Zoological Journal of the Linnean Society, 2019, **XX**, 1–30. With 11 figures.

Exploring the effects of character construction and choice, outgroups and analytical method on phylogenetic inference from discrete characters in extant crocodilians

ROLAND B. SOOKIAS

Museum für Naturkunde, Leibniz-Institut für Evolutions- und Biodiversitätsforschung, Invalidenstraße 43, 10115 Berlin, Germany

Received 20 December 2018; revised 28 January 2019; accepted for publication 5 February 2019

Phylogenies for fossil taxa must be inferred from morphology, but accuracy of inference is questionable. Here, morphological characters for extant crocodilians are investigated to assess how to improve inference accuracy. The homoplasy of characters is assessed against a DNA-based phylogenetic tree. Cranial characters are significantly less homoplastic, but this result is perhaps confounded by research effort. Meristic characters are significantly more homoplastic and should be used with caution. Characters were reassessed first hand and documented. Those characters passing tests of robust construction are significantly less homoplastic. Suggestions are made for means to improve coding of discrete characters. Phylogenies inferred using only robust characters and a reassessed matrix, including corrected scorings, were not overall closer to the DNA tree, but did often place the gharial (*Gavialis*) in a position agreeing with or closer to it. The effects of the choice of analytical method were modest, but Bayesian analysis of the reassessed matrix placed *Gavialis* and *Mecistops* (slender-snouted crocodile) in DNA-concordant positions. Use of extant rather than extinct outgroups, even with the original matrix, placed *Gavialis* in a more DNA-concordant position, as did factoring out 3D skull shape. The morphological case for placement of *Gavialis* outside other extant crocodilians is arguably overstated, with many characters linked to skull shape.

ADDITIONAL KEYWORDS: Crocodylia – Archosauria – phylogeny – character construction – morphology.

INTRODUCTION

Phylogeny is the historical branching pattern of evolving lineages (Edwards, 2009), often depicted as a tree (O'Malley & Koonin, 2011). It is through phylogeny that we understand the evolutionary history of life on Earth (Forest *et al.*, 2015). In itself, understanding how taxa originated and relate to each other is of basic interest, because it helps us to place ourselves and other taxa in a wider evolutionary framework (Huelsenbeck *et al.*, 2001). Phylogenies also allow us to investigate a multitude of scientific problems (Huelsenbeck *et al.*, 2001), including the mode and tempo of evolution (Nee *et al.*, 1992; Benson *et al.*, 2018), diversification (Ruta *et al.*, 2007; Stadler, 2011) and extinction (Crisp & Cook, 2009; Volkmann *et al.*, 2014), and their interaction with intrinsic and extrinsic factors (Sookias *et al.*, 2012a, b; Slater *et al.*, 2017). Furthermore, phylogenies are increasingly used in conservation biology (Davis *et al.*, 2010; Forest *et al.*, 2015), allowing conservation efforts to be directed toward species particularly at risk (Davis *et al.*, 2010) or toward species that represent evolutionarily isolated lineages (Collen *et al.*, 2011). Accurate estimation of phylogeny is thus of great importance both from a theoretical and a pragmatic viewpoint.

For most taxa, the true phylogeny is unknown because divergence occurred outside the control and observation of humans (Hillis *et al.*, 1992), often in the distant past. Estimates of this true phylogeny are thus made based on available data. For the inference of phylogeny of extant species, nucleotide and amino acid data are generally used and analysed with varying cladistic and distance-based methods (Yang & Rannala, 2012; Garamszegi, 2014); other types of molecular data and analytical approaches have largely been abandoned (Garamszegi, 2014). These data have the advantage of being increasingly quick and easy to obtain (Metzker, 2010) and objectively scorable

Corresponding author. E-mail: sookias.r.b@gmail.com

(Scotland *et al.*, 2003). DNA datasets are becoming increasingly large, often of whole-genome scale (Metzker, 2010; Pyron 2015; Lyubetsky *et al.*, 2016). There remain many uncertainties in our understanding of the phylogeny of extant taxa, but even for groups that have previously been controversial, larger scale datasets have begun to show signs of converging on a stable topology (Delsuc *et al.*, 2005; Suh, 2016), although there may remain relationships that are irresolvable due to rapid divergence and incomplete lineage sorting (Delsuc *et al.*, 2005; Pyron, 2015; Suh, 2016). Because datasets are increasingly large, and because we have an increasingly good understanding of how to model the evolution of DNA (Delsuc *et al.*, 2005), we have increasing reasons to suppose that the relationships yielded by analysis of DNA data are accurate. These relationships often correspond with biogeography (Stanhope *et al.*, 1998; Yonezawa *et al.*, 2017), providing a second line of evidence to indicate that they may be correct. Furthermore, some experimental studies also exist, which demonstrate DNA-based cladistic phylogenies to be accurate (Atchley & Fitch, 1991; Hillis *et al.*, 1992).

Prior to the advent of molecular systematics in the latter half of the 20th century, morphological data had generally been used to infer the evolutionary relationships of taxa (Brown, 2002). These data are still used, but have been increasingly analysed quantitatively, first using phenetic and now generally using cladistic methods (Brown, 2002; Hull, 2010). Although in many cases the evolutionary relationships inferred from molecular and morphological data agree with each other (e.g. the monophyly of many intermediate-sized bird and mammal clades: Livezey & Zusi, 2007; O'leary *et al.*, 2013; Prum *et al.*, 2015), there are notable cases where they conflict. A striking example is the placement of golden moles and tenrecs with European moles and hedgehogs, respectively, using morphology, whereas molecular evidence firmly places the former two taxa in Afrotheria and the latter two in Laurasiatheria (Stanhope *et al.*, 1998; Madsen *et al.*, 2001; Rasnitsyn, 2006; O'leary *et al.*, 2013). In this case, ecomorphological convergence leads to the morphologically based placement, but molecular evidence agrees with biogeography. Another example is the sister-taxon status of flamingos and grebes consistently found using molecular data, but never based on morphological data, which instead place grebes with the ecomorphologically similar divers and flamingos with the long-legged, wading storks and ibises (Mayr & Clarke, 2003; Mayr, 2004). Although in both cases morphological characters have subsequently been found supporting the placement based on molecular data (Mayr, 2004; Manegold, 2006; Asher & Lehmann, 2008), this placement is not found when all morphological characters are analysed together, at least for mammals (O'Leary *et al.*, 2013).

Although DNA-based phylogenies can be used for modern taxa, by far the majority of taxa that have lived on Earth are now extinct (Wiens, 2004; Benton, 2007) and we can only rely on morphological data from fossils to place these taxa phylogenetically (Wiens, 2004). Given that the phylogenies inferred from morphological data appear in many cases to be further from the true phylogeny than those inferred from DNA, this presents the problem that our phylogenies for these taxa are inaccurate. Not only are the phylogenetic relationships of fossil taxa the subject of continued and strong scientific interest in themselves (e.g. dinosaurs: Baron *et al.*, 2017; Langer *et al.*, 2017), but inference of these relationships underlies a huge body of work examining diversity and evolution through time and thereby our knowledge of the history of life on Earth (Wiens, 2004). Attempting to ensure that these phylogenies are as accurate as possible is thus of great importance.

This study attempts to address the question of how to improve the accuracy of phylogenies based on hard tissue morphology and thereby those using fossils. The group Crocodylia is taken as a test case. Crocodylia is an ancient clade found throughout the tropics (Markwick, 1998) and historically had a wider distribution and much greater diversity than today (Markwick, 1998; Mannion *et al.*, 2015; roughly 23 extant species: Erickson *et al.*, 2012). Crocodylia, together with living birds, Aves, forms the crown group of Archosauria (Gauthier *et al.*, 1988). The archosaurs formed the major larger-bodied component of the land fauna for most of the Mesozoic (>150 million years; Sookias *et al.*, 2012b) and have dominated the skies from some 230 million years ago until today in the form of pterosaurs and birds (Andres & Myers, 2012; Benson *et al.*, 2014). Understanding the relationships of extinct members of this group is thus of special importance, with dinosaur relationships, in particular, the source of recent controversy (Xu *et al.*, 2015; Baron *et al.*, 2017; Langer *et al.*, 2017). There also exist ongoing debates about the placement of extant crocodilian taxa using morphological and molecular data. The gharial – *Gavialis gangeticus* Gmelin, 1789– has generally been considered to be so morphologically divergent from other extant taxa (and variously similar to extinct, non-crown taxa) that a placement outside all other crown crocodilians has been considered most parsimonious using morphology alone (Frey *et al.*, 1989; Tarsitano *et al.*, 1989; Brochu, 1997; Gold *et al.*, 2014). Molecular – from immunological to genomic – data have consistently contradicted this placement, grouping the gharial as the sister taxon to the false gharial – *Tomistoma schlegelii* Müller, 1838– and these in turn as the sister taxon to Crocodylinae (Hass *et al.*, 1992; Brochu, 1997, 2003; Piras *et al.*, 2010; Gold *et al.*, 2014; Green *et al.*, 2014). Another

discrepancy is the placement of the slender-snouted crocodile – *Mecistops*– Gray, 1844 which is shown to be the sister taxon to the dwarf crocodile *Osteolaemus* using multigene DNA analyses (Oaks, 2011; Erickson *et al.*, 2012; analyses based on single genes have also placed it as sister to *Crocodylus* Laurenti, 1768 or in a polytomy with *Crocodylus* and *Osteolaemus* Cope, 1861: McAliley *et al.*, 2006), but it is grouped as the sister to *Crocodylus* using morphology and was long placed in that genus (Brochu, 2003; McAliley *et al.*, 2006). Given these discrepancies, the phylogenetic importance of this group, as well as its relatively limited extant diversity, Crocodylia provides a highly pertinent but manageable group through which to begin to investigate the problem of discrepancy between molecules and morphology.

For the purposes of the work presented here, it is assumed that the phylogeny for Crocodylia based on DNA is accurate. There does, of course, remain the possibility that this may not be the case. As mentioned, morphological evidence continues to contrast strongly with DNA in some respects (e.g. Brochu 1997; Gatesy *et al.* 2003; Gold *et al.*, 2014; Lee & Yates, 2018) and single-gene studies found slightly different placements to larger-scale, more recent work (e.g. McAliley *et al.*, 2006). However, the DNA evidence is of increasingly high quality (Erickson *et al.*, 2012; Green *et al.*, 2014). Regarding the placement of *Gavialis* Oppel, 1811, non-coding (i.e. less convergence-prone) regions also support the same topology and molecular trees are significantly more likely (Harshman *et al.*, 2003), stratigraphic data are concordant with the molecular hypothesis (Lee and Yates, 2018) and some morphological data are also beginning to concur better with DNA than previously (Iijima & Kobayashi, 2019). DNA also yields phylogenies that are more consistent with biogeography and less prone to convergence in major clades bracketing or close to Crocodylia (Pyron *et al.* 2013; Jarvis *et al.* 2014; Crawford *et al.* 2015; Prum *et al.* 2015). Altogether, these facts lead to the assumption that DNA-based inferences are largely accurate. DNA thus provides a second line of evidence, with fewer of the difficulties inherent in morphological data, against which we can compare the results of morphology-based analyses and the effectiveness of different methods. Total evidence analysis is becoming increasingly sophisticated (Gavryushkina *et al.*, 2017) and is an appropriate approach for extant clades. However, total evidence analyses usually yield topologies broadly consistent with DNA (e.g. Brochu, 1997; Gatesy *et al.* 2003; Gold *et al.*, 2014), and for wholly fossil clades other methods are probably needed. Irrespective, that the assumption of accuracy of DNA-based inference has been made should always be borne in mind when considering the work at hand.

In order to investigate which kinds of morphological characteristics best support similar relationships to those yielded by DNA, the fit of different groups of morphological characters from a recent study to a composite phylogeny based on DNA evidence is assessed using three different homoplasy metrics. Comparisons are made between braincase, palate, cranial and other anatomical regions, and between meristic and non-meristic characters. Futhermore, all characters in the dataset are individually assessed, through first-hand observation of specimens, in terms of their robustness of construction using three tests. Those characters passing these tests are used alone to construct a phylogeny, as is a dataset including these characters, alongside rescored and redelimited versions of some of the characters failing these tests. Phylogenetic inference is carried out using parsimony, Bayesian and neighbour-joining methods for comparison, and the effect of using extinct and extant outgroups is assessed along with differences in certainty level of outgroup scoring. A direct attempt is also made to exclude the effects of convergence by weighting/excluding characters based on their correlation with overall skull shape, as measured from 3D cranial scans. The wider implications of findings for inferring phylogenies, and specific considerations regarding the placement of *Gavialis gangeticus* are discussed.

METHODOLOGY

Note on supplement

Supplementary methodological details, the detailed results of all analyses, and full verbal and photographic documentation of all anatomical observations are given in the electronic supplement. This can be downloaded as a zip file and is best accessed through the explanatory file 'Explanation of supplement.docx' in the root folder. Because of its large size, due to high-quality images, the supplement is uploaded to the Open Science Framework repository as the file Supplement.zip (doi: 10.17605/OSF.IO/MGH48), accessible at the following persistent URL: https://osf.io/mgh48/

Morphological character data

The morphological phylogenetic matrix of Narváez *et al.* (2015), based on that of Brochu & Storrs (2012) with many characters illustrated in Brochu (1999a), was used as the source of all morphological characters examined in this analysis. This was chosen because it is one of the most recent morphological analyses of crocodilian relationships and includes nearly all hard tissue and dermal characters used in previous morphological phylogenetic works. This matrix contains 103 taxa, including 16 extant species, and 189 characters. All characters that were uninformative or unscored for the extant taxon set were excluded. All soft-tissue characters

(e.g. muscle characteristics, eye colour) were also excluded because these do not fossilize. Scale characters were not excluded because scales do, on occasion, fossilize and the cut-off between ossified osteoderms and scales is not always clear-cut. This left a total of 117 characters.

Molecular phylogeny

A single phylogenetic tree, based on the most recent molecular phylogenies for crocodilians, was used (Fig. 1A). This tree was based on the phylogeny of Erickson *et al.* (2012), which used the gene database of Gatesy *et al.* (2004) consisting of nine nuclear genes and portions of six different mitochondrial genes. The position of *Caiman yacare* Daudin, 1802 [not included in the analysis of Erickson *et al.* (2012)] was based on the work of Hrbek *et al.* (2008). Although using more limited genetic material (three nuclear and one mitochondrial genes), the position of *Ca. yacare* was strongly supported in the work of Hrbek (2008) and grouped with some populations previously considered *Caiman crocodilus* Linnaeus, 1758; the two taxa were historically in the same species and the only historical taxonomic doubt has been regarding whether *Ca. yacare* is a full species or subspecies (Hrbek *et al.*, 2008). This tree was compatible with the genome-based phylogeny of Green *et al.* (2014), which only included three taxa (*Alligator mississippiensis* Daudin, 1802, *Crocodylus porosus* Schneider, 1801 and *Gavialis gangeticus*), but which represents the only genome-scale crocodilian phylogeny to date.

Homoplasy indices

Three homoplasy metrics were calculated for each character in the matrix using the phylogenetic tree based on molecular data: the consistency index (CI; Kluge & Farris, 1969), the retention index (RI; Farris, 1989) and the number of extra steps (H). CI and RI were calculated using MESQUITE v.3.2 (Maddison & Maddison, 2017). H was calculated using values from MESQUITE using the formula $H = s–(\sigma–1)$, where s is the number of steps on the tree for that character and σ is the number of states for that character present in the matrix. RI has the advantage that it is less strongly affected by the number of states of the character and is thus more suitable for comparing between characters in a dataset (Hoyal Cuthill, 2015). The maximum value of H decreases with the number of characters, but with a taxon sample and maximum number of states of the size used here, this theoretical maximum is much larger than the number of changes observed (Hoyal Cuthill *et al.*, 2010), and the index is thus also unlikely to be affected by the number of states. Indeed, only CI was significantly correlated with the number of states for the dataset and H was the least correlated (see below).

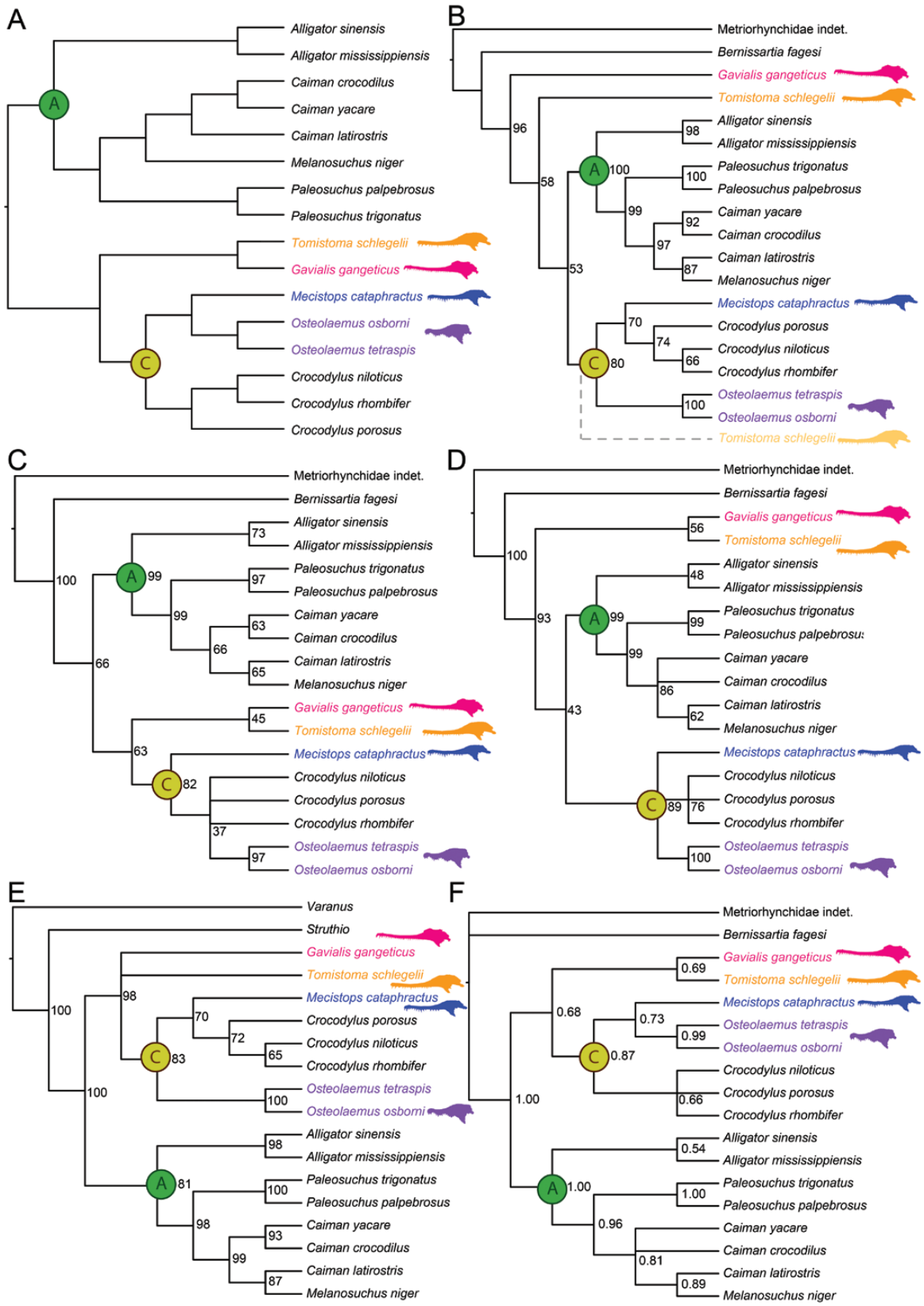
Comparative homoplasy level tests

In order to assess whether the homoplasy of one group of morphological characters when plotted onto to the phylogeny was significantly greater than that of another group, a Mann–Whitney U test was employed as all samples were non-normal (Jarque–Bera and Shapiro–Wilk test, $P<0.01$ – see supplement), with different character groups treated as different samples. Cranial (including mandibular and dental) characters were compared against postcranial characters. Braincase characters, palatal characters, and braincase and palatal characters together were compared against all characters and those from the rest of the cranium (including mandible and teeth). Observable, correctly scored and clearly delimited (henceforth 'robust' – see below) characters were also compared against other characters. Association of number of states with homoplasy indices was examined with Kendall's rank correlation. All tests were conducted in PAST v.3.15 (Hammer *et al.*, 2001).

