APPENDIX 2: PHYLOGENETIC ANALYSES AND RESULTS

The morphological and combined (morphological and molecular) datasets were each analysed using parsimony, undated Bayesian, and tip-dated Bayesian methods. The these 6 analyses, and the three subsequent exploratory analyses, are all discussed below, and the relevant executable files are on Dryad [Link provided automatically after MS upload].

1. Morphological data alone.

1.1. Parsimony Analysis.

Parsimony analysis used PAUP* (Swofford 2003), with all morphological character changes assigned unit weight (ie "unweighted"). Morphological characters which formed morphoclines were ordered (see Character List).

Goniopholis, an extinct taxon, was set as the outgroup to all other taxa (the ingroup), based on global analyses of crocodylomorphs (e.g. Turner and Pritchard 2015). The root position of the ingroup network (based on both morphology and, when included, DNA) is thus driven by the scoreable (56%) morphological characters in *Goniopholis*. Using a more distant, extant outgroup (e.g. from Aves) would allow the root position of the ingroup network to be determined by both morphological and DNA data. This advantage, however, is offset by several issues: (1) the problem of scoring morphological characters across highly divergent ingroups and outgroups ie crocodylians and birds, (2) the need to greatly increase taxon sampling – e.g. all major groups of dinosaurs - to capture highly transformed homologies between the ingroup and outgroup, and (in tip-dated Bayesian analyses) to satisfy the assumption of relatively even taxon sampling across clades and across time-slices, (3) the disparate divergences between the ingroup and outgroup – birds and crocodylians diverged ~250Ma, whereas crown crocodylians are <<120Ma – creates difficulties for molecular analyses, especially for reconstructing branch lengths. Molecular loci which are appropriate for reconstructing relationships and branch-lengths within crocodylians would tend to be saturated in crocodilian-bird comparisons (e.g. see Figure 2 in Gatesy et al. 2003).

Heuristic searches with multiple (20) random starting points were used, saving only 5000 trees per tree island (to avoid memory overflow on large tree island).

2248 most-parsimonious trees of length 1655 were found, and the strict consensus obtained (Fig. S1). As not all trees from all islands were saved, the number of MPTs will varied slightly when this search was repeated, but the same tree length and strict consensus was found. This analysis retrieved *Gavialis* as basal to all other living crocodylians, an arrangement robustly refuted by molecular data (e.g. Gatesy et al. 2003; Harshman et al. 2003; Oaks 2011). Thoracosaurs were placed with *Gavialis* (Fig. S5).

Bootstrapping (200 reps) was used to evaluate clade support; to reduce computation time to weeks rather than months, the above search setting were used but the number of random starting points was reduced to 10, and the number of saved trees per island was reduced to 1000. Bootstrap frequencies for each of the clades on the strict consensus are shown on Fig. S1. Trees for this and all other analyses were visualised and formatted with the aid of FigTree (Rambaut 2016).

The PAUP file with all MP search and bootstrapping settings is in Dryad (zip file: DataAndResults).

1.2. Undated Bayesian Analysis

Undated Bayesian analyses used MCMC as implemented in MrBayes 3.2 (Ronquist et al. 2013). The morphological characters were analysed using the Lewis (2000) model, with correction for non-sampling of constant characters. Stepping-stone sampling favoured the inclusion of gamma parameter for rate variation across characters (Bayes Factor *sensu* Kass and Raftery 1995 > 400). For rooting, *Goniopholis* was the outgroup to all other taxa.

Four independent runs (each with 4 incrementally-heated chains) were performed, each of length 20M with sampling every 20 000, resulting in 1000 samples/run. Burnin (50%) and convergence in numerical parameters was confirmed by superimposed traces and high (>>100) effective sample sizes in Tracer (Rambaut et al. 2014) as well as similar within- and between-run variances in MrBayes (PSRF or potential scale reduction factor was ~1). Convergence in topology was confirmed via MrBayes (standard deviation of split ie clade frequencies between runs was low, <0.02). The post-burnin samples were combined in MrBayes to generate summary statistics and a majority-rule consensus tree with posterior probabilities. This analysis retrieved *Gavialis* as basal to all other living crocodylians, an arrangement robustly refuted by molecular data (e.g. Gatesy et al. 2003; Harshmann et al. 2003; Oaks et al. 2012). Thoracosaurs were placed with *Gavialis* (Fig. S2).

