least 10 of which show frequencies indicative of no contemporary gene flow [6]. Even though this study is the largest examination of intraspecies discrete cranial morphological variation in a crocodilian species, no other study has found as many morphological characters differentiating lineages previously believed to represent a single species. For example, only up to four cranial characters apparently separate Osteolaemus osborni from O. tetraspis, which have substantially higher molecular divergence than seen with the two Mecistops species [19,54,55]. Two skulls analysed as part of this study presented confounding results, unfortunately both listed with provenance of Nigeria. Further examination of the acquisition data revealed that they were purchased from dealers and it is highly likely that the locality data are inaccurate—a common problem with older museum specimens [56].

The strength of the multiple lines of evidence presented here is threefold [2]. First, the multi-inferential molecular approach to species delimitation overcomes the risk of confounding gene trees with species trees [57]. Second, support from fixed morphological characters should convince those sceptical of strictly molecular taxonomy [58,59]. Third, our geographically thorough in-group and taxonomically comprehensive outgroup sampling ensures that the confounding effects of incomplete taxon sampling are avoided [60]. Existing gaps in the in-group sampling will not impact the inference of distinct species [61]. For example, comparison of inter-individual genetic and geographical distances precludes the possibility of isolation by distance effects despite the molecular sampling gap (see electronic supplementary material, figure S5). Undetected genetic variation will, instead, probably be informative on intraspecies phylogeographic process and fine-scale delimitation of species geographical range limits. For example, COI sequences from zoo Mecistops revealed new haplotypes and demonstrated that they all originated exclusively from West Africa (M. H. Shirley 2013, unpublished data). Incidentally, previous studies focusing on the molecular systematics of Mecistops [62] or using multiple Mecistops individuals in phylogenetic and biogeographic analyses [41] never detected multiple Mecistops species because they exclusively sampled zoo individuals.

(b) Divergence timing and the biogeography of Mecistops and Osteolaemus

Divergence times estimated within the crown Crocodylia generally agree with previous results [20,41] and confirm baseline expectations for the MRCA between crocodilian sister species pairs and intracrocodilian species lineage diversification. Virtually all pairs of extant crocodilian sister species shared MRCA in the Mid to Late Miocene (± 6–15 Ma). This establishes a timeline of millions of years as a reasonable expectation for crocodilian speciation [63] and suggests that climatic events driving crocodilian mass extinctions in the Miocene [64,65] also served to diversify tenacious crocodilian lineages through allopatric isolation of populations. Slight differences between our estimated divergence dates and those estimated by others are best explained by our additional calibration points closer to the terminal nodes [66,67], different molecular markers and/or prior hyperparameter specification, resulting in, if anything, biased underestimates of divergence time [68]. The incongruent phylogeny hypothesized by the nDNA dataset is best explained by the inconsistent ability to amplify msat flanking sequences in all outgroup species.

This study is the first to estimate divergence time between all pairs of established and putative sister species from the genera Mecistops and Osteolaemus. Not surprisingly, the estimated divergence timing for Mecistops and Osteolaemus lineages was highly congruent (see figure 4; electronic supplementary material, table S5 and figures S2–S4), indicating a regional vicariant process, and further this timing is congruent with the earliest fossil appearance of Mecistops in Africa, stratigraphically dated to the Late Miocene [62,69,70]. While no samples were available for genetic analysis from Nigeria, our molecular divergence dates and morphological classification of one skull from Nigeria and four skulls from Cameroon to the West and Central species, respectively, and the molecular identification of one Cameroon skull to the Central species, coincide well with the formation of the Cameroon Volcanic Line (CVL) and/or the Benue Trough as the likely biogeographic barrier. Formation of the CVL began in the Late Mesozoic but continued into the Cenozoic, with much of the middle volcanoes and associated highlands (e.g. Mt. Cameroon, Bioko and Bambouto) not arising until the Mid to Late Miocene with final uplifts into the Pleistocene [71,72]. While further molecular sampling may be desired before this biogeography is ultimately accepted, the CVL is a significant geological feature in the region acting as a zoogeographic barrier for many other taxa [20,22,73]. The additional speciation event inferred for Osteolaemus relative to Mecistops in Central Africa is likely to be the result Osteolaemus’s more direct dependence on forest habitat and the cyclical isolation of humid forest cores throughout the Pliocene and Pleistocene.

(c) Implications for crocodilian taxonomy and conservation

Morphological variation thought to characterize a phenotypically diverse species can often be parsed into monomorphic (and even reciprocally monophyletic) groups with the help of DNA evidence and consideration of distribution-wide biogeography and species- or population-level ecology [74]. Despite this, apparent morphological stasis may be a
reasonable expectation even in the face of significant evolutionary change [75]. These ideas are both supported by the divergent genetic diversity and cranial morphologies found in this study and congruent with the prevailing, albeit simplistic, view of crocodiles as morphologically static entities. That we are not the first to find highly divergent, cryptic crocodilian species emphasizes that the crocodile conservation community (and systematists in general) should recognize crocodilians as evolutionarily dynamic species rather than as conveniently binned entities subjected to our limited perception of morphological divergence [76]. Unfortunately, this may not be so straightforward as crocodilian taxonomy often underscores the limitations of many traditional species criteria. For example, Caiman crocodilus and C. yacare are not easily distinguished molecularly [77] or morphologically [78]. Similarly, Crocodylus intermedius and C. acutus, despite being readily distinguished morphologically, may not be genetically reciprocally monophyletic [79]. Further, rampant hybridization has been detected between wild C. acutus and its sympatric congener C. moerletti [80,81] and C. rhombifer [82,83], suggesting that species boundaries between these taxa may yet be porous [26].

Recognizing well-supported, cryptic crocodilian species is not only of theoretical importance but can also have significant implications for species conservation. The slender-snouted crocodile is being evaluated for the 2013 IUCN Red List and may be critically endangered, in recognition of the fact that this species is on the verge of extinction in West Africa and merits considerable efforts to ensure its future [84,85]. By contrast, Central African Mecistops has several robust populations, notably in Gabon [85]. Consistent with the age of taxonomic progress [86], the findings discussed here for African and other crocodilians should encourage the crocodilian systematists and, especially, conservation communities to reflect critically on how we are willing to conceptualize species within the Crocodylia.


Data accessibility. All sequence alignments and the morphological character matrix will be available for download from Dryad six months after the print date of this article or from the corresponding author by request before that time (doi:10.5061/dryad.sh3sm0).

Funding statement. This work was supported by a National Science Foundation Doctoral Dissertation Improvement Grant (DDIG; agreement no. 1010574), Riverbanks Zoo and Gardens Conservation Support Fund, US Fish and Wildlife Service Wildlife Without Borders program (agreement no. 96200-1-G003), Conservation, Food and Health Foundation, Columbus Zoo, Idea Wild Foundation, St. Augustine Alligator Farm Zoological Park, IUCN/SSC Crocodile Specialist Group, AZA Crocodile Advisory Group, Minnesota Zoo, Fresno Chaffee Zoo, San Diego Zoological Society, Mohamed bin Zayed Species Conservation Fund, Cleveland Metroparks Zoo, Oklahoma City Zoo, Aspinall Foundation, Wildlife Conservation Society, WWF, Rotary International and Colin Stevenson. The Rare Species Conservation Foundation served as a zero-overhead funding fiduciary for all grants other than the DDIG.

References