Categorization of characters as non-robust and robust

To categorize characters as robust, sufficient taxa to examine each state adequately were observed. A character was not considered robust if (see Fig. 2): (1) the variation encoded in the matrix could not be observed at all (i.e. the features described could not be seen in any taxa) or its distribution was completely at odds with the scoring of the matrix (i.e. even if some variation in a feature could be observed and could conceivably correspond to the states described, the distribution was so different from that coded that the coding could not have been intended to refer to this); (2) one or more scorings did not match the observed state distribution; (3) if there did not appear to be reasonable justification for considering the morphology referred to by at least all but one of the states (σ) present in the taxon sample (σ–1; e.g. one state in a binary character, two states in a three-state multistate character, where all three states were present in the taxon sample – if correctly rooted, characters should produce an accurate topology if σ–1 states represent homologous inherited features) to be homologous in each state. All attempts were made to understand character formulation, including contacting their originators where possible, before considering test (1) to have been failed.

Criterion (3) was potentially the most subjective, but consisted of comparing the morphology between taxa in each state and assessing whether the morphology in each state was sufficiently distinct and similar between taxa with the same scoring that the hypothesis of this morphology indicating a shared inheritance was reasonable. The cut-off between the states must be unambiguous and non-arbitrary (i.e. not breaking



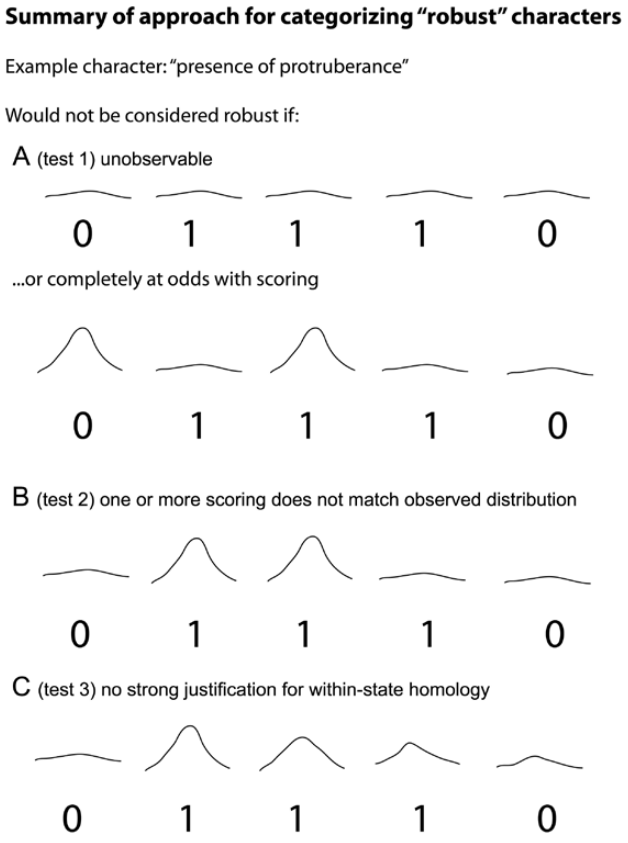


Figure 2. Schematic diagram summarizing how characters were classified as 'robust' and 'non-robust'. To be considered robust, characters had to pass three tests: A (test 1), that the variation indicated was broadly observable and largely corresponded to (was not completely at odds with) the scoring in the matrix; B (test 2), that scoring in the matrix fully corresponded to the observed variation; C (test 3), that the delimitation between states was clear, with at least all but one of the states corresponding to a plausible apomorphy, the separation of which from other states was clear from observing the data (i.e. the state delimitation was not arbitrary or simply based on breaking continuous variation down into equal portions). Characters passing these tests were found to be less homoplastic than other characters and were analysed alone for comparison with the original data.

variation down into equal or convenient portions, but rather with there being a clear discontinuity in the sample), with no taxa in the sample showing an intermediate morphology. All observations and assessments are fully photographically documented and are discussed in the supplementary material to facilitate further assessment by other workers. Yet further photographs and also scan data are available from the author upon request.

Specimens and institutional abbreviations

All specimens used for assessment of characters (including all those photographed) are held in public research institutions or museums. Institutional abbreviations for all specimens referenced in the text and supplement are given below:

FMNH, Field Museum of Natural History, Chicago, USA; MUT SZ, Museum der Univerität Tübingen – Zoologische Sammlung, Tübingen, Germany; MNB, Museum für Naturkunde, Berlin, Germany; NHMUK, Natural History Museum, London, UK (formerly BMNH); NMB, Naturhistorisches Museum, Basel, Switzerland; NMP, National Museum (Národní muzeum), Prague, Czech Republic; RBINS, Royal Belgian Institute of Natural Sciences, Brussels, Belgium; SMNS, Staatliches Museum für Naturkunde, Stuttgart, Germany; TMM, Texas Memorial Museum, Austin, USA; UF, Florida Museum of Natural History, Gainesville, USA; ZMB, Museum für Naturkunde, Berlin, Germany (formerly Zoologisches Museum, Berlin); ZMH, Zoologisches Museum Hamburg, Hamburg, Germany; ZSM, Zoologische Staatssammlung, Munich, Germany.

Phylogenetic analysis

Table 1 outlines the different combinations of matrix scoring, analytical technique and outgroup choice and scoring used in all of the analyses carried out. A full list of all the analyses conducted, with details of the data and analytical protocol, are given in the supplement.

Figure 1. Selected trees summarizing results of phylogenetic analysis undertaken: A, DNA-based tree based on Erickson *et al.* (2012) and Hrbek *et al.* (2008) used as comparator for assessment of accuracy of morphology-based trees; B, tree inferred from all discrete morphological characters of Narváez *et al.* (2015), with unweighted parsimony and two fossil outgroups with certain scoring (see text) – using neighbour-joining and the same dataset with either fossil outgroup, the only difference was in the position of *Tomistoma schlegelii* and this position is indicated with the paler (dashed) branch and outline; C, tree inferred as (B) but using only robust characters; D, tree inferred as (B) but with both robust characters and reassessed previously non-robust characters and two new characters; E, tree inferred in the same way as (B) but using extant instead of fossil outgroups; F, tree inferred using the same dataset as (D), but using Bayesian analysis. Only the original dataset under maximum parsimony (B) failed to recover *Gavialis* and *Tomistoma* as sister taxa and several analysis placed *Gavialis* and *Tomistoma* in positions agreeing with the DNA based phylogeny (tree A). Bayesian analysis of the reassessed dataset (tree F) was the only analysis to place both *Gavialis* / *Tomistoma* and *Osteolaemus* / *Mecistops* in positions concordant with the DNA-based tree. Trees B–E are strict consensus trees and tree F a majority rule consensus. Circle containing 'A' at node indicates base of Alligatoridae, circle containing 'C' indicates base of Crocodylinae. Numbers at nodes are standard bootstrap values (B–E) or posterior probabilities (F).

Hypotheses of relationships were constructed using all characters and only robust characters, using three different methods – maximum parsimony, Bayesian probabilistic inference and neighbour-joining – to allow comparison between methods. Additionally, a new matrix (henceforth 'reassessed matrix') was created, including all robust characters but also some rescored and redelimited previously non-robust characters. Characters that could not suitably be reformulated continued to be excluded. This matrix, with explanations of all rescorings and redelimitations (including two additional characters from within the variation previously referred to by other characters), is given in the supplement.

Maximum parsimony analysis was conducted in TNT v.1.5 (Goloboff & Catalano, 2016), using a traditional search with 1000 replications with 10 000 trees held per replication, followed by a branch and bound (= implicit enumeration) search on the trees from RAM. Other settings were default (e.g. branch swapping, tree collapsing and random seed off). Analyses were undertaken with equal and implied weights, with k (concavity constant) values of 3, 12 and 100. Implied weighting has recently been considered to yield more accurate results than unweighted parsimony or Bayesian analysis (Goloboff *et al.* 2018a – with a k value of 12 found to be optimal) and for this reason it was considered prudent to test this approach; it could be considered that elimination/rescoring of 'non-robust' characters is also a form of elimination of homoplastic morphology and thus implied weighting is superfluous, but it should be remembered that more plausible homology does not necessarily imply lower homoplasy. Neighbour-joining (NJ) trees and NeighbourNet networks were constructed in SplitsTree v.4.14.6 (Huson & Bryant, 2005). Bayesian analysis was conducted using MrBayes v.3.2.6 (Huelsenbeck & Ronquist, 2001), with 1×10^6 generations, 0.25 of the sample discarded as burn-in and a stop value of 0.01 standard deviations of split frequencies. Parsimony and Bayesian analyses were carried out with either fossil crocodylomorph outgroup or extant outgroup pairs, each consisting of a more distantly related taxon set as the outgroup and a less distantly related taxon; extant taxa were constrained to be monophyletic to effectively allow a multi-taxon (and thus more broadly representative) outgroup. NJ analyses were conducted with single outgroups, because extant monophyly could not be enforced. NJ analyses with both taxa in each outgroup pair, and parsimony analyses with single outgroups, were conducted for comparison.

Outgroups

The fossil outgroup taxa used were an indeterminate metriorhynchid MNB P0048 (more distant relative; chosen largely for its convenient location at the MNB, but representing a major non-crown lineage) and *Bernissartia fagesi* Dollo, 1883 RBINS R46 (closer relative). The extant outgroup taxa were the genus *Varanus* Mellem, 1820 (more distant relative) and *Struthio camelus* Linnaenus, 1758 (closer relative). Use of an extant morphological outgroup has precedence in crocodilian phylogenetics (Brochu 1997) and was done here specifically to allow more direct comparison to the phylogeny inferred from DNA. Scorings for outgroups with photographic or scan documentation are given in the supplement. Outgroups were scored in two different manners: (1) with maximum certainty, that is, with as few '?' scorings as possible—taxa were assigned scorings, even if the morphology differed slightly from other taxa with that scoring or if preservation limited (but did not fully prevent) assessment, and new states were created for

Table 1. Table summarizing the different analytical approaches, matrix scorings and outgroups employed. All possible combinations were tested except: (i) any involving the reassessed matrix with uncertain outgroup scoring; (ii) individual outgroups with Bayesian or any of the additional analyses. GPSA-based weighting/exclusion was only tested with *Varanus*, and *Varanus* and *Struthio* outgroups (with certain scoring) and the original matrix scoring. Full results are given in the supplementary information.

Analytical methods	Matrix scoring	Outgroup combinations	Outgroup scorings
Main analyses	Original	*Varanus*	Certain
Parsimony, unweighted	Robust characters only	*Struthio camelus*	Uncertain
Bayesian inference	Reassessed	*Varanus* and *Struthio camelus*	
Neighbour-joining		*Bernissartia*	
Additional analyses		Metriorhynchidae indet.	
Parsimony, implied weights, k = 3, 12 and 100		*Bernissartia* and *Metriorhynchidae* indet.	
Parsimony, inverse weighting by axis 1 of GPSA ordination			
Parsimony, unweighted, characters significantly correlated with axis 1 of GPSA ordination excluded			

morphologies that were clearly homologous with the subject of the character but not with any of the states, and (2) with full uncertainty, that is, scoring the taxon with a '?' whenever there was any plausible doubt as to which state it should be assigned. The uncertain scoring regime precipitated a very large number of inapplicable scorings (93, i.e. 79%, in *Varanus* and 98, i.e. 84%, in *Struthio camelus*; these were scored as missing, because inapplicable and missing are treated the same by TNT), which is a potential hazard of using extant outgroups; with certain scoring these were reduced to 1 and 0 inapplicable scorings for *Varanus* and *Struthio camelus*, respectively. Missing (whether inapplicable or truly missing) scorings were also very high for the fossil outgroups used, both with certain (59 missing, 50%, for *Bernissartia* Dollo, 1883; 52 missing, 44%, for the metriorhynchid) and uncertain (87 missing, 74%, for *Bernissartia*; 81 missing, 69%, for the metriorhynchid) scoring, with preservation a major impediment to scoring fossil outgroups. Both fossil outgroups contained cranial and postcranial material, with *Bernissartia* including more or less the entire skeleton. Using outgroups that are too far from (and too near to) the clade in question has been considered to yield potentially spurious results (Wilberg, 2015) and this criticism could be levelled at the employment of extant outgroups. However, use of extant outgroups was done with the specific objective of comparison to results of DNA-based phylogenies, which by their nature rely on extant taxa. The two fossil outgroups chosen were clearly outside of the crown and represented varied morphotypes (see below). All analyses were conducted with both scorings except using the redelimited matrix, where only certain scoring was used. Robinson–Foulds distances between inferred trees (after removing outgroups) and the DNA-based tree were calculated in R v.3.4.3 (R Core Team, 2014) using phangorn (Schliep, 2011) and number of nodes per tree extracted using ape (Paradis *et al.*, 2004) – R code is given in the supplement.

Factoring out overall skull shape

Attempting to examine whether explicitly removing the effects of convergent evolution on character scoring could yield a topology more concordant with DNA, a dataset composed of a 3D surface scan of each of the extant crocodilian taxa included in the matrix was ordinated following generalized Procrustes surface analysis (GPSA: Pomidor *et al.*, 2016; see Fig. 3). R^2 and P values for each character for these taxa were calculated based on their correlation with the principal coordinates' ordination projection scores for axis one (representing 58.8% of the sample variance – see supplement), using the linear model function in R. Characters were then weighted on a scale of 0–100, inversely to their correlation with the ordination values, and analysed in TNT. Separately, characters showing a significant ($P = 0.01$) correlation with ordination values were excluded and only remaining characters were analysed. Ordination values and R code for returning a string of weights and a list of excluded characters from this data are included in the supplement.

RESULTS

Comparative homoplasy tests

Robust characters are significantly less homoplastic than non-robust characters using all three metrics (CI, $P = 1 \times 10^{-4}$; RI, $P = 1.39 \times 10^{-6}$; H, $P = 3.9 \times 10^{-7}$). Cranial characters are significantly less homoplastic than postcranial characters using CI ($P = 4.23 \times 10^{-7}$), H ($P = 3.91 \times 10^{-7}$) and RI ($P = 1.39 \times 10^{-6}$). Meristic characters are significantly more homoplastic than other characters using H ($P = 0.214$). Other comparisons are not significant at $P = 0.05$ or 0.01 level. The number of states is significantly positively correlated with CI ($P = 1.35 \times 10^{-4}$), but not with RI ($P = 0.0745$) or H ($P = 0.0849$). Full results are given in the supplement in 'Homoplasy index statistics.xlsx'.

Phylogenetic analysis

The results of all phylogenetic analyses (trees as files and images, support values for clades, synapomorphies mapping at nodes, number of most parsimonious trees, etc.) are given in full in the supplement in 'Phylogenetic analyses undertaken, and results.xlsx', with a visual overview in Fig. 1B–F.

Parsimony analysis

Using fossil outgroups and all characters scored with maximum certainty, the same topology is yielded as in the original analysis of Narváez *et al.* (2015; Fig. 1B; analysis 1 in supplement). This differs from the molecular phylogeny (Fig. 1A) in that it places *Gavialis gangeticus* as the sister taxon to all other extant crocodilians (rather than with *Tomistoma schlegelii* as the sister taxon to Crocodylinae), *Mecistops cataphractus* Cuvier, 1825 as the sister taxon to *Crocodylus* (rather than to *Osteolaemus*) and *Melanosuchus* niger Spix, 1825 as the sister taxon to *Caiman latirostris* Daudin, 1801 (rather than outside *Caiman* Spix, 1825). With only robust characters (Fig. 1C; analysis 2), *Gavialis* and *Tomistoma* Müller, 1846 are sister taxa and Longirostres is recovered, approaching the DNA topology more closely, but all *Crocodylus* species form a polytomy with *Osteolaemus,* and *Caiman latirostris* remains the sister taxon to *Melanosuchus*. Analysis of the reassessed matrix

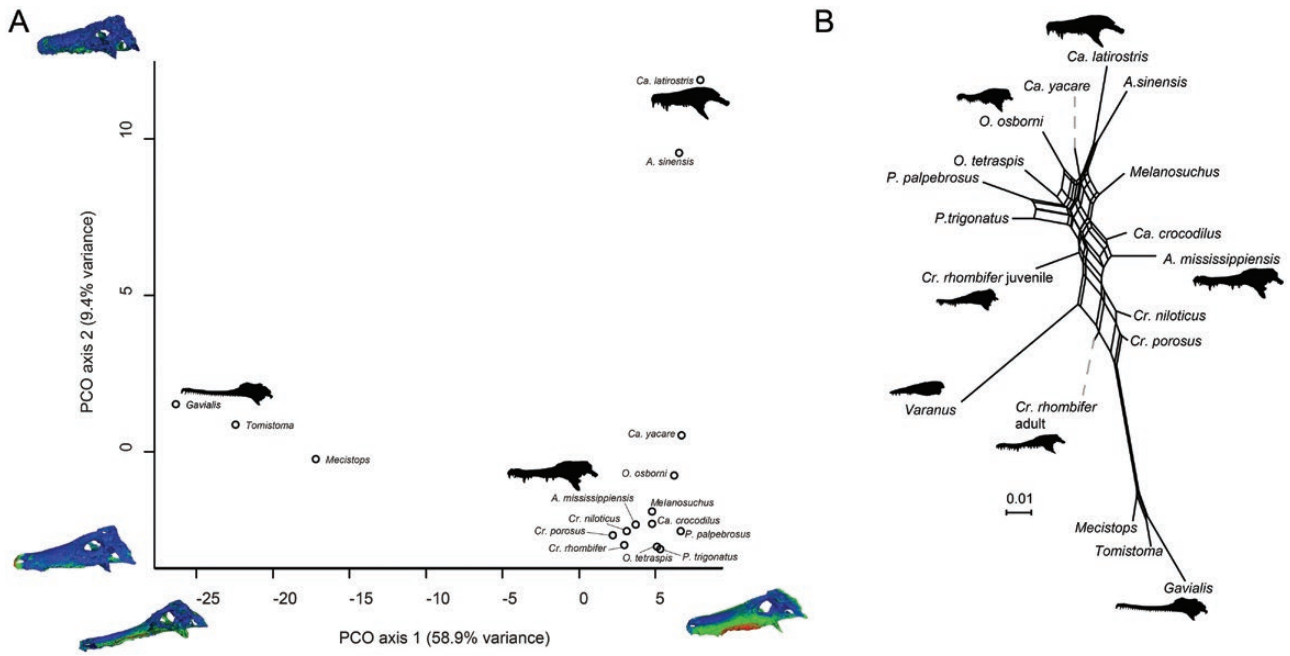


Figure 3. A, plot showing 3D photogrammetric surface scans of crocodilian taxa projected on the first two axes of the principal coordinates ordination space from generalized Procrustes surface analysis (GPSA). A parsimony phylogenetic analysis was carried out with characters downweighted proportionately to their correlation with the first axis of the ordination and separately with characters significantly correlated therewith excluded; the latter analysis was slightly closer to the DNA-based results than the original analysis in that *Gavialis* was sister to *Tomistoma*. Coloured images outside plot are visualizations of the surfaces at the extremes of the ordination axes, with heatmap colours showing most (red) and least (blue) variable regions of the surface along the axis. B, NeighbourNet network diagram showing overall similarity of surface scan shapes to one another, based on a distance matrix from GPSA. Interestingly *Paleosuchus* and *Osteolaemus* group closely together, in turn close to (other) alligatorids and a juvenile individual of *Crocodylus rhombifer* was most similar to dwarf crocodile and alligatorid taxa in overall shape; this demonstrates the paedomorphic processes (in these cases probably postdisplacement) probably underlying dwarf crocodile and alligatorid skull shape (and/or peramorphosis in other taxa). In (A) and (B) silhouettes are the taxa the names of which they are closest to. *A.* = *Alligator*; *Ca.* = *Caiman*; *Cr.* = *Crocodylus*; *O.* = *Osteolaemus*; *P.* = *Paleosuchus*.