The full MrBayes file with all MCMC run settings is in Dryad (zip file: DataAndResults).

1.3. Tip-Dated Bayesian Analysis

The tip-dated Bayesian analyses used used MCMC as implemented in BEAST 1.8 and related packages (Drummond et al. 2012). In all tip-dated Bayesian analyses, the morphological characters were analysed using the Mk-model with correction for non-sampling of constant characters (Lewis 2001; Alekseyenko et al. 2008), which has been well-tested (Wright and Hillis 2014; O'Reilly et al. 2016). To ensure all state transitions were weighted identically, even for characters with different numbers of observed states, a 6-state Mk model was employed for all traits. Partitioning characters according to the number of observed states can result in better model fit better has minimal effect on phylogeny, dates or ancestral state reconstruction (e.g. Gavryushkina et al. 2017; King et al. 2017). Furthermore, such state-partitioned models have potential drawbacks. (1) They increase the influence of characters with more states (e.g. King et al 2017). (2) Unless homoplasy is exceptionally rampant, the observed states of a character will not represent all possible states (Hoyall Cuthill 2015). Indeed, in related outgroups (Kespka et al. 2012), increased state-space can be observed for many of the characters modelled in Gavryushkina et al. (2017). As a further example, most sites in amino acid alignments will not exhibit all 20 possible states, and many DNA alignments will have sites with fewer than 4 states. Yet, amino acid or DNA alignments are not typically analysed by assuming that the observed states at each variable position are the *only* possible states.

Correction for non-sampling of constant characters (Alexsevenko et al. 2008; Wright and Hillis 2014) was implemented, along with the gamma parameter for rate variation

across characters (see previous section). However, correction for under-sampling of autapomorphies is impossible in most empirical datasets, though preliminary analyses suggest that this correction might not be vital (Matzke and Irmis 2017). Stepping-stone analysis favoured a relaxed (uncorrelated lognormal) clock over a strict clock (BF > 200), i.e there was significant variation in evolutionary rates across branches, hence an uncorrelated lognormal relaxed clock (Drummond et al. 2006) was implemented.

The most appropriate (birth-death serial-sampling) tree prior in BEAST 1.8 was used. Tip calibrations were employed using the stratigraphic age for each taxon, implemented as a point age midway between the upper and lower limits of any uncertainty (Table S2). Extant taxa were assigned an age of 0. No node calibrations were employed, not even for root age. For rooting, *Goniopholis* was the outgroup to all other taxa.

Four independent runs were performed, each of length 50M with sampling every 50 000, resulting in 1000 samples/run. Burnin (20%) and convergence in numerical parameters was confirmed by superimposed traces and high (>100) effective sample sizes in Tracer (Rambaut et al. 2014). Convergence in topology was confirmed via AWTY, with all clades appearing in similar frequencies across runs. The post-burnin samples were combined in LogCombiner, and TreeAnnotator was used to generate a consensus tree (maximum clade credibility) and associated summary statistics. This analysis retrieved *Gavialis* as basal to all other living crocodylians, an arrangement robustly refuted by molecular data (e.g. Gatesy et al. 2003; Harshman et al. 2003; Oaks 2011). Thoracosaurs were placed on the *Gavialis* lineage (Fig. S3).

The full MrBayes file with all MCMC run settings is in Dryad (zip file: DataAndResults).

2. Morphological and Molecular Data Combined.

2.1. Parsimony Analysis.

Parsimony analysis used PAUP* (Swofford 2003), with all morphological and molecular character changes assigned unit weight (ie "unweighted"). Morphological characters which formed morphoclines were ordered (see Character List). Heuristic searches with multiple (20) random starting points were used, saving only 5000 trees per tree island (to avoid memory overflow on large tree island).