(Fig. 1D; analysis 3) places *Tomistoma* and *Gavialis* together as a clade, but as sister to other extant crocodilians; it finds *Crocodylus* to be monophyletic, but does not resolve relationships within the genus and places *Mecistops*, *Osteolaemus* and *Crocodylus* in a polytomy. The node uniting *Caiman yacare* and *Ca. crocodilus* collapses, but *Ca. latirostris* remains sister to *Melanosuchus* Gray, 1862. With fully uncertain outgroup scoring and the original character set (analysis 7), *Gavialis* is placed in a polytomy with *Tomistoma*, Crocodylinae and Alligatoridae; uncertain fossil outgroup scoring (analysis 8) does not affect the topology using only robust characters.

Using extant outgroups, even with the original dataset (Fig. 1E; analysis 4), the clade Longirostres is recovered, but *Gavialis* and *Tomistoma* are placed in a polytomy with Crocodylinae; other relationships are unchanged. With only robust characters (analysis 5), the results are the same as with fossil outgroups, except that the node separating *Mecistops* from other members of Crocodylinae is not resolved. Analysis of the reassessed matrix (analysis 6) is identical to using fossil outgroups except that *Gavialis*+*Tomistoma* are placed as sister to Crocodylinae. Uncertain extant outgroup scoring has no effect on topology using all characters (analysis 9) except that it places *Gavialis* as sister to *Tomistoma*; using uncertain outgroup scoring with only robust characters (analysis 10), the base of the tree is much less resolved, with a basal polytomy including *Gavialis*, both *Alligator* Daudin, 1809 species, *Tomistoma*, Crocodylinae and Alligatoridae.

For simplicity, only results for implied weights with k = 12 (given as analyses 100–109 in supplement) will be reported in full here, as this value was found previously to be optimal (Goloboff *et al.*, 2018a); results with k = 3 and k = 100 (analyses 46–65) are given in full in the supplement and are briefly summarized here. Overall, implied weighting yields fewer, better resolved trees than unweighted analysis, but they are not closer to the DNA phylogeny. Using the original dataset and

fossil outgroups with certain and uncertain scoring, the topology is the same as unweighted parsimony except that *Tomistoma* and *Gavialis* are successive outgroups of Crocodylinae, rather than of all other crown taxa. Using only robust characters and the same outgroups, the phylogeny is the same as unweighted parsimony except that *Crocodylus* is resolved as monophyletic with *Cr. rhombifer* Cuvier, 1807 and *Cr. niloticus* Laurenti, 1768 as sister taxa, rather than all *Crocodylus* species in a polytomy with *Osteolaemus*. Using the reassessed matrix *Caiman yacare* and *Ca. crocodilus* were resolved as successively more distant outgroups to *Caiman latirostris*+*Melanosuchus*, rather than being placed in a polytomy with this clade. With the original dataset and extant outgroups with certain scoring, the tree differs from unweighted parsimony in that *Tomistoma* and *Gavialis* are successively more distant outgroups to Crocodylinae (sister taxa with uncertain scoring), rather than both in a polytomy with Crocodylinae. With only robust characters, *Crocodylus* is resolved as sister to *Mecistops* with *Cr. niloticus* and *Cr. rhombifer* as sisters (or in a polytomy with *Cr. porosus* with uncertain scoring), rather than all *Crocodylus* species in a polytomy with *Mecistops* and *Osteolaemus*, and with uncertain scoring *Gavialis* and *Tomistoma* were in a polytomy with Crocodylinae, rather than sister taxa. With the reassessed dataset, the tree differs from unweighted parsimony in that *Caiman yacare* and *Ca. crocodilus* are sister taxa, rather than in a polytomy with *Caiman latirostris*+*Melanosuchus*. Results with k = 100 were identical except with uncertain scoring and fossil outgroups, where k = 100 yields the same topology as with certain scoring and k = 12. Results with k = 3 differ in several cases for the exact placement of *Gavialis* and *Tomistoma*, the resolution of relationships with Crocodylinae and *Crocodylus* and in Jacarea (caimans except *Paleosuchus* Gray, 1862).

Unweighted parsimony analyses were also carried out with single outgroups (analyses 66–85 in supplement; Fig. 4). Most differences from the two-outgroup analyses involve the placement of *Gavialis* and *Tomistoma*:

(a) With a single, inner fossil outgroup (*Bernissartia fagesi*; Fig. 4A), all characters and certain outgroup scoring, the topology is unchanged from with two outgroups except that *Gavialis* is sister to Crocodylinae+*Tomistoma*, rather than to all extant taxa. With robust characters, the results do not differ from with two outgroups. With the reassessed matrix, *Gavialis*+*Tomistoma* are placed as sister to Crocodylinae, rather than outside the rest of the crown. With uncertain outgroup scoring and all characters, *Gavialis*+*Tomistoma* are placed as sister to *Crocodylus*+*Mecistops*, rather than individually in a polytomy with Crocodylinae and Alligatoridae; with only robust characters there was no difference to with two outgroups.

(b) With a single, outer fossil outgroup (Metriorhynchidae indet.; Fig. 4B), all characters and certain outgroup scoring, the results are the same as with two fossil outgroups. With only robust characters, *Gavialis* is placed as sister to the rest of the crown, instead of together with *Tomistoma* as sister to Crocodylinae. With the reassessed matrix, *Gavialis*, *Tomistoma* and the rest of the crown form a polytomy, rather than *Gavialis*+*Tomistoma* being sister to the rest of the crown. With uncertain outgroup scoring and all characters, *Gavialis* is outside *Tomistoma*+rest of the crown, rather than in a polytomy with these taxa. With robust characters only, *Tomistoma* and *Gavialis* are successive outgroups to the rest of the crown, rather than sister taxa in turn sister to Crocodylinae.

(c) With a single, inner extant outgroup (*Struthio camelus*; Fig. 4C), all characters and certain outgroup scoring, the results differ from those with two extant outgroups in that *Gavialis* and *Tomistoma* form a clade sister to Crocodylinae, rather than being placed separately in a polytomy with Crocodylinae. Using only robust characters, the topology differs in that *Crocodylus* is resolved as sister to *Mecistops* and relationships in *Crocodylus* are resolved, instead of all *Crocodylus* species being placed in a polytomy with *Mecistops* and *Osteolaemus*. With the reassessed matrix, all *Crocodylus* species are in a polytomy with *Mecistops* and *Osteolaemus*, instead of being resolved together as an unresolved clade, and *Caiman yacare* and *Ca. crocodilus* are sister taxa, instead of in a polytomy with *Caiman latirostris*+*Melanosuchus*. With uncertain scoring and all characters, trees do not differ between a single inner extant and two extant outgroups, but with robust characters *Alligator* is resolved (at the base of Alligatoridae), and *Tomistoma* and *Gavialis* are placed as successive outgroups to other crown taxa, rather than all these taxa being placed in a polytomy.

(d) With a single, outer extant outgroup (*Varanus*; Fig. 4D), all characters and certain outgroup scoring, *Tomistoma* and *Gavialis* are successive outgroups to Crocodylinae, rather than placed in a polytomy with the latter. Using robust characters, *Alligator* is in a polytomy with a clade composed of other alligatorids and with Longirostres, rather than at the base of Alligatoridae. With uncertain outgroup scoring and all characters, *Alligator* is placed as the sister to other extant taxa, rather than at the base of Alligatoridae. With only robust characters, *Gavialis* and *Tomistoma* are placed in a polytomy with *Alligator* species and other Alligatoridae+Crocodylinae, rather than together as sister to Crocodylinae.

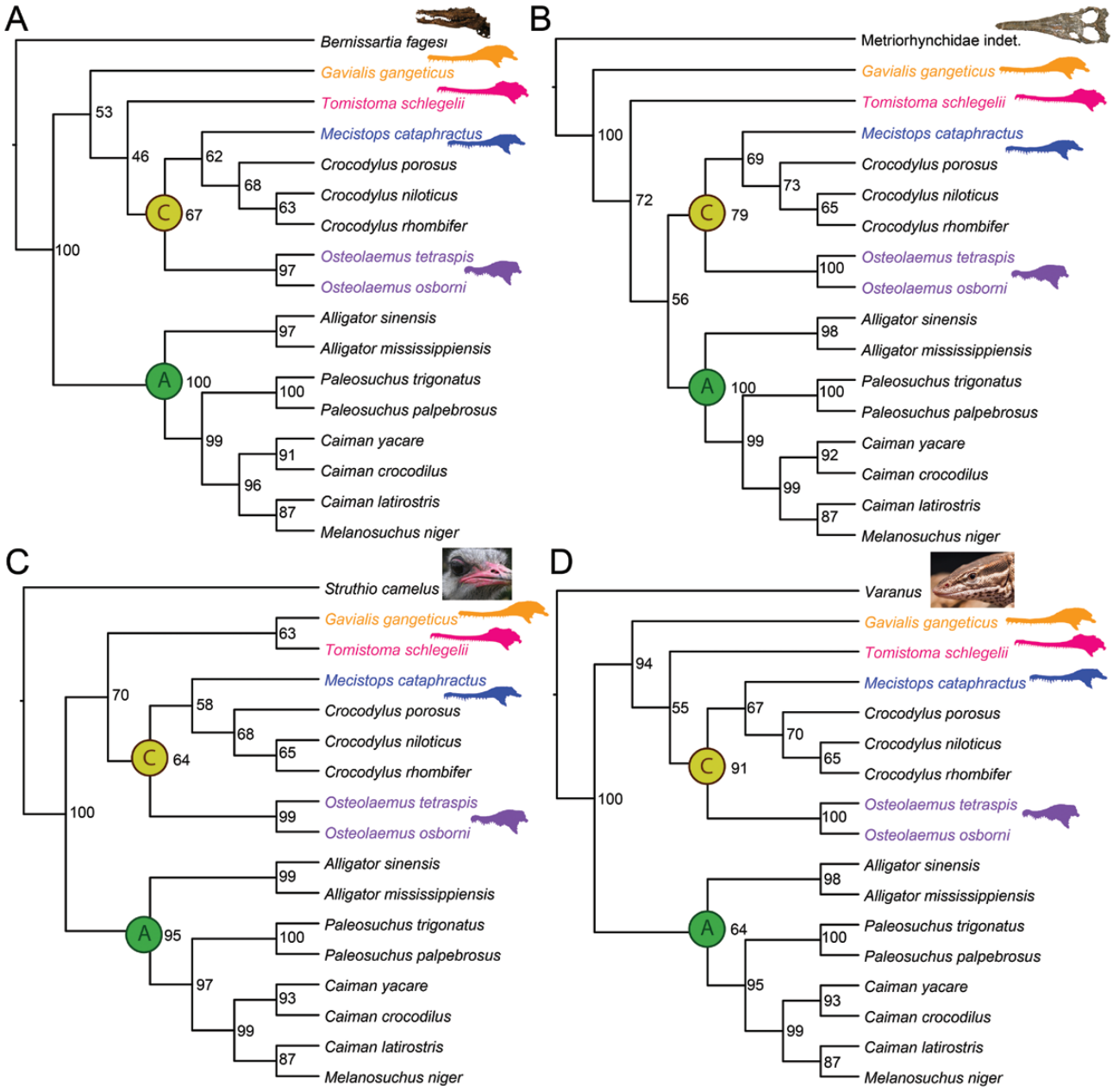


Figure 4. Single most parsimonious trees resulting from unweighted maximum parsimony analyses using single outgroups, using all characters and certain outgroup scoring. Outgroups were: A, *Bernissartia fagesi*; B, an indeterminate metriorhynchid MNB P0048; C, *Struthio camelus*; D, *Varanus.* With extant outgroups (*Varanus*, *Struthio*) or the relatively brevirostrine fossil taxon *Bernissartia* as the outgroup, *Gavialis* is inside the rest of the crown, and is sister to *Tomistoma* with *Struthio camelus* as the outgroup, whereas with the longirostrine fossil metriorhynchid as the outgroup, *Tomistoma* and *Gavialis* are successive outgroups to the rest of the crown. This emphasizes the impact outgroup skull shape and taxonomic distance can have on results. Circle containing 'A' at node indicates base of Alligatoridae, circle containing 'C' indicates base of Crocodylinae. Numbers at nodes are standard bootstrap values. Images of *Struthio*, *Varanus* and *Bernissartia* from Wikimedia Commons.

Boostrap values are given in full in the supplement for appropriate analyses, but are generally relatively high (>70). Nodes with support <70 are those placing *Mecistops* in relation to *Crocodylus* and *Osteolaemus*, those within *Crocodylus* and those placing *Gavialis* and *Tomistoma* in relation to each other and other taxa. With

extant outgroups and with uncertain scoring for fossil outgroups, support values for and within Alligatoridae (excepting for Caimaninae and *Paleosuchus*) drop to <70 in several analyses, reflecting the correspondingly lower amount of certainty regarding character polarity. Decay indices vary between 1 and 15, with *Paleosuchus* and *Osteolaemus*, in particular, remaining well supported (≥9) in all analyses; as would be expected with characters removed, decay indices are generally lower with robust characters only than with all characters. However, they are yet lower with the reassessed matrix, indicating slightly more conflicting signal. Use of uncertain outgroup scoring also reduces decay indices.

Factoring out overall skull shape

No difference in topology is observed between the analysis where characters are inversely weighted by their correlation to their GPSA ordination projection values on axis 1 (analysis 96 in supplement) and an unweighted analysis with a single *Varanus* outgroup (analysis 79 in supplement); *Gavialis* is the sister taxon to *Tomistoma*+Crocodylinae and other relationships are as in Narváez *et al.* (2015) and, thus, also as in an unweighted parsimony analysis with all characters and fossil outgroups (Fig. 1B). Using only those characters not significantly correlated with the first axis of the GPSA ordination, *Tomistoma* and *Gavialis* are sister taxa, in turn sister to Crocodylinae, with other aspects of topology as in Narváez *et al.* (2015).

Bayesian analysis

Bayesian analysis results are analyses 11–20 in the supplement. With all characters and fossil outgroups with certain outgroup scoring, *Gavialis* is sister to *Tomistoma*+Crocodylinae, rather than to all other extant taxa; indeed all Bayesian analyses recover Longirostres either with *Gavialis* or *Gavialis*+*Tomistoma* the most basal taxon. *Crocodylus* species, *Mecistops* and *Osteolaemus* are placed in a polytomy, instead of the resolved relationships of *Osteolaemus*+(*Mecistops*(*Cr. porosus*(*Cr. rhombifer*+*Cr. niloticus*))) under unweighted parsimony. With uncertain outgroup scoring, the genus *Crocodylus* and relationships within it are resolved as with parsimony, but the genus is in a polytomy with *Mecistops* and *Osteolaemus*. Using only robust characters, results differ from unweighted parsimony in that *Gavialis* is sister to *Tomistoma*+Crocodylinae, rather than *Gavialis* and *Tomistoma* being sister taxa (with uncertain outgroup scoring, Bayesian results do not differ from unweighted parsimony). Using the reassessed matrix, *Mecistops* and *Osteolaemus* are sister taxa, rather than in a polytomy with *Crocodylus*, alongside *Gavialis* and *Tomistoma* being placed as in the DNA-based tree, the only analysis to recover these relationships (Fig. 1F). Using extant outgroups (certain and uncertain scoring) and all characters, *Gavialis*+*Tomistoma*, *Mecistops*, *Crocodylus* and *Osteolaemus* form a polytomy, rather than the *Gavialis*+*Tomistoma*+Crocodylinae polytomy with parsimony. Using robust characters and certain outgroup scoring, the tree differs from unweighted parsimony in that *Mecistops* is sister to other Crocodylinae, rather than in a polytomy with *Crocodylus* genera and *Mecistops*. With uncertain outgroup scoring, *Gavialis* is sister to *Tomistoma*+Crocodylinae, rather than to *Tomistoma*. Using the reassessed matrix, *Ca. yacare* and *Ca. crocodilus* are resolved in a polytomy with *Ca. latirostris*+*Melanosuchus*, rather than as sister taxa using unweighted parsimony, relationships within *Crocodylus* were unresolved, and *Mecistops* is sister to *Osteolaemus* rather than to *Crocodylus*.

Posterior probabilities at nodes range from 0.53 to 0.99. The worst supported nodes (in all cases <0.80) are, as with parsimony analysis, those separating or grouping *Tomistoma* and *Gavialis* with each other and other taxa, and those relating to the relative placement of *Osteolaemus*, *Mecistops* and *Crocodylus*. *Alligator* also have relatively low (consistently <0.80, in most cases <0.70) support and the node separating *Caiman*+*Melanosuchus* from *Paleosuchus* receives support of <0.80 in all analyses.

Neighbour-joining analyses

Analyses with an indeterminate metriorhynchid non-crown crocodilian taxon are 21–23, 27 and 28 in the supplement. With all characters, the topology differs from the DNA-based tree in that *Gavialis* is placed as sister to other crown taxa, *Mecistops* is sister to *Crocodylus*, and *Caiman latirostris* and *Melanosuchus* are sister taxa. It differs from the parsimony results in that *Tomistoma* is sister to other Crocodylinae, rather than to Alligatoridae+Crocodylinae (Fig. 1B – position of *Tomistoma* with neighbour-joining indicated with lighter colours and a dashed line). Topology differs using only robust characters, in that *Tomistoma* and *Gavialis* are successive outgroups to other crown taxa. With the reassessed matrix, *Gavialis* and *Tomistoma* form a clade sister to other extant taxa and *Caiman yacare* and *Ca. crocodilus* are successive sister taxa to *Ca. latirostris*+*Melanosuchus*. Uncertain outgroup scoring does not affect topology with all characters, but with robust characters *Tomistoma* and *Gavialis* are successive outgroups to other crown taxa and *Ca. crocodilus* and *Ca. yacare* form a clade once more.

Analyses with *Bernissartia* as non-crown taxon are 31–33, 37 and 38 in the supplement. Using all characters and certain scoring, topology differs from using the metriorhynchid in that *Tomistoma* and

Gavialis are successive outgroups to Crocodylinae. Using robust characters and the reassessed matrix, *Gavialis* and *Tomistoma* form a clade sister to Crocodylinae and in each case *Caiman yacare* and *Ca. crocodilus* form a clade. This same arrangement is found with all characters and robust characters with uncertain scoring.

Analyses with *Varanus* as a single non-crocodilian taxon are numbers 24–26, 29 and 30 in the supplement. With all characters and certain scoring, *Tomistoma* and *Gavialis* are successive outgroups of Crocodylinae and *Mecistops* is sister to *Crocodylus* and *Caiman yacare* and *Melanosuchus*. With only robust characters and the reassessed matrix, *Gavialis* and *Tomistoma* form a clade sister to Crocodylinae, but with the reassessed matrix *Ca. yacare* and *Ca. crocodilus* are successive outgroups to *Ca. latirostris+Melanosuchus*. Uncertain outgroup scoring does not affect the topology using all characters, but with robust characters *Alligator* is placed as sister to other extant taxa.

With *Struthio camelus* as a single non-crocodilian taxon (analyses 34–36, 39 and 40 in supplement), *Tomistoma* and *Gavialis* form a clade sister to Crocodylinae in all cases, except using robust characters and uncertain outgroup scoring, where they form successive outgroups to other crown taxa. *Caiman yacare* and *Ca. crocodilus* form a clade in all cases, except with the reassessed matrix and certain scoring, where they are successive sister taxa to *Ca. latirostris+Melanosuchus*. Other relationships are as in the original matrix of Narváez *et al.* (2015), differing from the DNA-based tree in that *Mecistops* is sister to *Crocodylus* and *Caiman latirostris+Melanosuchus* is a clade.

When both fossil non-crown taxa (*Bernissartia* and the indeterminate metriorhynchid) are included (analyses 86–88, 92 and 93 in supplement), *Gavialis* is the sister taxon to the metriorhynchid in all cases, except with the reassessed matrix and certain outgroup scoring, where *Gavialis* and *Tomistoma* are sisters, in turn sister to Crocodylinae; *Tomistoma* is sister to Crocodylinae using all characters and sister to *Gavialis*+Metriorhynchidae indet. with robust characters (both outgroup scorings). Other relationships are again as in the original matrix of Narváez *et al.* (2015), differing from the DNA-based tree in that *Mecistops* is sister to *Crocodylus* and *Caiman latirostris+Melanosuchus* is a clade. *Bernissartia* is placed between Crocodylinae and Alligatoridae.

When both extant taxa (*Varanus* and *Struthio camelus*) are included (analyses 89–91, 94 and 95 in supplement), with certain outgroup scoring they form a clade in all cases, between Longirostres and Alligatoridae, with *Gavialis* and *Tomistoma* forming a clade sister to Crocodylinae, except with all characters, where *Tomistoma* is one node closer to Crocodylinae. *Caiman yacare* and *Ca. crocodilus* are successive sister taxa to *Melanosuchus+Caiman latirostris* using the reassessed matrix. Other relationships are the same as in Narváez *et al.* (2015). Using uncertain outgroup scoring and all characters, *Varanus* and *Struthio* Linnaeus, 1758 do not form a clade, but are closest to one another and other relationships are unchanged. With only robust characters, *Varanus* is sister to *Alligator* and *Struthio* to *Gavialis* – other relationships are unchanged from using certain outgroup scoring.

Robinson–Foulds index and resolution

The Robinson–Foulds (RF) index of the inferred trees compared to the DNA-based phylogeny are generally similar (column 'R–F distance from DNA' in 'Phylogenetic analyses undertaken and result.xlsx' in supplement; minimum 4, maximum 8, mode 4). Using unweighted parsimony, the RF index is lowest using all characters (mean 4.7), then with the reassessed matrix (mean 5), then with only robust characters (mean 6.25). On average, for analyses including two outgroups, neighbour-joining trees have the lowest RF distance (5), followed by implied weights parsimony (5.2), unweighted parsimony (5.3), and Bayesian analyses (5.6). Using a single outgroup, unweighted parsimony slightly outperforms neighbour-joining (4.95 vs. 5).

The number of resolved nodes in the trees ranges from 8 to 15 (mean 13.65). For unweighted parsimony, the number of nodes is highest using all characters (mean 14), followed by the reassessed matrix (mean 13) and just robust characters (mean 12.67), as expected given the decreasing number of characters. The mean number of nodes resolved by unweighted parsimony is highest (12.7), followed by implied weights parsimony (12.5), and Bayesian analysis and neighbour-joining (both 12.4). Using a single outgroup, unweighted parsimony trees have very slightly more nodes than neighbour-joining trees (mean 13.25 versus 13.2).

DISCUSSION

These results do not present a simple solution to the problem of inferring a more accurate phylogeny from discrete morphological characters. However, various aspects of the results, especially regarding the placement of *Gavialis* and other problems related specifically to extant crocodilian systematics (e.g. placement of *Mecistops*), may provide some indication of the reasons for the discrepancies seen and clues as to how to circumvent them. The results also take the longstanding discussion of morphological support for the placement of *Gavialis* a step further (see discussion of characters below). Furthermore, the results provide

further information regarding the usefulness of characters from particular anatomical regions and of particular types.

Character coding robustness

The significantly lower homoplasy of those characters passing the three parts of the robustness test – (1) observability; (2) correct scoring; (3) biologically plausible state delimitation – suggests that improving our methods of character formulation and scoring may allow more accurate phylogenies to be inferred from morphological characters than hitherto was the case. That characters should be observable and correctly scored is, of course, not controversial, but methods to ensure these criteria are always met should perhaps be more widely employed. For example, the use of verbal scorings, in addition to or subsequently translated to numbers, may reduce errors, because they will be more immediately apparent. Increased inclusion of reference images with scorings should also reduce such errors and allow reviewers to assess observability. These approaches are already beginning to be taken (e.g. O'leary *et al.*, 2013), but there continue to be notable exceptions (e.g. Baron *et al.*, 2017; Langer *et al.*, 2017) and they should become standard practice.

Regarding the third criterion of biologically plausible state delimitation, the fundamental point that should be born in mind is that character states must represent justifiable plausible biological homology (or an 'other' state implying no homology within that state), not simply grouping based on a geometric or other artificial cut-off. For example, characters based on the placement of a foramen on one of two adjacent elements (e.g. character 52, placement of the foramen for the anterior ramus of cranial nerve V on the splenial or dentary – Fig. 5; character 69, lingual foramen for articular artery on surangular/angular suture or only surangular) were homoplastic against DNA data in the dataset examined (assuming that the DNA tree is correct), probably because the cut-off point at the suture is artificial. Although the suture is a natural boundary, in some cases the foramina in taxa differing in which bone the foramen was placed on were actually arguably more similar in overall position relative to the jaw as a whole than in taxa where the foramen was on the same element, or there was no particular similarity in the placement except for the nominal placement on that element (Fig. 5), making the sutural cut-off effectively artificial and unlikely to represent a shared genetic inheritance.

Some tentative suggestions for conservatively coding characters and states to try to avoid false homology statements can be made (see Fig. 6). One is initially to ensure that a clear discontinuity exists between the states as first defined. This is analogous to the idea of gap coding (Mickevich & Johnson, 1976) or simple gap coding (Almeida & Bisby, 1984), but in the absence of quantitative data. If no such discontinuity exists, variation may be better included as a quantitative character, perhaps in a separate or subordinate analysis. Another is to always name a reference specimen in the character description (preferably with photograph or other visual documentation) for each state, with which additional taxa can be compared. The verbal descriptions can then be considered as guides, to help taxonomists identify which aspects of morphology were considered to differ in the taxon sample, and can be appropriately updated, but must not rigidly define the states that are instead

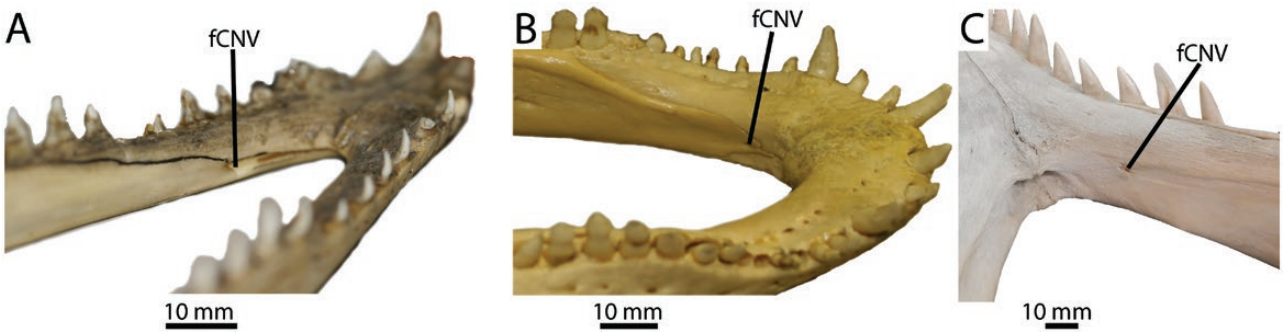


Figure 5. Images showing placement of foramen for cranial nerve V (fCNV): A, *Crocodylus niloticus* MUT SZ Rept. 80 mandible in right dorsomedial and slightly posterior view; B. *Alligator sinensis* SMNS 4915 mandible in right dorsomedial view; C, *Gavialis gangeticus* ZMB 36678 right mandible in left dorsomedial and posterior view. The split between the two character states of character 52 in the matrix of Narváez *et al.* (2015) is placed at the border of the splenial, with this foramen either on the splenial or elsewhere (on dentary or at the splenial-dentary border). However, the biological justification for this cut-off is not clear, with, for example, the placement in *Alligator sinensis* not overall more similar to that in *Gavialis* (which is scored the same) than to *Crocodylus*, except that the foramen is nominally on the splenial.

defined based on the specimens. This should prevent focus on nominally 'fitting' taxa into states and rather divert focus to ensuring plausible homology statements. These considerations relate to the idea of continuously updating character states when adding new taxa, but again may be easier to employ in the absence of quantitative data.

Even if the morphology of a taxon newly scored for a character fits the state description (e.g. X shorter than Y), if—because it differs in some other noticeable way from the reference specimen—the homology of this morphology is doubtful, it should not receive the same scoring. For example, imagine a character that is defined based on a process X, which is shorter (state 0) or longer (state 1) than process Y (Fig. 6A). In the initial delineation, there should ideally be a clear distinction between the states, with a noticeable discontinuity in the sample, and each state should be assigned a reference specimen (α, β, Fig. 6). On scoring a new taxon (yellow panel, Fig. 6A), even if in terms of length it fits nominally in one of the two state categories (in this case state 1, represented by reference specimen β), if the morphology does not fall within or very close to (such that the clear discontinuity is maintained) the range of other taxa with that scoring/is not clearly closer to one of the two reference specimens, then it should not receive the same scoring because homology is doubtful. Instead the character must be either redefined to be based on some other part of the variation or the new taxon should receive another state (in this case 2) to avoid potentially incorrect homology statements.

The situation is further complicated when another aspect of morphology covaries (Fig. 6B), in this case the process shape – blunt or pointed. If the character is defined based on one aspect of morphology (in this case length) and a new taxon (Fig. 6B, case a in yellow panel) is added that is not clearly (only nominally, i.e. it is not in fact markedly closer to either reference specimen or sample group) assignable to either state based on this, but does appear to be closer to one of the scorings based on other aspects of morphology combined with this (in this case to state 1, because of its pointed shape), use of a reference specimen would mean that the fundamental homology statement (homologous with α or β – with the new taxon assigned as homologous to β) can remain intact and only the verbal character description needs to be extended/rewritten to flag up the pointed/blunt dichotomy along with length. If the new taxon nominally fits one of the two states (case b – a long process, and thus state 1), but the morphology is in other respects (here shape) so different that homology is doubtful, then I would suggest that the new taxon should receive a new scoring (here 2) to avoid a false homology statement; again, the reference specimens can remain the

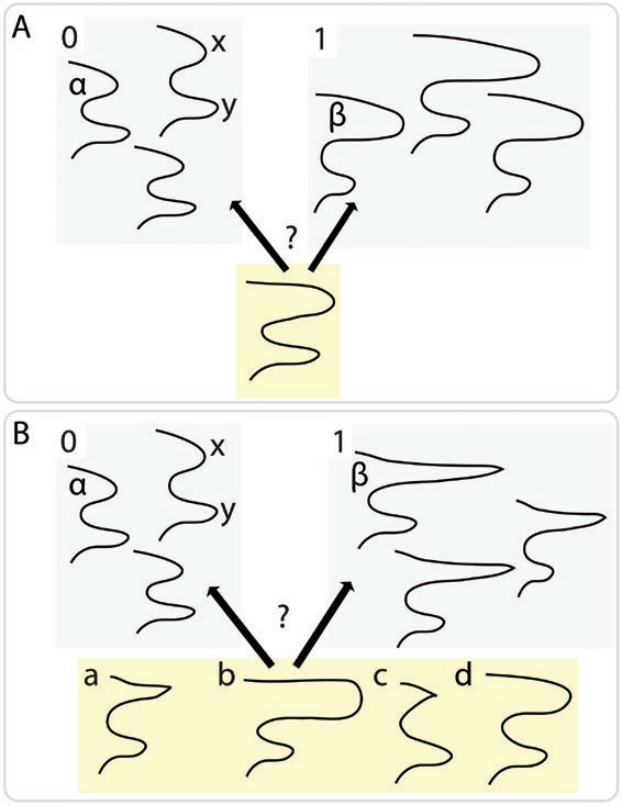


Figure 6. Schematic diagram illustrating difficulties encountered when assigning a state scoring to a new taxon and ways of conservatively coding, using reference specimens, to avoid false homology statements (see text for full explanation). States 0 and 1 (grey panels) are verbally defined based on the length of process X (0 shorter, 1 longer) compared to process Y. In A, the only difference between the states is length; in a conservative scoring approach to ensure false homology statements are avoided, when a new taxon (yellow panel) showing an intermediate morphology is added to the matrix, it should receive a new state, rather than be 'forced' into one of the existing states, but reference specimens (α, β) for states 0 and 1 can remain. In B, the morphologies also differ in shape (pointed versus blunt). If the states as defined (length, in this example) now do not place the new taxon based on a clear discontinuity but only nominally (case a, yellow panel), use of a reference specimen allows the homology statements to be maintained, with only a rewording of the description (to include shape in this case) necessary. If the feature of the new taxon nominally fits one of the two states, but is very different in other aspects of morphology (case b, yellow panel), it is considered prudent to create a new state and verbally redefine the other states (in this case to include shape, but maintaining the reference specimens), rather than try to force it into either state. If a new taxon (cases c, d) fits one reference taxon/sample in terms of one aspect of morphology (here length), but the other reference taxon/sample in another (here shape), it is suggested to create a new state, or define the character based on the more information rich (complex) part of morphology.

same (with the addition of one for state 2), but with appropriate rewording of the character description to aid scoring. It could also be argued that the character could be broken down into separate length and shape characters, with only the shape character requiring a new state for the taxon. However, the evidence for homology within the states of the (lower information - see below) length character may be less strong than with both aspects combined. If the new taxon nominally fits one state (cases c – nominally it fits state 0, shorter X than Y – and d – nominally fits state 1), but has another aspect of morphology that is more similar to the other state (pointed/round shape), then I would suggest that it either receives a new state scoring to avoid false homology, or the aspect of morphology that is arguably more complex (in this case 2D shape rather than 1D length) replaces the old wording (but, again, the reference specimens can remain the same). The other alternative would be to create two characters (shape, length) and this is often the approach taken. A thorough assessment of this coding strategy is beyond the scope of this work, but based on my observations in the sample at hand I would voice that I fear that this would precipitate false homology statements, because more information-rich features seem more likely to represent true homology.

Despite the lower homoplasy of more robustly constructed characters, indicating that morphological data may be broadly compatible with DNA (and that the DNA-based tree is likely to be correct), the phylogenies inferred using only characters passing these tests were not, on average, closer to the DNA-based phylogeny when compared as a whole, in part because these trees were, on average, slightly less resolved due to the lower number of characters. Even when excluded characters were rescored and redelimited to attempt to include any useful phylogenetic information contained in them, the phylogenies inferred were still no closer, on average, to that inferred from DNA. However, the relatively small size of the dataset, especially regarding the taxon sample, and the general similarity in terms of RF indices of all trees, makes drawing firm conclusions using such overall metrics potentially misleading.

Although robust-only and reassessed matrices did not represent an improvement on the original matrix, when comparing the whole tree using RF distances, these datasets did yield some relationships closer to, or the same as, those inferred from DNA, and which differed from the original dataset. Robust-only and reassessed datasets consistently place *Gavialis* either as the sister taxon to *Tomistoma* or in a polytomy including *Tomistoma*. Given that morphological analyses excluding stratigraphic data (both cladistic and descriptive: Frey *et al.*, 1989; Tarsitano *et al.*, 1989; Gold *et al.*, 2014; Lee & Yates, 2018) have often failed to support this relationship, this is noteworthy. Furthermore, using Bayesian methods, Longirostres *sensu* Harshman *et al.* (2003), i.e. a clade including Crocodylinae, *Tomistoma* and *Gavialis*, was consistently recovered. Furthermore, *Mecistops*, the slender-snouted crocodile, was placed in a polytomy including *Osteolaemus*, the dwarf crocodile (its sister taxon based on DNA: Erickson *et al.*, 2012), in unweighted parsimony analyses using the reassessed matrix, and with Bayesian analysis was placed as the sister taxon to *Osteolaemus*, agreeing with molecular data; thus the Bayesian analyses carried out here using the reassessed dataset, both with fossil and extant outgroups, represents the first morphology-based analyses to place both *Gavialis* and *Mecistops* in positions agreeing with molecular data and are as close overall (RF distance = 4) to the DNA-based tree as the original dataset. However, the grouping of caimans was still at odds with the DNA-based tree, with *Caiman yacare* and *Caiman crocodilus* not forming a clade (unlike using the original matrix), but rather in a polytomy with *Caiman latirostris* and *Melanosuchus*.

The effects of convergence

Among the characters that appeared to be accurately scored and clearly delimited, the major reason for homoplasy against the DNA phylogeny appears to be convergent evolution. This is clearly the case for several characters shared between the dwarf caiman *Paleosuchus* and dwarf crocodile *Osteolaemus*, including a dorsoventrally, very deep parabasisphenoid exposure on the braincase in posterior view (original character 173, also shared with all alligatorids – see below), and the shape of the lateral projection of the palatines into the suborbital fenestrae (original character 117). Furthermore, caimans as a whole, and to a lesser extent alligatorids, share features with *Osteolaemus*, and these taxa group together, and with a juvenile *Crocodylus rhombifer*, in terms of overall skull shape similarity (Fig. 3B). Postdisplacement has been shown to have occurred in cranial ontogeny at the base of all of these groups (Morris *et al.*, 2019), yielding a 'paedomorphic' skull shape, and morphological convergences beyond overall skull shape itself may well be connected to this ontogenetic shift. Similarly, the development of 'peramorphic' longirostry, either through acceleration or predisplacement (Morris *et al.*, 2019), has likely precipitated many convergent characters in *Mecistops* and *Gavialis*/*Tomistoma* (see: Gatesy *et al.*, 2003). *Crocodylus* can be considered intermediate between these extremes, with reduction in dental differentiation compared to more 'paedomorphic', shorter snouted forms and with an intermediate braincase height.