3696 most-parsimonious trees of length 7560 were found, and the strict consensus obtained (Fig. S4). As not all trees from all islands were saved, the number of MPTs will varied slightly when this search was repeated, but the same tree length and strict consensus was found. This analysis still retrieved the problematic position of thoracosaurs within living gavials (Fig. S5).

Bootstrapping (200 reps) was used to evaluate clade support; to reduce computation time to weeks rather than months, the above search setting were used but the number of random starting points was reduced to 10, and the number of saved trees per island was reduced to 1000. Bootstrap frequencies for each of the clades on the strict consensus are shown on Fig. S4.

The PAUP file with all MP search and bootstrapping settings is in Dryad (zip file: DataAndResults).

2.2. Undated Bayesian Analysis

Undated Bayesian analyses used MCMC as implemented in MrBayes (Ronquist et al. 2013), with settings for the morphological data as discussed in the Morphology Only analyses above.

For the molecular data, partitioning schemes and substitution models were selected using PartitionFinder (Lanfear et al. 2012). Candidate partitions evaluated were: each codon for each locus (exons), each locus (introns, tRNAs and D-loop). This generated 53 candidate partitions. BIC with unlinked branch lengths and common ("MrBayes") substitution models were used to ascertain the appropriate number of partitions; these settings were used to avoid selecting overly complex partitioning schemes and models, which might be analytically redundant and also cause problems with achieving convergence in Bayesian analysis. The favoured partition scheme involved 3 partitions (substantially fewer than the maximum 53 possible) with the following substitution models. This scheme was intuitively sensible, with partition 1 including nearly all the nuclear partitions, partition 2 mainly including slower mtDNA partitions, and partition 3 including faster mtDNA partitions.

Partition 1 K80+G	ACTB_exon34_codon1, ACTB_exon34_codon2, ACTB_intron_codon3, ACTC_exon45_codon1, ACTC_exon45_codon2, ACTC_exon45_codon3, ACTC_intron_4, AChR_exon78_codon1, AChR_exon78_codon2, AChR_exon78_codon3, AChR_intron_7, CTB_exon34_codon3, GAPDH_exon1112_codon1, GAPDH_exon1112_codon2, GAPDH_intron_codon11, LDHA_intron_7, LDHB_exon67_codon1, LDHB_exon67_codon2, LDHB_exon67_codon3, LDHB_exon78_codon1, LDHB_exon78_codon2, LDHB_exon78_codon3, LDHB_intron_6, RHO_exon23_codon1, RHO_exon23_codon2, RHO_exon23_codon3, RHO_intron_codon2, aTROP_exon56_codon1, aTROP_exon56_codon2, aTROP_exon56_codon3, aTROP_intron_5, cmos_codon1, cmos_codon2, cmos_codon3, rag1_codon1, rag1_codon2, rag1_codon3
Partition 2 GTR+I+G	CYTB_codon1, CYTB_codon2, DLOOP, GAPDH_exon1112_codon3, ND2_codon1, ND2_codon2, ND3_codon1, ND3_codon2, tRNAarg, tRNAglu, tRNAgly, tRNAmet, tRNAtrp
Partition 3 GTR+I+G	CYTB_codon3, ND2_codon3, ND3_codon3

The combined analyses used 4 independent partitions (morphology and the 3 molecular partitions), with above 4 models, with separate relaxed (igr) clocks for the morphological and molecular partitions. Different overall rates for the 3 molecular partitions were accommodated using the rate scalar (ratepr=variable). For rooting, *Goniopholis* was the outgroup to all other taxa.

Four independent runs (each with 32 incrementally-heated chains) were performed, initially each of length 30M with sampling every 10 000, resulting in 1000 samples/run; because of slow convergence, this was extended to 50M and then 80M. Burnin (25%) and convergence in numerical parameters was confirmed by superimposed traces and high (>100) effective sample sizes in Tracer (Rambaut et al. 2014) as well as similar within- and between-run variances in MrBayes (PSRF or potential scale reduction factor was ~1). Convergence in topology was confirmed via near-identical consensus trees from different

runs (split frequencies across runs could not be monitored as each 32-chain run had to be allocated a different cluster). The post-burnin samples were combined in MrBayes to generate summary statistics and a majority-rule consensus tree with posterior probabilities. This analysis still retrieved the problematic position of thoracosaurs within living gavials (Fig. S5).