Eliminating the effects of convergence continues to be, arguably, the greatest challenge in morphology-based phylogenetics. With the current dataset, when attempts were made to take convergence into account by inversely weighting characters by their association with overall skull shape (axis 1 of the principal coordinates analysis of cranial surface; Fig. 3A), the same topology was recovered. When characters significantly correlated with cranial shape were excluded from the analysis, *Gavialis* was placed as the sister taxon to *Tomistoma*, in turn sister to Crocodylinae, which is the same as its position based on DNA, but other relationships were unaffected. Factoring out overall shape in this way is thus no panacea. The relatively small sample probably impedes accurate correlations between skull shape and character scorings. Furthermore, a large proportion of cranial shape *is* inherited, and thus characters removed could be phylogenetically informative, although this does not appear to have been a problem here, given that closely related taxa with similar skulls still grouped together. Another approach would be to explicitly exclude ecology, but because inferences regarding ecology of fossil taxa often rest on the morphology itself, factoring out ecology more explicitly is challenging. Focusing on particularly informative regions (see below) may be more fruitful.

The effect of analytical techniques

The small taxonomic scale of the current study means it provides limited data on the relative usefulness of different analytical techniques. Overall, none of the methods tested was strongly divergent in results yielded, and data appear to make more difference than analytical method. Bayesian analyses were on average slightly further from the DNA-based topology than parsimony; this conflicts with data indicating that Bayesian methods outperform parsimony in terms of accuracy, especially using small, non-consistent datasets with missing data (O'Reilly *et al.*, 2016, 2018; Puttick *et al.*, 2018), but the very small difference (0.6 RF units), the fact that under some circumstances (e.g. low amount of missing data – Wright & Hillis, 2014; low homoplasy – O'Reilly *et al.*, 2018; differing mechanisms of evolution between characters – Goloboff *et al.*, 2018b) Bayesian analysis does not yield an improvement, and evidence actually indicating parsimony may outperform Bayesian methods (Goloboff *et al.*, 2018b; Sansom *et al.*, 2018), this is less surprising. However, the number of nodes was slightly higher using parsimony, which is what might be expected given indications that parsimony trees tend to be more resolved (O'Reilly *et al.*, 2016). Bayesian analyses with the reassessed dataset remain the only method to have correctly placed both *Gavialis* and *Mecistops*, but the same result was found in two of the four most parsimonious trees using parsimony; in the strict consensus this relationship was thus unresolved, as probably would have been the case in the Bayesian tree, if it was constructed using a strict, rather than majority rule, approach. Although it could be considered a 'phenetic' method because it is based on overall similarity (although still, through the matrix, on homology statements), neighbour-joining generally yielded extremely similar results to parsimony and Bayesian analyses, with the average RF distance from the DNA-based tree the lowest of the methods implemented. Inability to enforce topological constraints in SplitsTree (Huson & Bryant, 2005) led to spurious placement of outgroups in several NJ trees, but otherwise results were broadly very similar to other methods. Similar results were evidenced when topologies from phylogenetic morphometrics and NJ were compared (Catalano & Torres, 2017), but this may not, however, be the case with larger discrete character datasets and further investigation is required.

The placement of *Gavialis*

The phylogenetic placement of the gharial *Gavialis* has been a subject of controversy at least since molecular data began to indicate that it was the sister taxon to the false gharial *Tomistoma*, rather than the sister taxon to all extant taxa closer in some ways to extinct 'mesosuchians' (i.e. a paraphyletic grade of non-eusuchian mesoeucrocodylians, for which Tarsitano *et al.*, 1989, used the teleosaurid *Steneosaurus bollensis* Jaeger, 1828 as an example) than other extant forms (Tarsitano *et al.*, 1989). Some authors have argued that there is in fact relatively little morphological evidence contradicting the molecular data and that the two datasets are consistent (Buffetaut, 1985), while others considered that morphological data unequivocally placed the gharial outside the rest of the crown and linked it to 'mesosuchian'-grade fossil taxa (Frey *et al.*, 1989; Tarsitano *et al.*, 1989). Combined analyses have supported molecular results (Brochu, 1997; Gatesy *et al.*, 2003; Gold *et al.*, 2014), but morphological character partitions (Brochu, 1997; Gatesy *et al.*, 2003) and a geometric morphometric analysis of the braincase (Gold *et al.*, 2014) have continued to place the gharial as sister to other extant taxa. However, a secondary phylogenetic signal has been recovered among morphological characters supporting the molecular phylogeny (Trueman, 1998), although whether this is recovered depends greatly on the dataset used (Brochu, 1999b). Furthermore, including stratigraphy along with morphological characters has yielded a phylogeny more compatible with DNA (Lee & Yates, 2018) and recent morphological evidence has been found arguably bridging that gap between *Gavialis*

to *Tomistoma* (Iijima & Kobayashi, 2019; although parsimony analysis still recovered the 'traditional' morphological topology). The results presented here indicate that when characters with scoring mistakes and lacking strongly plausible biological homology are removed, discrete morphological characters are in fact compatible with DNA data regarding *Gavialis*. This placement is yet more strongly supported following rescoring or redelineation of some characters, where scoring and state delimitation were deemed problematic. This corresponds to the observations of Harshman *et al.* (2003) that character formulation may, to some extent, be responsible for the discrepancies seen between morphological and molecular phylogenies. To allow others to easily examine and test reassessments made, these features will be discussed individually:

(a) ***Capitulum of anterior sacral rib* (Fig. 7; original character 22).** In the matrix of Narváez *et al.* (2015), *Gavialis* is scored as showing a capitulum of the anterior sacral rib exposed in dorsal view, which is the alligatorid condition, while *Tomistoma* is scored as showing the capitulum obscured by the tuberculum in dorsal view. On observation, the homologization of the form in *Tomistoma* with that in Crocodylinae was found to be implausible. Although the medialmost portion of the capitulum is obscured, laterally it is exposed. Furthermore, the anterior margins of the capitulum and tuberculum are sharply defined. This is the case in alligatorids and *Gavialis*, but contrasts with the smooth, rounded margin in Crocodylinae. For this reason, *Tomistoma* was rescored to the same state as *Gavialis* and alligatorids (0).

(b) ***Number of dorsal osteoderms in longest rows* (Fig. 8A, B; original character 40).** In the original matrix, *Tomistoma* was scored as showing six contiguous osteoderms in its longest rows, whereas *Gavialis* was scored as showing four. On observation, this distinction could not be seen. The lateralmost osteoderms (i.e. third from midline) in *Tomistoma* are rounded, their keels directed at an angle to the midline and they do not abut the medially adjacent row along their whole lengths. In *Gavialis* the arrangement is very similar and, if anything, the lateral rows are *more* integrated, abutting the medially adjacent row along almost the entire medial length of each osteoderm. Given the general similarity in arrangement, *Tomistoma* was rescored to the same condition as *Gavialis*.

(c) ***Arrangement of nuchal osteoderms* (Fig. 8C–E; original character 41).** In the original matrix, *Tomistoma* was scored as showing eight nuchal (nape) osteoderms in two parallel rows, while *Gavialis* was scored as showing a nuchal osteoderm shield grading into the dorsal shield (i.e. no distinct nuchal osteoderms). On observation, the nuchal shield of *Tomistoma* was also found to not be clearly separated from the dorsal shield and the taxon was thus rescored to the same condition as *Gavialis* (0). While there were indeed eight osteoderms in *Tomistoma* that were distinguishable from the more posterior shield by the reduced mediolateral width of the more posterior osteoderms of the eight, there was no noticeable gap between the most posterior row and the dorsal shield. *Gavialis* also showed a nuchal region of osteoderms, although it graded into the dorsal shield, consisting of large median osteoderms reducing in width posteriorly, bordered by smaller lateral osteoderms. The anteriormost lateral osteoderms slightly posterior to the first row of medial osteoderms and the fourth row back was bordered by two, rather than one, lateral osteoderm. This rescoring was considered expedient because the nuchal shield of *Gavialis* was not more similar to that of the fossil outgroup *Bernissartia fagesi* than to *Tomistoma*, whereas the lack of a clear separation between nuchal and dorsal osteoderms in *Tomistoma* and *Gavialis* is plausibly a plesiomorphy shared with *Bernissartia*.

(d) ***Follicle glands in ventral scales* (Fig. 8F–H; original character 44 in list, 46 in matrix).** In the original matrix *Gavialis* was scored as absent (0) for possession of follicle glands on its ventral scales (the alligatorid condition; using scorings from char. 46 in the matrix, with which it seems to have been confused). This is incorrect, with follicle glands clearly observable, as in *Tomistoma* and Crocodylinae.

(e) ***Palatine anterior processes* (Fig. 9; original character 116).** The anterior ends of the palatines are scored as tapering to form a thin wedge in *Gavialis* and *Mecistops* (0), whereas other taxa are scored as having anteriorly broad palatines (1). On observation, the palatines of *Tomistoma* were also found to taper anteriorly and the taxon was rescored to state 0.

(f) ***Quadrate medial hemicondyle* (Fig. 10; original character 181).** The medial quadrate hemicondyle was scored as 'small' and 'ventrally reflected' in *Gavialis*, whereas *Tomistoma* was scored as having an 'expanded' hemicondyle, as were all members of Crocodylinae. On observation, the hemicondyles of *Gavialis* and *Tomistoma* were extremely similar, both being relatively small compared to the size of the lateral hemicondyle, anteromedially slanted and with the distinct articular surface forming a raised border (in posterior view) with a pronounced proximal expansion of this border at the interface between posterior and medial surfaces of the quadrate. In Crocodylinae this border is less clear, and the hemicondyle is larger and less anteriorly

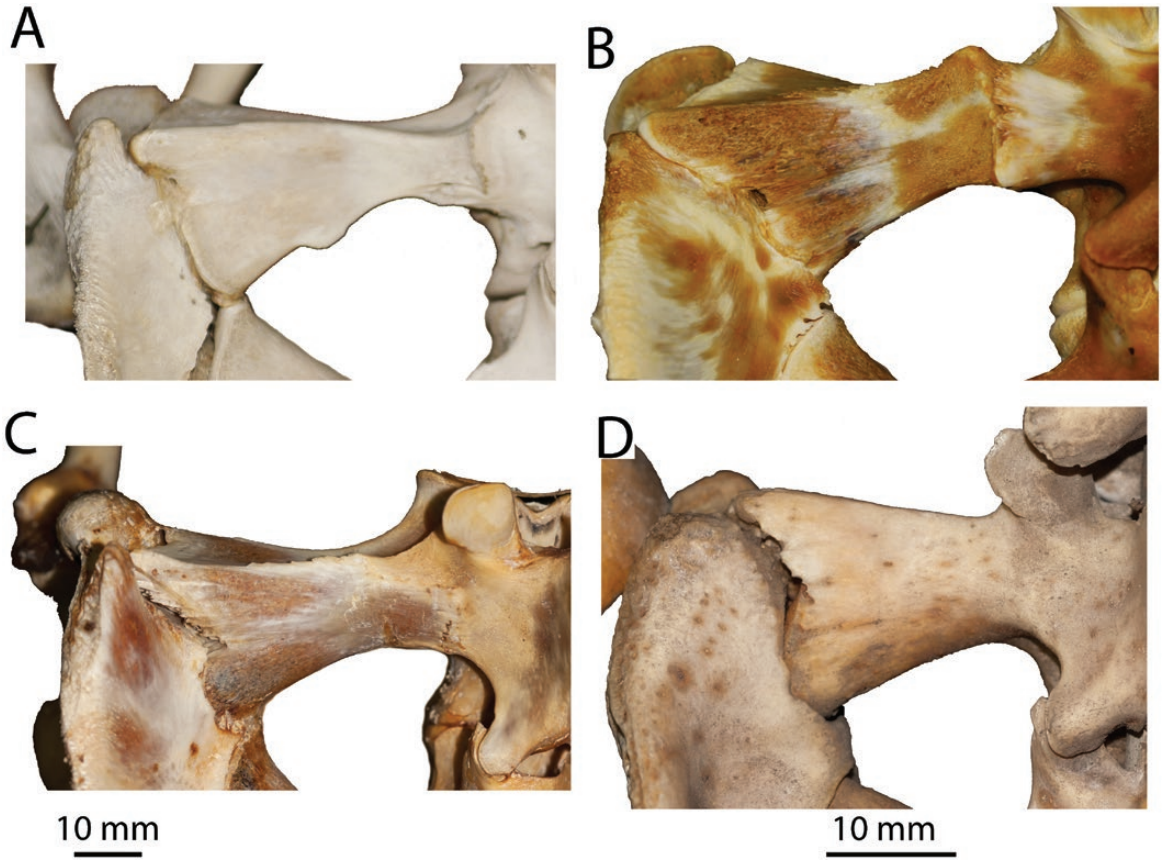


Figure 7. Left anterior sacral rib in dorsal view of: A, *Gavialis gangeticus* MUT SZ Rept. 53; B, *Tomistoma schlegelii* NMP6V 73906/1; C, *Caiman latirostris* NHMUK 1886.10.4.2; D, *Osteolaemus tetraspis* NHMUK 1983.1130. Contra the scoring of Narváez *et al.* (2015), *Tomistoma* showed a morphology closer to the alligatorid condition. Scale bars = 10 mm, that below C applies to A–C, that below D applies to D.

deflected. This deflection is even less pronounced in alligatorids, with no clear distinction between posterior and medial surfaces.

Assuming these differences from the original scoring do indeed represent improved observations, these discrepancies serve to illustrate the impact that individual characters can have on the outcomes of phylogenetic analyses and underline the point that, unfortunately, errors and biologically implausible state delimitation will often not get cancelled out, even in large datasets. Indeed, the most controversial phylogenetic placements often rest on relatively few characters, even using very large matrices (e.g. major dinosaur clades: Baron *et al.*, 2017; Langer *et al.*, 2017); thus, it is essential that *all* characters are accurately observed and plausibly scored. If this must come at the expense of dataset size, this may be preferable.

However, it must be noted that the current analysis made no attempt to include a fossil sample, except the outgroups, and doing so may change the conclusions drawn. State delimitation plays a major role and observation of fossil taxa may precipitate different conclusions about appropriate state delimitations. Harshman *et al.* (2003), for example, underlined how scoring of some snout characters (e.g. degree of splenial participation on the symphysis – character 54 here) using delimitations based on fossil groups may have created fewer putative synapomorphies of *Gavialis*+*Tomistoma* than otherwise. Equally, many characters (e.g. 'square', highly angular dorsal osteoderms) that support a *Tomistoma*+*Gavialis* grouping with extant taxa alone were found to be more widespread in fossils and to have been lost in other taxa, and thus presumed to have been lost in extant Crocodylinae and Alligatoridae (Gatesy *et al.*, 2003). Furthermore, fossil gharial-like eusuchians, for example, the thoracosaurs, were not included, which generally group with *Gavialis* using morphological characters alone (Brochu, 1997; Lee & Yates, 2018). It is possible, even following reassessment of the characters uniting them, that they would still be placed together. As discussed below, *Gavialis* shared several features with the longirostrine, aquatic

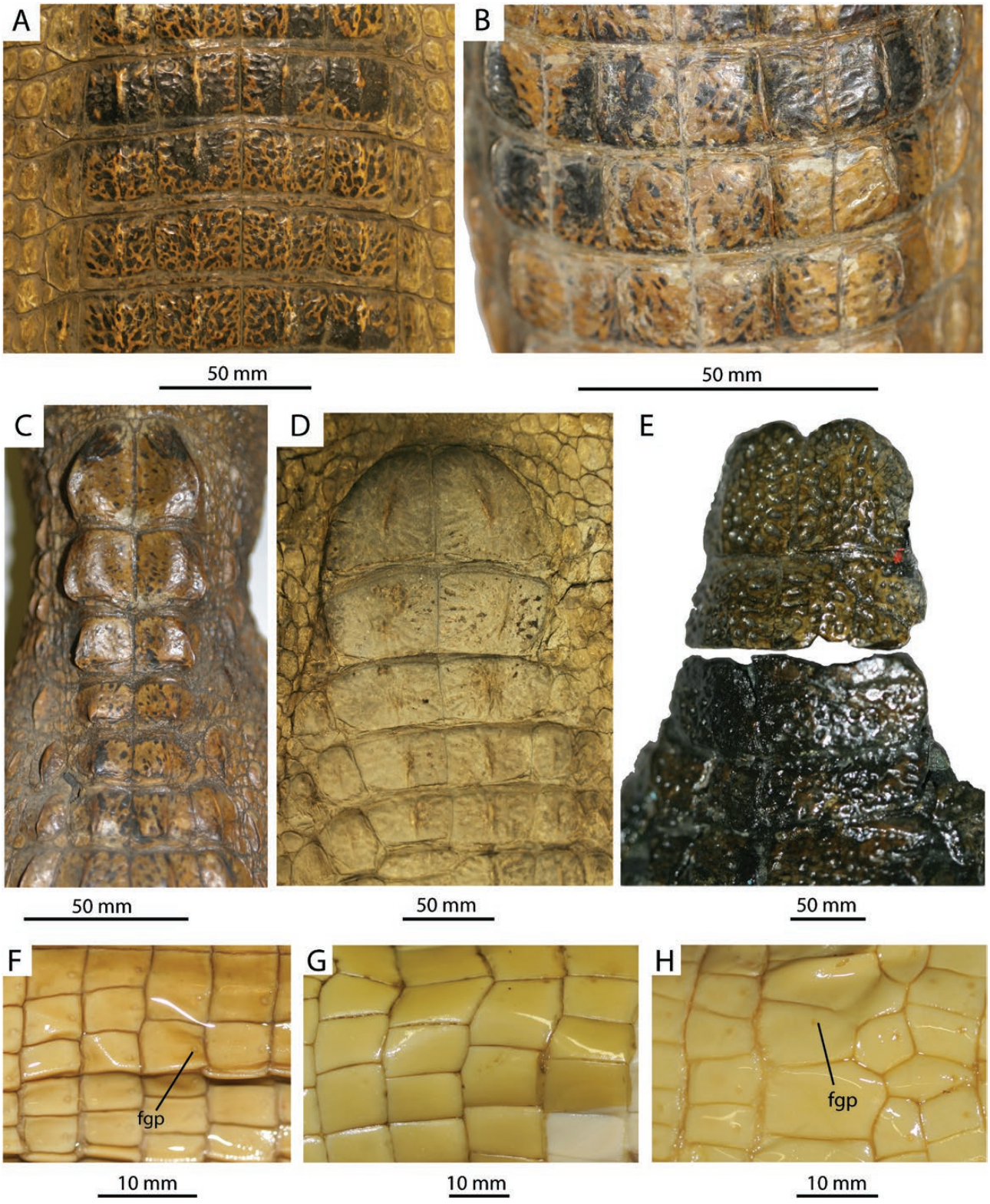


Figure 8. Illustration of osteoderm and scale characters that differ on observation from their original scoring for *Gavialis gangeticus* or *Tomistoma schlegelii*. Upper row: dorsal osteoderms of (A) *Gavialis gangeticus* ZMB uncatalogued 2018 specimen (dorsal view); and (B) *Tomistoma schlegelii* ZMB 36684 (dorsal and slightly anterior view). Both taxa show four

metriorhynchid outgroup used. The use of this outgroup alone precipitated a more basal placement of *Gavialis*. However, on reassessment of *Bernissartia fagesi*, fewer features shared with *Gavialis* were found than in the original scoring, and reassessment and full documentation of characters uniting *Gavialis* and thoracosaurs – and indeed more broadly in fossil crocodilians – is warranted.

Snout/dental characters obviously associated with longirostry are relatively easy to identify, but even many of the other characters supposedly uniting *Gavialis* with 'mesosuchians' (Frey *et al.*, 1989; Tarsitano *et al.*, 1989) plausibly relate to overall skull shape and lifestyle, as indicated by work finding correlation of cranial character changes with shifts in ecomorphology (Sadleir & Makovicky, 2008). Tarsitano *et al.* (1989) made great play of the similarity in braincase morphology between *Gavialis,* 'mesosuchians' and posthatchlings of other extant crocodilians, considering that in all the braincase was not verticalized. Gold *et al.* (2014) confirmed the low verticalization of the *Gavialis* braincase, but demonstrated that there was no clear cut-off in overall braincase morphology, with *Tomistoma* and *Gavialis* overlapping. However, the Eustachian system of *Gavialis* was distinct, but was not compared to 'mesosuchians' or any other outgroup, and was not considered to be evidence against the DNA-based topology (Gold *et al.*, 2014). Leaving aside the potentially autapomorphic Eustachian morphology of *Gavialis*, overall braincase height and shape is intimately linked to overall skull shape. For example, in highly brevirostrine dwarf taxa – *Paleosuchus* and *Osteolaemus* (Fig. 11) – the braincase is strongly verticalized [as also confirmed by Tarsitano *et al.* (1989) and Gold *et al.* (2014)], with the parabasisphenoid extensively exposed below the basioccipital. As expected, given its less extreme longirostry in comparison to *Gavialis*, the braincase of *Tomistoma* is indeed more verticalized, but this verticalization does not represent a markedly different organization but rather a point on a continuum and is not indisputably more similar to, for example, *Crocodylus* than *Gavialis* [Fig. 11; in fig. 4 of Tarsitano *et al.* (1989) the skull of *Tomistoma* is photographed in slightly ventral view, exaggerating the height of the braincase]. *Mecistops* equally shows a less verticalized braincase floor than *Crocodylus*, again in accordance with its longer rostrum (Fig. 11).

Given that the 'mesosuchians' (e.g. *Steneosaurus* Geoffroy, 1825: Tarsitano *et al.*, 1989: fig. 11), with which *Gavialis* was classically compared, were longirostrine (some more so than *Gavialis*), an overall similarity in braincase shape and arrangement is only to have been expected, based on the observations above. It should also be noted that ontogenetic trajectories have been used to indicate *Gavialis* to be further from *Tomistoma* than the latter is from other taxa (Piras *et al.*, 2010; *Crocodylus* and *Mecistops*). However, not only does overall similarity (as with skull shape) not provide sufficient grounds for inferring relationships between taxa, but the phylogenetic utility of such trajectories has subsequently been shown to be minimal (Watanabe & Slice, 2014). Alongside cranial characters, musculature has been used extensively to support placement of *Gavialis* outside of other extant crocodilians (Frey *et al.*, 1989). Soft tissues were not investigated in the present study and none of the skeletal characters analysed referred to muscular correlates relevant in this debate, so no conclusions in this regard can be drawn. However, if musculature were to be investigated once again, its functional importance and association with overall morphology must be borne in mind, with *Gavialis* being arguably the most specialized piscivore (McCurry *et al.*, 2017) (and certainly the most longirostrine taxon) among extant crocodilians, with potentially convergent similarities to extinct, yet more specialized (Ballell *et al.*, 2019) piscivores to be expected.

Outgroup choice

The results presented demonstrate how outgroup choice and scoring can affect topologies, as already demonstrated by other studies (e.g. Wilberg, 2015), which has particular relevance again for the placement of *Gavialis*. Use of extant, rather than fossil, outgroups resulted in resolution of Longirostres, even using all characters as originally scored. Given that DNA-based analyses must use extant outgroups, this could potentially explain the discrepancies seen in both this and other cases. Assuming that DNA-based phylogenies are indeed more accurate, the relative difficulty in scoring characters in fossils due to preservation and

main rows of large osteoderms with a much smaller row on each side lateral to these, contra the scoring of Narváez *et al.* (2015). Middle row: nuchal osteoderms in dorsal view of (C) *Tomistoma schlegelii* ZMB 36684; (D) *Gavialis gangeticus* ZMB 36736; (E) *Bernissartia fagesi* RBINS R 46. The nuchal shield of *Tomistoma* is not clearly separated from the dorsal shield, *contra* the scoring of Narváez *et al.* (2015). Lower row: ventral scales in ventral view showing presence or absence of follicle gland pores for (F) *Gavialis gangeticus* ZMB 36552; (G) *Alligator mississippiensis* ZMB 5383; (H) *Crocodylus niloticus* ZMB 36552. *Gavialis* shows follicle gland pores, as do *Tomistoma* and Crocodylinae, contra the apparent scoring in Narváez *et al.* (2015). Anterior is orientated upwards in A–E and left in F–H. Scale bars top and middle rows = 50 mm, lower row = 10 mm.

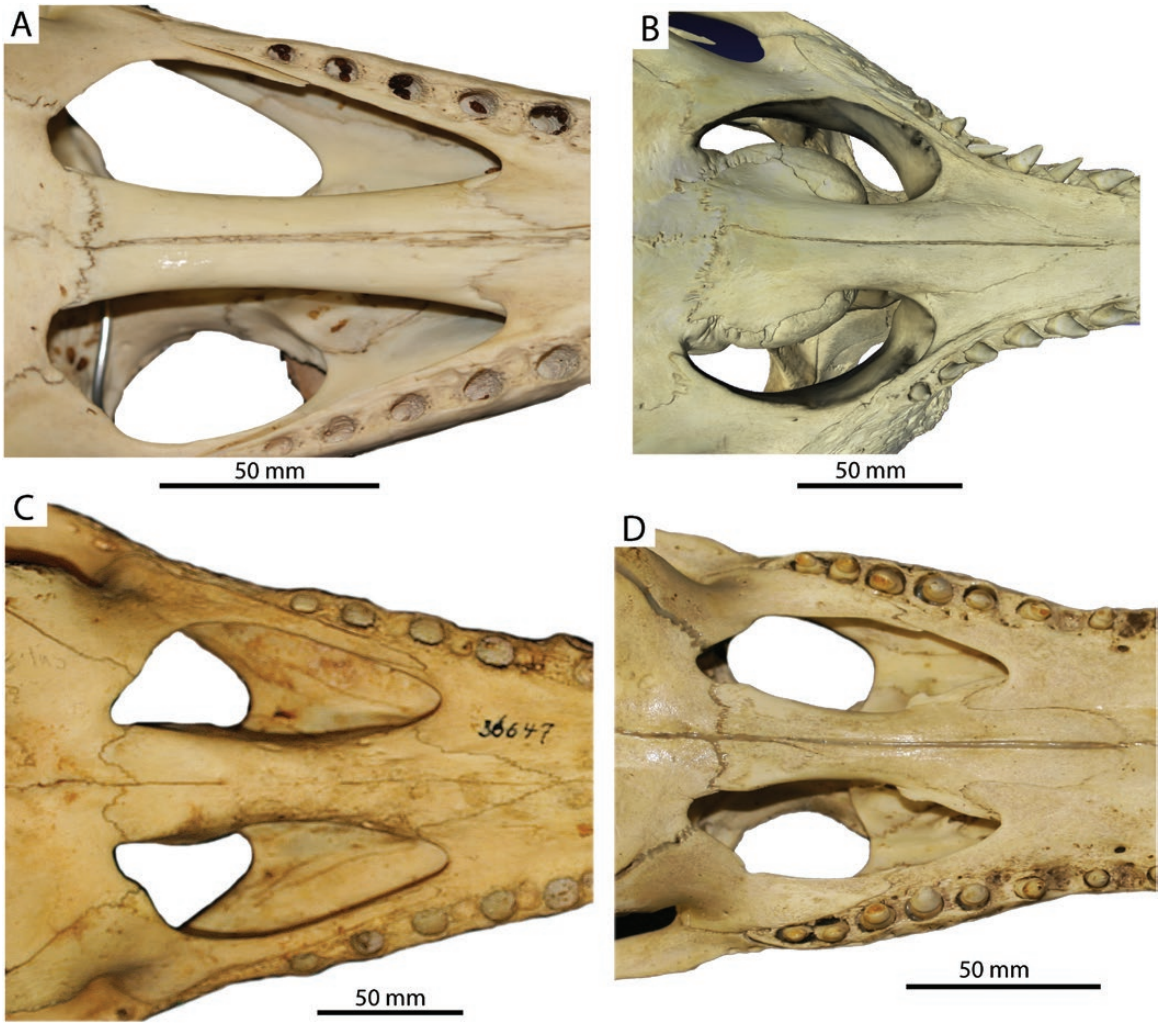


Figure 9. Palate in ventral view showing anterior palatine processes of: A, *Tomistoma schlegelii* ZMB 37221; B, *Gavialis gangeticus* ZMB 36678 (light surface scan rendering); C, *Mecistops cataphractus* ZMB 33647; D, *Crocodylus niloticus* MUT SZ 9486. The palatine processes of *Tomistoma* taper anteriorly as much, if not more, than those of *Gavialis*, contra the scoring of Narváez *et al.* (2015). Scale bars = 10 mm.

incompleteness may explain why inclusion of fossils could precipitate incorrect topologies. The use of near fossil (or indeed extant, though not applicable here) outgroups is also potentially problematic because a priori assignment of a taxon as outside the ingroup may be incorrect. However, these results should be treated with caution given that closer (although definitively non-ingroup) outgroups have been shown to yield potentially more accurate topologies (Wilberg, 2015), and the result may also have been precipitated due to the skull shape of the extant outgroup used, with *Varanus* approaching brevirostrine forms in overall skull shape and being concordantly scored in the matrix. This is supported by the fact that when the fossil outgroup *Bernissartia fagesi* was used alone, following first-hand reassessment of the taxon [the taxon was also included in the original matrix of Narváez *et al.* (2015), and when used, as scored by these authors, as a sole outgroup resulted in the same topology as presented by those authors], Longirostres was recovered, albeit with *Gavialis* sister to other taxa. Conversely, using the metriorhynchid outgroup alone, *Gavialis* was placed as the sister to all other extant taxa and *Tomistoma* as sister to Crocodylinae+Alligatoridae. This was due to the convergent features shared by the outgroup, *Gavialis*, and, in fewer cases, *Tomistoma*, relating to piscivory/longirostry (e.g. homodont dentition; lack of contact between nasals and premaxillae; using NJ the three taxa form a group with *Gavialis* sister to the metriorhynchid – see, for example, analysis 76 in supplement). This indicates that including more than one taxon in the outgroup, ideally with differing

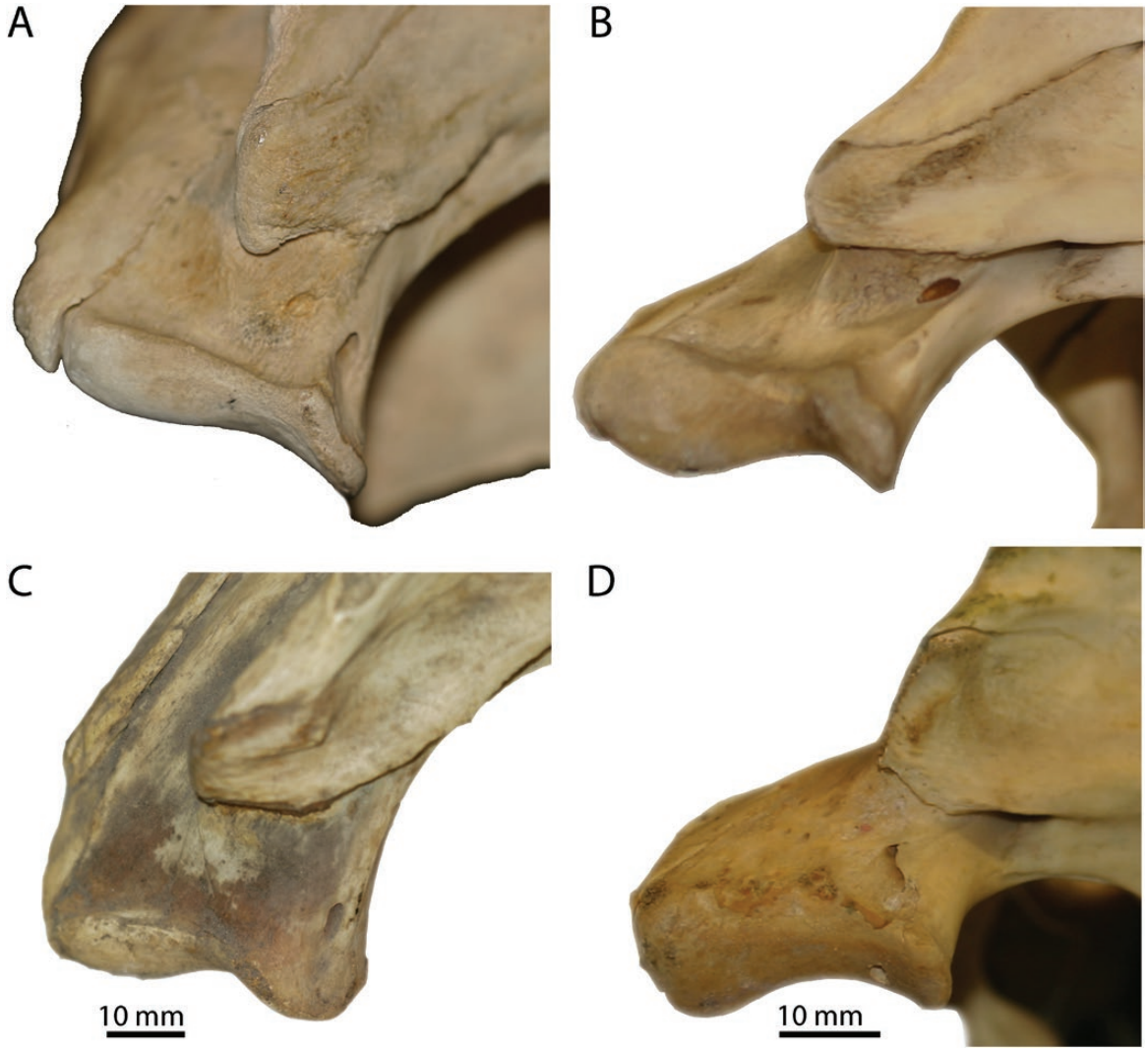


Figure 10. Distal quadrate showing medial (and lateral) hemicondyle of: A, *Gavialis gangeticus* ZMB 36678; B, *Tomistoma schlegelii* ZMB 37221; C, *Crocodylus niloticus* ZMB 36658; D, *Alligator mississippiensis* ZMB 36623. The morphology of *Gavialis* and *Tomistoma* is broadly similar, and distinct from other extant taxa, *contra* the scoring of Narváez *et al.* (2015). Scales bars = 10 mm, that below C applies to A–C, that below D applies to D. A and D show left quadrates and B and C right quadrates inverted for comparison. All are in broadly posterior view, but B is also slightly medial, and C slightly dorsal.

ecomorphologies, may be advisable. Supporting this influence of skull shape, the collapse of Alligatoridae when *Varanus* alone is used as the outgroup, with robust-only and reassessed datasets, appears to be due to characteristics of *Alligator* that are also probably plesiomorphic for amniotes (e.g. dentary tooth occluding laterally to maxillary teeth, fully bisected naris) shared with *Alligator*; inclusion of *Struthio* in the outgroup rectifies this 'problem', but due to its highly derived morphology minimizing symplesiomorphies. Indeed, using NJ with the reassessed matrix and uncertain outgroup scoring, *Alligator* and *Varanus* form a group – see analyses 30 and 95 in supplement – although this is not the case in any other analyses.

It must also be considered that there remains the possibility that a lack of close enough outgroups may lead to polarization mistakes in DNA-based phylogenetic inference. For example, the placement of *Gavialis* with crocodylids using DNA data could be due to both the gharial and Crocodylinae sharing genetic innovations that were common to stem crocodilians, but lost or even reversed in alligatorids. The potential to use a fossil outgroup may, thus, potentially make the morphology-based topology *more* accurate. Similarly, ability to include fossils potentially on the stem of *Tomistoma* and *Gavialis* (see i.a. Brochu, 1997; Lee & Yates, 2018) may also potentially mean morphology-based inferences are more accurate, because these

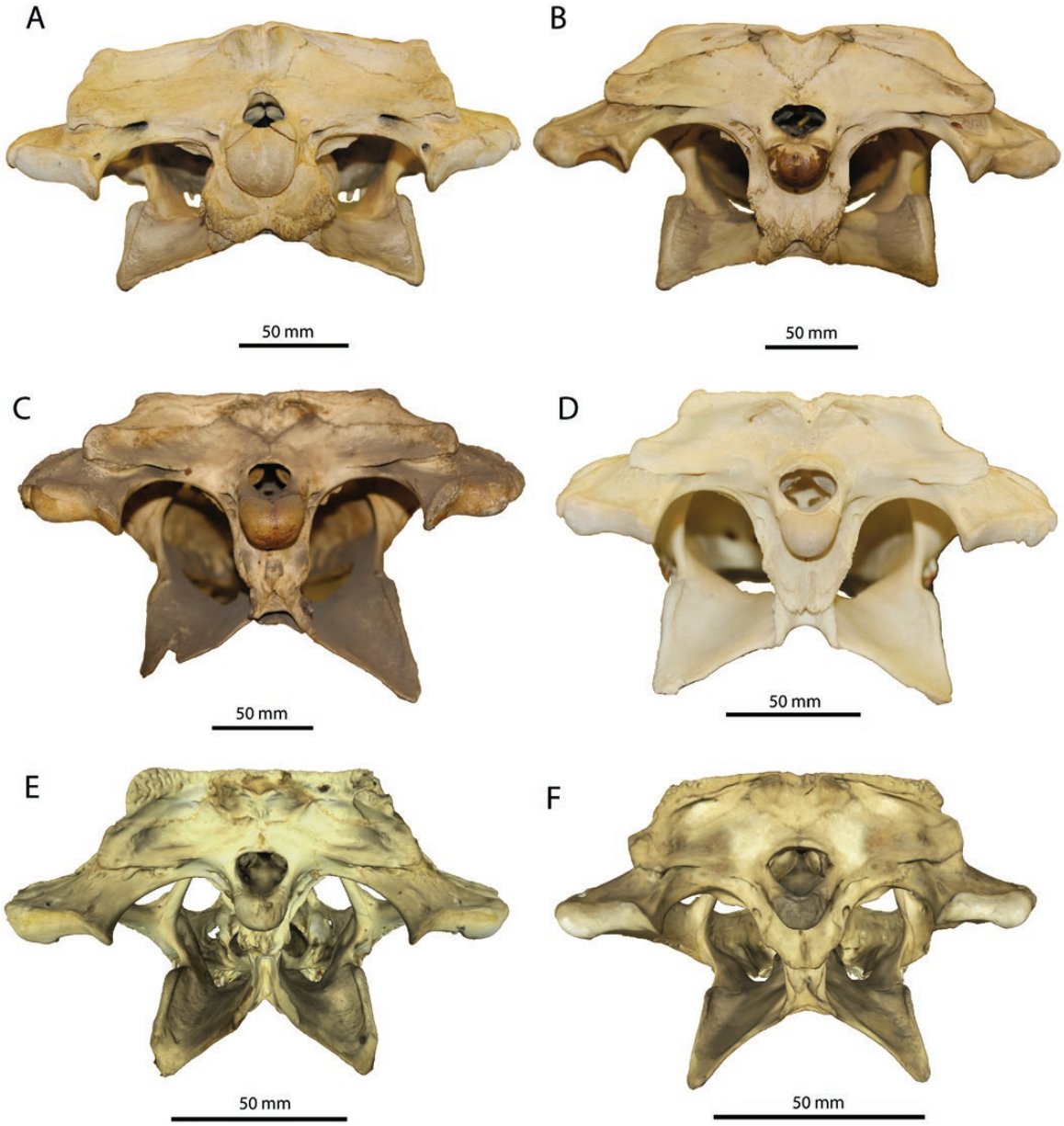


Figure 11. Posterior view of the skull of extant crocodilian taxa including *Gavialis* and *Tomistoma*, showing the height and form of the skull and braincase. Braincase height is intimately linked to overall cranial shape, with the lower braincase of *Gavialis* which has been used as evidence for its placement as sister taxon to other extant crocodilians plausibly due to its extreme longirostry. A, *Gavialis gangeticus* ZMB 36678; B, *Tomistoma schlegelii* ZMB 37221; C, *Mecistops cataphractus* ZMB 33647; D, *Crocodylus niloticus* ZMB 80690; E, *Paleosuchus trigonatus* ZMB 36636, light surface scan rendering; F, *Osteolaemus tetraspis* ZMB 25495, light surface scan rendering. Scale bars = 50 mm.

forms may show fewer derived features that mask their relationships to other clades. Although the percentage of missing data was not higher in the extant outgroups than fossil outgroups (see Methodology), lack of key information in extant outgroups may lead to a spurious grouping, with important polarizations lacking. However, the fact that a more DNA-concordant result was also obtained using *Bernissartia* points against this to some extent and rather towards the importance of skull shape.