The full MrBayes file with all MCMC run settings is in Dryad (zip file: DataAndResults).

2.3. Tip-Dated Bayesian Analysis

The tip-dated Bayesian analyses used used MCMC as implemented in BEAST 1.8 and related packages, and the same BEAST morphological models as used for the tip-dated morphology-only analysis. The molecular data were combined and analysed using the same (PartitionFinder) models as in above undated analysis.

The most appropriate (birth-death serial-sampling) tree prior in BEAST 1.8 was used. Separate relaxed (uncorrelated lognormal: Drummond et al. 2006) clocks were allocated to the morphological and molecular partitions. Stepping-stone analyses favoured (BF ~60-370) this complex clock model over other possible clock models tested (a strict clock for morphology and molecules, or a strict clock for morphology and a relaxed clock for molecular data, or vice versa). For rooting, *Goniopholis* was again the outgroup to all other taxa.

Four independent runs were performed, each of length 50M with sampling every 50 000, resulting in 1000 samples/run. Burnin (20%) and convergence in numerical parameters was confirmed by superimposed traces and high (>100) effective sample sizes in Tracer (Rambaut et al. 2014). Convergence in topology was confirmed via AWTY, with all clades appearing in similar frequencies across runs. The post-burnin samples were combined in LogCombiner, and TreeAnnotator was used to generate a consensus tree (maximum clade credibility) and associated summary statistics.

In contrast to the other 5 analyses, this analysis robustly retrieved thoracosaurs as stem crocodylians, far removed from living gavials (Fig. S6). Thus, only tip-dated analysis of morphological and molecular data separates out these convergent forms: using tip-dating (without the molecular data), or using the combined morphological and molecular data (without tip dating) fails to identify this homoplasy. Rates of Evolution were fastest on the branch leading to *Gavialis* (Fig. S7).

The full BEAST file with all MCMC run settings is in Dryad (zip file: DataAndResults).

3. Testing the Temporal and Phylogenetic Signal in Thoracosaurs in the Tip-Dated Morphological+Molecular Analysis.

3.1. Removal of stratigraphic information from thoracosaurs causes them to again nest within gavials.

In order to test whether the basal position of thoracosaurs retrieved in the tip-dated, combined analysis (2.3) is at least partly driven by their stratigraphic age, this analysis was repeated, but with stratigraphic information deleted from thoracosaurs. The ages of the 7 thoracosaurs (*Argochampsa krebsi, Eogavialis africanus, Eosuchus lerichei, E. minor, Eothoracosaurus mississippiensis, T. macrorhynchus, T. neocesariensis*) were changed from fixed point ages, to a uniform prior spanning 0 to infinity, with relevant operators

introduced to allow the MCMC to explore this range. All other tip ages were retained. This analysis thus evaluates the phylogenetic position of thoracosaurs based only on their morphology, and allocates them an age most consistent with the phylogenetic position estimated from morphology alone.

All other aspects of the BEAST analysis remained the same as in the original analysis. Initial analyses using BEAST 1.8 did not converge, presumably due to the much greater array of possible trees allowed by the removal of stratigraphic information from thoracosaurs. However, analyses converged successfully in BEASTMC3, which employs multiple chains and heating (metropolis-coupling): each run employed 4 chains, incrementally heated at a value of 0.1.

This analysis retrieved all 7 thoracosaurs as crown gavials, related to *Gavialis,* with high support (pp=1.0, Fig. S8). Accordingly, the stratigraphic ages most consistent with this deeply nested position are young, with all 7 thoracosaurs being assigned ages of <15Ma.

The full BEAST MC3 file with all MCMC run settings is in Dryad (zip file: DataAndResults).

3.2. Thoracosaurs and the rest of the crocodylian fossil record have fundamentally inconsistent temporal signals

The tendency for the ancient thoracosaurs to nest within gavials seems inconsistent with the rest of the crocodylian fossil record, which suggests much shallower divergences across crocodylians. To test this, the BEAST analysis of the tip-dated, combined data (section 2.3) was repeated, but with two mutually-exclusive sets of fossils: thoracosaurs only (i.e. all non-thoracosaurs deleted), and non-thoracosaurs only (e.g. thoracosaurs deleted).