However, given the size and quality of molecular datasets now available (Green *et al.*, 2014), the concordance between (Erickson *et al.*, 2012; Green *et al.*, 2014) and within (Harshman *et al.*, 2003) datasets and the relatively good understanding we have of how to model molecular evolution (in marked contrast

to morphology – Goloboff *et al.*, 2018b), the greater objectivity in homologization and scoring of molecular data (Scotland *et al.*, 2003), and the fact that the more robustly constructed morphological characters are more concordant with the DNA-based topology, such incorrect inferences seem increasingly unlikely. More plausible – as is generally accepted (Brochu, 1997; Harshman *et al.*, 2003; Gold *et al.*, 2014; Lee & Yates, 2018) – is that the use of fossil outgroups either convergent on (Sadleir & Makovicky, 2008) or showing plesiomorphies retained (or 'regained' as atavisms – Gatesy *et al.*, 2003; although this can arguably be subsumed in convergence) by *Gavialis* has resulted in the topologies hitherto inferred. However, it should also be noted that hybridization/gene flow between species may have affected molecular (and potentially morphological) phylogenies, with hybridization at least within genera known to occur (Fitzsimmons *et al.*, 2002; Cedeño-Vázquez *et al.*, 2008; Rodríguez *et al.*, 2008; Weaver *et al.*, 2008; Milián-García *et al.*, 2015). Although there is no evidence for this currently, the geographic proximity of, for example, *Gavialis* and *Tomistoma,* and *Mecistops*, *Osteolaemus* and the African *Crocodylus* species, does not rule out hybridization or introgression. Occurrence of incomplete lineage sorting is also plausible, having been found extensively in birds (Suh *et al.*, 2015). Further analyses, both molecular and morphological, with attempts to incorporate hybridization (Reich *et al.*, 2009; Patterson *et al.*, 2012; Bromham, 2017) or incomplete lineage sorting (Wang *et al.*, 2018) may be of interest.

Anatomical region, number of states and meristic characters

Only limited inferences regarding the usefulness of different anatomical regions in phylogenetic inference can be drawn from this dataset. The only significant comparison in homoplasy of characters from different anatomical regions against the DNA-based phylogeny was that cranial characters were, on average, less homoplastic. This may be because of greater variability in the cranium than the postcranium of extant crocodilians, which are a relatively ecomorphologically limited group (i.e. all semi-aquatic, generally freshwater: Ross & Garnett, 1989; Benton, 2014; with this limitation clear when compared to Crocodyliformes more widely: Godoy *et al.*, 2016), allowing phylogenetic information to be more easily extracted. It plausibly may also reflect the greater number of structures, including ossified elements, nerve and blood vessel passages (i.e. arguably greater overall complexity), of the cranium in comparison to the postcranium, meaning greater information content. Although the number of ossified elements per se of the postcranium is larger, many of these are clear serial homologues (vertebrae, ribs) and, thus, unlikely to yield additional phylogenetic information. However, it must be noted that different patterns have been found for other groups, including the arguably more ecomorphologically diverse squamates, birds and mammals (Callender-Crowe, pers. comm.; Sansom *et al.*, 2017), and including the postcranium has been found to increase phylogenetic accuracy in phylogenetic morphometrics (Catalano *et al.*, 2014). There is also the possibility that the difference seen is because the postcranium is relatively understudied (Mounce *et al.*, 2016), with many museum collections being relatively depauperate in postcrania. Furthermore, Godoy *et al.* (2016) found that the cranium contributed more to the phylogenetic signal due to lower amounts of missing data in the cranial partition. However, in the dataset at hand, the percentage of missing data was very low overall (1.38%) and did not differ greatly between the cranial (1.18%) and postcranial (1.95%) partition and, thus, this is unlikely to be a factor. Rather, lack of study may have precipitated lack of appropriate and well-delimited postcranial characters. Given the limited sample, and heavy bias of the dataset towards the cranium, it is inadvisable to draw any firm conclusions about use of cranial and postcranial characters.

Also worthy of note is the apparent robustness of characters associated with interlocking structures. For example, the shape and placement of the foramen aerum on the quadrate and articular were robust characters grouping alligatorids in contrast to Longirostres, as was the articulation of the fourth dentary tooth (whether in a pit or a lateral groove). Whether this pattern is widely the case would need to be investigated, but such features could be envisaged to be more conserved because simultaneous change of both structures would be required if function were to be maintained. This draws a potential parallel to the extensive discussion and use of a similar interlocking structure – the ankle – in wider archosaur and diapsid phylogeny (Parrish, 1987; Sereno & Arcucci, 1990).

As expected, based on theoretical (e.g. Hoyal Cuthill *et al.*, 2010; Hoyal Cuthill, 2015) and empirical (Zou & Zhang, 2016) work, the consistency index (CI) was significantly correlated with the number of states of a character. However, this is not significant evidence for the claim made by Zou & Zhang (2016) that multistate morphological characters should be preferred. Simulations have shown characters with larger real state spaces to be more useful in phylogenetic inference (Simmons *et al.*, 2004), but real morphological characters differ fundamentally in that they can usually be broken down into more characters with fewer states per character, or amalgamated into multistate characters, and the state space is thus artificial. Rather, the correlation found here is probably indicative of the previously elucidated biased

nature of homoplasy indices, and the CI in particular (Hoyal Cuthill *et al.*, 2010; Hoyal Cuthill, 2015), when comparing characters with different numbers of states. Thus, although the only homoplasy index that was significantly greater in broadly meristic characters, i.e. characters either 'classically' meristic (e.g. number of premaxillary teeth) or those based on a meristic feature (i.e. element X extends beyond tooth *N*), was number of extra steps (H), this result is still of note because H is the index that was least correlated to the number of states of the characters, and meristic characters had significantly more states than other characters. Meristic characters can appear attractive because of their clear, natural delimitation. However, most such characters pertain to serial homologues (e.g. teeth, vertebrae), which tend to increase in number if the structure in which they are embedded or which they form increases in length or size (Lindsey, 1975).

In the current example, the dwarf caiman *Paleosuchus* and dwarf crocodile *Osteolaemus* both have four premaxillary teeth, as opposed to five in other taxa. The number of vertebrae with postatlantal hypapophyseal keels also seems to be, to some extent, connected to dwarfing, with the small-bodied Chinese alligator *Alligator sinensis* Fauvel, 1879 along with *Paleosuchus* and *Osteolaemus* both showing 12 as opposed to 11 such vertebrae, perhaps due to faster somitgenesis creating more, smaller vertebrae (Müller *et al.*, 2010), although the exception is *Crocodylus porosus* that also has 12 keeled vertebrae and is the largest extant reptile (Britton *et al.*, 2012). In any case, such characters do seem to be exceptionally labile, probably because of the ease of shifting the mechanisms controlling these characters, e.g. the speed of the 'segmentation clock' regulating somitogenesis and shifting *Hox* gene expression regions controlling regionalization (Müller *et al.*, 2010), without greatly impeding function or interfering with development of other structures. The highly conserved (especially cervical) vertebral count of mammals (Asher *et al.*, 2011) is a major exception to this pattern of lability, but is not replicated in other amniote (Müller *et al.*, 2010) or tetrapod (Inger, 1967) groups, or actinopterygian fish (Lindsey, 1975). For species-/genus-level identification, such characters may be more useful and no variation was observed for meristic characters in the genera observed. However, caution is warranted because even within-species variation has been observed in lissamphibian taxa (Inger, 1967). Overall, given the lability of meristic characters, at least outside Mammalia, their exclusion from phylogenetic analyses could be warranted.

CONCLUSIONS

How better to infer phylogenies using morphological data remains a question as yet to be fully answered. However, the results presented here provide some indications as to the ways in which morphology may be able to be used more effectively, or at least in a way that is more concordant with the results of DNA-based analyses, both for crocodiles and their relatives, and more broadly. Improving delimitation of characters to make certain that state delimitation is firmly based on plausible biological homologies may at least yield characters that are less at odds with molecular data. However, these may still not yield a more accurate topology, because information may not be sufficient to resolve clades and convergence (especially due to overall shape) can still occur. Analytical method appears to have relatively little effect on the topology yielded, but the results of a Bayesian analysis of a reassessed dataset are notable in being the first to place both *Gavialis* and *Mecistops* in a position concordant with multigene DNA analyses. Outgroup choice had a noticeable effect on the outcomes of phylogenetic analyses. It should be aimed to include as broad and representative an outgroup as possible, and the effects of using extant outgroups, rather than fossil outgroups, in DNA-based analyses should be borne in mind. Exclusion of meristic and other highly labile characters, and attempts to explicitly take convergence into account, may bear fruit. More general concerns aside, following reassessment of the morphological case supporting placement of *Gavialis* outside of the rest of the crown, it appears at least possible to conclude that it has been overstated, with many characters linked to morphological adaptations to a similar lifestyle in *Gavialis* and fossil taxa.

ACKNOWLEDGEMENTS

The author would like to thank the following people: F. Montefeltro (Universidade Estadual Paulista, Ilha Solteira), C. Brochu (University of Iowa) and P. Godoy (University of Birmingham) for generous help in understanding crocodilian characters, J. Hinz (Eberhard Karls University of Tübingen, EKUT) for sharing and providing information regarding 3D scan data, E. Ascarrunz (Université de Fribourg) and M. Ezcurra (Museo Argentino de Ciencias Naturales) for technical assistance and discussions, J. Anquetin (JURASSICA Museum) for sharing scan data, assisting in scanning further specimens and discussions, R. Felice (University College London) for sharing data, E. Gold (Suffolk University) and E. Wilberg (Stony Brook University) for assistance locating specimens, A. Resetar (FMNH) and I. Werneburg (EKUT) for photographing specimens, K. Mahlow and M. Kirchner (MNB) for assistance with CT data and discussions, J. Müller (MNB) for hosting and assisting with the project and discussions and last, but certainly not least, A. Folie, T. Walschaerts, O. Pauwels (all RBINS), F. Glaw and M. Franzen (ZSM),

J. Hallermann (ZMH), J. Moravec (NMP), E. Westwig (OUM), E. Weber and J. Rösinger (MUT), P. Campbell (NHMUK), D. Schwarz, F. Tillack, S. Frahnert and P. Eckhoff (all MNB) for access to specimens in their care. The author thanks two reviewers for their comments, which greatly improved the manuscript.

FUNDING

This work was funded by a research fellowship for postdoctoral researchers from the Alexander von Humboldt Foundation.

REFERENCES

Almeida M, **Bisby F. 1984.** A simple method for establishing taxonomic characters from measurement data. *Taxon* **33:** 405–409.

Andres B, **Myers TS. 2012.** Lone star pterosaurs. *Earth and Environmental Science Transactions of the Royal Society of Edinburgh* **103:** 383–398.

Asher RJ, **Lehmann T. 2008.** Dental eruption in afrotherian mammals. *BMC Biology* **6:** 14.

Asher R, **Lin K**, **Kardjilov N**, **Hautier L. 2011.** Variability and constraint in the mammalian vertebral column. *Journal of Evolutionary Biology* **24:** 1080–1090.

Atchley WR, **Fitch WM. 1991.** Gene trees and the origins of inbred strains of mice. *Science* **254:** 554.

Ballell A, **Moon BC**, **Porro LB**, **Benton MJ**, **Rayfield EJ**. **2019.** Convergence and functional evolution of longirostry in crocodylomorphs. *Palaeontology*.

Baron MG, **Norman DB**, **Barrett PM. 2017.** A new hypothesis of dinosaur relationships and early dinosaur evolution. *Nature* **543:** 501–506.

Benson RBJ, **Frigot RA**, **Goswami A**, **Andres B**, **Butler RJ. 2014.** Competition and constraint drove Cope's rule in the evolution of giant flying reptiles. *Nature Communications* **5:** 3567.

Benson RB, **Hunt G**, **Carrano MT**, **Campione N. 2018.** Cope's rule and the adaptive landscape of dinosaur body size evolution. *Palaeontology* **61:** 13–48.

Benton M. 2007. Biodiversity through time. In: Briggs D, Crowther PR, ed. *Palaeobiology II*. Malden (MA): Blackwell Science Ltd, 600.

Benton M. 2014. *Vertebrate palaeontology*. Hoboken (NJ): Wiley-Blackwell.

Britton A, **Whitaker R**, **Whitaker N. 2012.** Here be a dragon: exceptional size in a saltwater crocodile (*Crocodylus porosus*) from the Philippines. *Herpetological Review* **43:** 541–546.

Brochu CA. 1997. Morphology, fossils, divergence timing, and the phylogenetic relationships of *Gavialis*. *Systematic Biology* **46:** 479–522.

Brochu CA. 1999a. Phylogenetics, taxonomy, and historical biogeography of Alligatoroidea. *Journal of Vertebrate Paleontology* **19:** 9–100.

Brochu CA. 1999b. Taxon sampling and reverse successive weighting. *Systematic Biology* **48:** 808–813.

Brochu CA. 2003. Phylogenetic approaches toward crocodylian history. *Annual Review of Earth and Planetary Sciences* **31:** 357–397.

Brochu CA, **Storrs GW. 2012.** A giant crocodile from the Plio-Pleistocene of Kenya, the phylogenetic relationships of Neogene African crocodylines, and the antiquity of *Crocodylus* in Africa. *Journal of Vertebrate Paleontology* **32:** 587–602.

Bromham L. 2017. Curiously the same: swapping tools between linguistics and evolutionary biology. *Biology & Philosophy* **32:** 855–886.

Brown TA. 2002. *Genomes, 2nd edn*. Wiley-Liss.

Buffetaut E. 1985. The place of *Gavialis* and *Tomistoma* in Eusuchian evolution: a reconciliation of palaeontological and biochemical data. *Neues Jahrbuch für Geologie und Paläontologie, Monatshefte* **12:** 707–716.

Catalano SA, **Torres A. 2017.** Phylogenetic inference based on landmark data in 41 empirical data sets. *Zoologica Scripta* **46:** 1–11.

Catalano SA, **Ercoli MD**, **Prevosti FJ. 2014.** The more, the better: the use of multiple landmark configurations to solve the phylogenetic relationships in musteloids. *Systematic Biology* **64:** 294–306.

Cedeño-Vázquez JR, **Rodriguez D**, **Calmé S**, **Ross JP**, **Densmore LD**, **Thorbjarnarson JB. 2008.** Hybridization between *Crocodylus acutus* and *Crocodylus moreletii* in the Yucatan Peninsula: I. Evidence from mitochondrial DNA and morphology. *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology* **309:** 661–673.

Collen B, **Turvey ST**, **Waterman C**, **Meredith HMR**, **Kuhn TS**, **Baillie JEM**, **Isaac NJB. 2011.** Investing in evolutionary history: implementing a phylogenetic approach for mammal conservation. *Philosophical Transactions of the Royal Society B: Biological Sciences* **366:** 2611–2622.

Crawford NG, **Parham JF**, **Sellas AB**, **Faircloth BC**, **Glenn TC**, **Papenfuss TJ**, **Henderson JB**, **Hansen MH**, **Simison WB. 2015.** A phylogenomic analysis of turtles. *Molecular Phylogenetics and Evolution* **83:** 250–257.

Crisp MD, **Cook LG. 2009.** Explosive radiation or cryptic mass extinction? Interpreting signatures in molecular phylogenies. *Evolution* **63:** 2257–2265.

Davis CC, **Willis CG**, **Primack RB**, **Miller-Rushing AJ. 2010.** The importance of phylogeny to the study of phenological response to global climate change. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365:** 3201–3213.

Delsuc F, **Brinkmann H**, **Philippe H. 2005.** Phylogenomics and the reconstruction of the tree of life. *Nature Reviews Genetics* **6:** 361–375.

Edwards SV. 2009. Is a new and general theory of molecular systematics emerging? *Evolution* **63:** 1–19.

Erickson GM, **Gignac PM**, **Steppan SJ**, **Lappin AK**, **Vliet KA**, **Brueggen JD**, **Inouye BD**, **Kledzik D**, **Webb GJ. 2012.** Insights into the ecology and evolutionary success of crocodilians revealed through bite-force and tooth-pressure experimentation. *PLoS ONE* **7:** e31781.

Farris JS. 1989. The retention index and the rescaled consistency index. *Cladistics* **5:** 417–419.

Fitzsimmons NN, **Buchan JC**, **Lam PV**, **Polet G**, **Hung TT**, **Thang NQ**, **Gratten J. 2002.** Identification of purebred *Crocodylus siamensis* for reintroduction in Vietnam. *Journal of Experimental Zoology* **294:** 373–381.

Forest F, **Crandall KA**, **Chase MW**, **Faith DP. 2015.** Phylogeny, extinction and conservation: embracing uncertainties in a time of urgency. *Philosophical Transactions of the Royal Society B: Biological Sciences* **370:** 20140002.

Frey E, **Riess J**, **Tarsitano SF. 1989.** The axial tail musculature of recent crocodiles and its phyletic implications. *American Zoologist* **29:** 857–862.

Gatesy J, **Amato G**, **Norell M**, **DeSalle R**, **Hayashi C. 2003.** Combined support for wholesale taxic atavism in gavialine crocodylians. *Systematic Biology* **52:** 403–422.

Gatesy J, **Baker RH**, **Hayashi C. 2004.** Inconsistencies in arguments for the supertree approach: supermatrices versus supertrees of Crocodylia. *Systematic Biology* **53:** 342–355.

Garamszegi LZ. 2014. *Modern phylogenetic comparative methods and their application in evolutionary biology: concepts and practice*. Berlin/Heidelberg: Springer.

Gauthier J, **Kluge AG**, **Rowe T. 1988.** Amniote phylogeny and the importance of fossils. *Cladistics* **4:** 105–209.

Gavryushkina A, **Heath TA**, **Ksepka DT**, **Stadler T**, **Welch D**, **Drummond AJ. 2017.** Bayesian total-evidence dating reveals the recent crown radiation of penguins. *Systematic Biology* **66:** 57–73.

Godoy PL, **Bronzati M**, **Eltink E**, **Júlio CdA**, **Cidade GM**, **Langer MC**, **Montefeltro FC. 2016.** Postcranial anatomy of *Pissarrachampsa sera* (Crocodyliformes, Baurusuchidae) from the Late Cretaceous of Brazil: insights on lifestyle and phylogenetic significance. *PeerJ* **4:** e2075.

Gold MEL, **Brochu CA**, **Norell MA. 2014.** An expanded combined evidence approach to the *Gavialis* problem using geometric morphometric data from crocodylian braincases and Eustachian systems. *PLoS ONE* **9:** e105793.

Goloboff PA, **Catalano SA. 2016.** TNT version 1.5, including a full implementation of phylogenetic morphometrics. *Cladistics* **32:** 221–238.

Goloboff PA, **Torres A**, **Arias JS. 2018a.** Weighted parsimony outperforms other methods of phylogenetic inference under models appropriate for morphology. *Cladistics* **34:** 407–437.

Goloboff PA, **Pittman M**, **Pol D**, **Xu X. 2018b.** Morphological datasets fit a common mechanism much more poorly than DNA sequences and call into question the Mkv model. *Systematic Biology*. doi:10.1093/sysbio/syy077. Available at: https://academic.oup.com/sysbio/advance-article-abstract/doi/10.1093/sysbio/syy077/5184275

Green RE, **Braun EL**, **Armstrong J**, **Earl D**, **Nguyen N**, **Hickey G**, **Vandewege MW**, **John JAS**, **Capella-Gutiérrez S**, **Castoe TA. 2014.** Three crocodilian genomes reveal ancestral patterns of evolution among archosaurs. *Science* **346:** 1254449.

Hammer Ø, **Harper D**, **Ryan P. 2001.** PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica* **4:** 9.

Harshman J, **Huddleston CJ**, **Bollback JP**, **Parsons TJ**, **Braun MJ. 2003.** True and false gharials: a nuclear gene phylogeny of Crocodylia. *Systematic Biology* **52:** 386–402.

Hass CA, **Hoffman MA**, **Densmore III LD**, **Maxson LR. 1992.** Crocodilian evolution: insights from immunological data. *Molecular Phylogenetics and Evolution* **1:** 193–201.

Hillis DM, **Bull JJ**, **White ME**, **Badgett MR**, **Molineux IJ. 1992.** Experimental phylogenetics: generation of a known phylogeny. *Science* **255:** 589–592.

Hoyal Cuthill JF. 2015. The size of the character state space affects the occurrence and detection of homoplasy: modelling the probability of incompatibility for unordered phylogenetic characters. *Journal of Theoretical Biology* **366:** 24–32.

Hoyal Cuthill JF, **Braddy SJ**, **Donoghue PCJ. 2010.** A formula for maximum possible steps in multistate characters: isolating matrix parameter effects on measures of evolutionary convergence. *Cladistics* **26:** 98–102.

Hrbek T, **Vasconcelos WR**, **Rebelo G**, **Farias IP. 2008.** Phylogenetic relationships of South American alligatorids and the caiman of Madeira River. *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology* **309:** 588–599.

Huelsenbeck JP, **Ronquist F. 2001.** MRBAYES: Bayesian inference of phylogenetic trees. *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.

Hull DL. 2010. *Science as a process: an evolutionary account of the social and conceptual development of science*. Chicago: University of Chicago Press.

Huson DH, **Bryant D. 2005.** Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* **23:** 254–267.

Iijima M, **Kobayashi Y. 2019.** Mosaic nature in the skeleton of East Asian crocodylians fills the morphological gap between 'Tomistominae' and Gavialinae. *Cladistics*. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1111/cla.12372

Inger RF. 1967. The development of a phylogeny of frogs. *Evolution* **21:** 369–384.

Jarvis ED, **Mirarab S**, **Aberer AJ**, **Li B**, **Houde P**, **Li C**, **Ho SY**, **Faircloth BC**, **Nabholz B**, **Howard JT. 2014.** Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science* **346:** 1320–1331.

Kluge AG, **Farris JS. 1969.** Quantitative phyletics and the evolution of anurans. *Systematic Biology* **18:** 1–32.

Langer MC, **Ezcurra MD**, **Rauhut OW**, **Benton MJ**, **Knoll F**, **McPhee BW**, **Novas FE**, **Pol D**, **Brusatte SL. 2017.** Untangling the dinosaur family tree. *Nature* **551:** E1.

Lee MS, **Yates AM. 2018.** Tip-dating and homoplasy: reconciling the shallow molecular divergences of modern gharials with their long fossil record. *Proceedings of the Royal Society of London B: Natural Sciences* **285:** 20181071.

Lindsey C. 1975. Pleomerism, the widespread tendency among related fish species for vertebral number to be correlated with maximum body length. *Journal of the Fisheries Board of Canada* **32:** 2453–2469.

Livezey BC, **Zusi RL. 2007.** Higher-order phylogeny of modern birds (Theropoda, Aves: Neornithes) based on comparative anatomy. II. Analysis and discussion. *Zoological Journal of the Linnean Society* **149:** 1–95.

Lyubetsky V, **Piel WH**, **Stadler PF. 2016.** Molecular phylogenetics 2016. *BioMed Research International* **2016:**

9029306. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5227124/

Maddison WP, **Maddison DR. 2017.** *Mesquite: a modular system for evolutionary analysis, v.3.2*. Available at: http://mesquiteproject.org. (accessed March 13, 2018).

Madsen O, **Scally M**, **Douady CJ**, **Kao DJ**, **DeBry RW**, **Adkins R**, **Amrine HM**, **Stanhope MJ**, **de Jong WW**, **Springer MS. 2001.** Parallel adaptive radiations in two major clades of placental mammals. *Nature* **409:** 610.

Manegold A. 2006. Two additional synapomorphies of grebes Podicipedidae and flamingos Phoenicopteridae. *Acta Ornithologica* **41:** 79–82.

Mannion PD, **Benson RB**, **Carrano MT**, **Tennant JP**, **Judd J**, **Butler RJ. 2015.** Climate constrains the evolutionary history and biodiversity of crocodylians. *Nature Communications* **6:** 8438.

Markwick PJ. 1998. Crocodilian diversity in space and time: the role of climate in paleoecology and its implication for understanding K/T extinctions. *Paleobiology* **24:** 470–497.

Mayr G. 2004. Morphological evidence for sister group relationship between flamingos (Aves: Phoenicopteridae) and grebes (Podicipedidae). *Zoological Journal of the Linnean Society* **140:** 157–169.

Mayr G, **Clarke J. 2003.** The deep divergences of neornithine birds: a phylogenetic analysis of morphological characters. *Cladistics* **19:** 527–553.

McAliley LR, **Willis RE**, **Ray DA**, **White PS**, **Brochu CA**, **Densmore III LD. 2006.** Are crocodiles really monophyletic?—Evidence for subdivisions from sequence and morphological data. *Molecular Phylogenetics and Evolution* **39:** 16–32.

McCurry MR, **Evans AR**, **Fitzgerald EM**, **Adams JW**, **Clausen PD**, **McHenry CR. 2017.** The remarkable convergence of skull shape in crocodilians and toothed whales. *Proceedings of the Royal Society B: Biological Sciences* **284:** 20162348.

Metzker ML. 2010. Sequencing technologies—the next generation. *Nature Reviews Genetics* **11:** 31.

Mickevich M, **Johnson MS. 1976.** Congruence between morphological and allozyme data in evolutionary inference and character evolution. *Systematic Zoology* **25:** 260–270.

Milián-García Y, **Ramos-Targarona R**, **Pérez-Fleitas E**, **Sosa-Rodríguez G**, **Guerra-Manchena L**, **Alonso-Tabet M**, **Espinosa-López G**, **Russello M. 2015.** Genetic evidence of hybridization between the critically endangered Cuban crocodile and the American crocodile: implications for population history and in situ/ex situ conservation. *Heredity* **114:** 272.

Morris ZS, **Vliet KA**, **Abzhanov A**, **Pierce SE. 2019.** Heterochronic shifts and conserved embryonic shape underlie crocodylian craniofacial disparity and convergence. *Proceedings of the Royal Society B* **286:** 20182389.

Mounce RC, **Sansom R**, **Wills MA. 2016.** Sampling diverse characters improves phylogenies: craniodental and postcranial characters of vertebrates often imply different trees. *Evolution* **70:** 666–686.

Müller J, **Scheyer TM**, **Head JJ**, **Barrett PM**, **Werneburg I**, **Ericson PG**, **Pol D**, **Sánchez-Villagra MR. 2010.** Homeotic effects, somitogenesis and the evolution of vertebral numbers in recent and fossil amniotes. *Proceedings of the National Academy of Sciences* **107:** 2118–2123.

Narváez I, **Brochu CA**, **Escaso F**, **Pérez-García A**, **Ortega F. 2015.** New crocodyliforms from southwestern Europe and definition of a diverse clade of European Late Cretaceous basal eusuchians. *PLoS ONE* **10:** e0140679.

Nee S, **Mooers AO**, **Harvey PH. 1992.** Tempo and mode of evolution revealed from molecular phylogenies. *Proceedings of the National Academy of Sciences* **89:** 8322–8326.

O'leary MA, **Bloch JI**, **Flynn JJ**, **Gaudin TJ**, **Giallombardo A**, **Giannini NP**, **Goldberg SL**, **Kraatz BP**, **Luo Z-X**, **Meng J. 2013.** The placental mammal ancestor and the post–K-Pg radiation of placentals. *Science* **339:** 662–667.

O'Malley MA, **Koonin EV. 2011.** How stands the Tree of Life a century and a half after The Origin? *Biology Direct* **6:** 32.

O'Reilly JE, **Puttick MN**, **Parry L**, **Tanner AR**, **Tarver JE**, **Fleming J**, **Pisani D**, **Donoghue PC. 2016.** Bayesian methods outperform parsimony but at the expense of precision in the estimation of phylogeny from discrete morphological data. *Biology Letters* **12:** 20160081.

O'Reilly JE, **Puttick MN**, **Pisani D**, **Donoghue PC. 2018.** Probabilistic methods surpass parsimony when assessing clade support in phylogenetic analyses of discrete morphological data. *Palaeontology* **61:** 105–118.

Oaks JR. 2011. A time-calibrated species tree of Crocodylia reveals a recent radiation of the true crocodiles. *Evolution* **65:** 3285–3297.

Paradis E, **Claude J**, **Strimmer K. 2004.** APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20:** 289–290.

Parrish JM. 1987. The origin of crocodilian locomotion. *Paleobiology* **13:** 396–414.

Patterson NJ, **Moorjani P**, **Luo Y**, **Mallick S**, **Rohland N**, **Zhan Y**, **Genschoreck T**, **Webster T**, **Reich D. 2012.** Ancient admixture in human history. *Genetics* **192:** 1065–1093.

Piras P, **Colangelo P**, **Adams DC**, **Buscalioni A**, **Cubo J**, **Kotsakis T**, **Meloro C**, **Raia P. 2010.** The *Gavialis–Tomistoma* debate: the contribution of skull ontogenetic allometry and growth trajectories to the study of crocodylian relationships. *Evolution & development* **12:** 568–579.

Pomidor BJ, **Makedonska J**, **Slice DE. 2016.** A landmark-free method for three-dimensional shape analysis. *PLoS ONE* **11:** e0150368.

Prum RO, **Berv JS**, **Dornburg A**, **Field DJ**, **Townsend JP**, **Lemmon EM**, **Lemmon AR. 2015.** A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature* **526:** 569.

Puttick MN, **O'Reilly JE**, **Pisani D**, **Donoghue PC. 2018.** Probabilistic methods outperform parsimony in the phylogenetic analysis of data simulated without a probabilistic model. *Palaeontology* **62:** 1–17.

Pyron RA, **Burbrink FT**, **Wiens JJ. 2013.** A phylogeny and revised classification of Squamata, including 4161 species of lizards and snakes. *BMC Evolutionary Biology* **13:** 93.

Pyron RA. 2015. Post-molecular systematics and the future of phylogenetics. *Trends in Ecology & Evolution* **30:** 384–389.

R Core Team. 2014. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Rasnitsyn A. 2006. Ontology of evolution and methodology of taxonomy. *Paleontological Journal* **40:** S679–S737.

Reich D, Thangaraj K, Patterson N, Price AL, Singh L. 2009. Reconstructing Indian population history. *Nature* **461:** 489.

Rodríguez D, Cedeño-Vázquez JR, Forstner MR, Densmore III LD. 2008. Hybridization between *Crocodylus acutus* and *Crocodylus moreletii* in the Yucatan Peninsula: II. Evidence from microsatellites. *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology* **309:** 674–686.

Ross C, Garnett S. 1989. *Crocodiles and alligators*. New York: Facts on File Inc.

Ruta M, Pisani D, Lloyd GT, Benton MJ. 2007. A supertree of Temnospondyli: cladogenetic patterns in the most species-rich group of early tetrapods. *Proceedings of the Royal Society of London B: Biological Sciences* **274:** 3087–3095.

Sadleir RW, Makovicky PJ. 2008. Cranial shape and correlated characters in crocodilian evolution. *Journal of Evolutionary Biology* **21:** 1578–1596.

Sansom RS, Wills MA, Williams T. 2017. Dental data perform relatively poorly in reconstructing mammal phylogenies: morphological partitions evaluated with molecular benchmarks. *Systematic Biology* **66:** 813–822.

Sansom RS, Choate PG, Keating JN, Randle E. 2018. Parsimony, not Bayesian analysis, recovers more stratigraphically congruent phylogenetic trees. *Biology Letters* **14:** 20180263.

Schliep KP. 2011. Phangorn: phylogenetic analysis in R. *Bioinformatics* **27:** 592.

Scotland RW, Olmstead RG, Bennett JR. 2003. Phylogeny reconstruction: the role of morphology. *Systematic Biology* **52:** 539–548.

Sereno PC, Arcucci A. 1990. The monophyly of crurotarsal archosaurs and the origin of bird and crocodile ankle joints. *Neues Jahrbuch für Geologie und Paläontologie, Abhandlungen* **180:** 21–52.

Simmons MP, Reeves A, Davis JI. 2004. Character state space versus rate of evolution in phylogenetic inference. *Cladistics* **20:** 191–204.

Slater GJ, Goldbogen JA, Pyenson ND. 2017. Independent evolution of baleen whale gigantism linked to Plio-Pleistocene ocean dynamics. *Proceedings of the Royal Society of London B: Natural Sciences* **284:** 20170546.

Sookias RB, Benson RB, Butler RJ. 2012a. Biology, not environment, drives major patterns in maximum tetrapod body size through time. *Biology Letters* **8:** 674–677.

Sookias RB, Butler RJ, Benson RB. 2012b. Rise of dinosaurs reveals major body-size transitions are driven by passive processes of trait evolution. *Proceedings of the Royal Society of London B: Biological Sciences* **279:** 2180–2187.

Stadler T. 2011. Mammalian phylogeny reveals recent diversification rate shifts. *Proceedings of the National Academy of Sciences* **108:** 6187–6192.

Stanhope MJ, Waddell VG, Madsen O, De Jong W, Hedges SB, Cleven GC, Kao D, Springer MS. 1998. Molecular evidence for multiple origins of Insectivora and for a new order of endemic African insectivore mammals. *Proceedings of the National Academy of Sciences* **95:** 9967–9972.

Suh A, Smeds L, Ellegren H. 2015. The dynamics of incomplete lineage sorting across the ancient adaptive radiation of neoavian birds. *PLoS Biology* **13:** e1002224.

Suh A. 2016. The phylogenomic forest of bird trees contains a hard polytomy at the root of Neoaves. *Zoologica Scripta* **45:** 50–62.

Tarsitano SF, Frey E, Riess J. 1989. The evolution of the Crocodilia: a conflict between morphological and biochemical data. *American Zoologist* **29:** 843–856.

Trueman JW. 1998. Reverse successive weighting. *Systematic Biology*: 733–737.

Volkmann L, Martyn I, Moulton V, Spillner A, Mooers AO. 2014. Prioritizing populations for conservation using phylogenetic networks. *PLoS ONE* **9:** e88945.

Wang K, Lenstra JA, Liu L, Hu Q, Ma T, Qiu Q, Liu J. 2018. Incomplete lineage sorting rather than hybridization explains the inconsistent phylogeny of the wisent. *Communications biology* **1:** 169.

Watanabe A, Slice DE. 2014. The utility of cranial ontogeny for phylogenetic inference: a case study in crocodylians using geometric morphometrics. *Journal of Evolutionary Biology* **27:** 1078–1092.

Weaver JP, Rodríguez D, Venegas-Anaya M, Cedeño-Vázquez JR, Forstner MR, Densmore III LD. 2008. Genetic characterization of captive Cuban crocodiles (*Crocodylus rhombifer*) and evidence of hybridization with the American crocodile (*Crocodylus acutus*). *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology* **309:** 649–660.

Wiens JJ. 2004. The role of morphological data in phylogeny reconstruction. *Systematic Biology* **53:** 653–661.

Wilberg EW. 2015. What's in an outgroup? The impact of outgroup choice on the phylogenetic position of Thalattosuchia (Crocodylomorpha) and the origin of Crocodyliformes. *Systematic Biology* **64:** 621–637.

Wright AM, Hillis DM. 2014. Bayesian analysis using a simple likelihood model outperforms parsimony for estimation of phylogeny from discrete morphological data. *PLoS ONE* **9:** e109210.

Xu X, Zheng X, Sullivan C, Wang X, Xing L, Wang Y, Zhang X, O'Connor JK, Zhang F, Pan Y. 2015. A bizarre Jurassic maniraptoran theropod with preserved evidence of membranous wings. *Nature* **521:** 70.

Yang Z, Rannala B. 2012. Molecular phylogenetics: principles and practice. *Nature Reviews Genetics* **13:** 303.

Yonezawa T, Segawa T, Mori H, Campos PF, Hongoh Y, Endo H, Akiyoshi A, Kohno N, Nishida S, Wu J. 2017. Phylogenomics and morphology of extinct paleognaths reveal the origin and evolution of the ratites. *Current Biology* **27:** 68–77.

Zou Z, Zhang J. 2016. Morphological and molecular convergences in mammalian phylogenetics. *Nature Communications* **7:** 12758.