Non-Thoracosaurs only. This analysis retrieved relationships and divergence dates across the entire crocodylian tree which were essentially identical to the analysis including all fossils (non-thoracosaurs and thoracosaurs) (Fig. 4A). Deletion of thoracosaurs thus has very little effect on other phylogenetic relationships and divergence dates, exemplifying their wildcard nature.

Thoracosaurs only. This analysis retrieved thoracosaurs within gavials, as relatives of *Gavialis.* Accordingly, divergence dates across the entire crocodylian tree are very old (compared to the analysis where all fossils are included). For example, the *Gavialis-Tomistoma* divergence is nearly 90Ma, and crown Crocodylia is ~120Ma (Fig. 4B). Initial analysis in BEAST 1.8 failed to converge, but convergence was achieved in BEASTMC3, which employs multiple chains and heating (metropolis-coupling). Each run employed 4 chains, incrementally heated at a value of 0.1.

These two analyses illustrate that the two suites of fossils (non-thoracosaur and thoracosaur) imply very different timescales for crocodylian phylogeny. When all fossils are included, the resultant time scale is *not* an average of both; rather, the temporal signal from the larger suite of fossils (non-thoracosaurs) is preserved. The dissenting thoracosaurs are repositioned to where they are best able to conform to this tree and timescale.

The full BEAST and BEASTMC3 files with all MCMC run settings is in Dryad (zip file: DataAndResults).

4: Morphological Evidence for a Basal Position of Thoracosaurs

To evaluate the suites of characters supporting a basal versus nested position of thoracosaurs, the morphological data were optimised (using PAUP*) onto two alternative trees: the combined tip-dated tree (where thoracosaurs were basal to all living crocodylians) and the combined parsimony tree (where thoracosaurs were nested within gavials).

Both Templeton's test (p=0.009), and the winning-sites test (0.048), identified the morphological data as favouring the nested position of gavials, which implied many fewer morphological changes (lengths of 1676 vs 1714).

56 characters favoured a nested position: 2, 3, 6, 9, 11, 14, 15, 16, 17, 18, 20, 31, 37, 40, 42, 50, 54, 56, 59, 60, 64, 68, 69, 71, 74, 81, 83, 97, 105, 117, 128, 129, 132, 133, 138, 142, 144, 146, 152, 158, 166, 170, 174, 175, 180, , 184, 187, 200, 206, 212, 231, 233, 240, 254, 255, 268.

36 characters favoured a basal position: 12, 24, 33, 41, 44, 66, 67, 70, 72, 73, 78, 90, 95, 100, 104, 108, 111, 121, 124, 126, 154, 163, 165, 168, 183, 189, 216, 219, 225, 230, 232, 245, 256, 259, 265, 266.

Most of the above characters are homoplastic under both hypotheses. However, a few of the characters which favour a basal position are very compelling plesiomorphies that are occur in thoracosaurs and all other stem crocodylians, but largely or entirely absent in crown crocodylians including *Gavialis* (see main text). Thus, the basal position of thoracosaurs is not only consistent with the global crocodylian stratrigraphic record, but explains certain plesiomorphic traits of thoracosaurs that cannot be explained even as atavisms shared with *Gavialis*.

5: References for Phylogenetic Analyses

Alekseyenko, A.V., Lee, C., Suchard, M.A. Wagner and Dollo: a stochastic duet by composing two parsimonious solos. *Syst. Biol.* 57: 772-784. (2008)

Drummond, A.J., Ho, S.Y.W., Phillips, M.J., Rambaut, A. Relaxed phylogenetics and dating with confidence. PLoS Biology. 4: e88. DOI:10.1371/journal.pbio.0040088 (2006).

Drummond, A.J., Suchard, M.A. Bayesian random local clocks, or one rate to rule them all. BMC Biology 8: 114 DOI: 10.1186/1741-7007-8-114. (2010).

Drummond A.J., Suchard M.A., Xie D., Rambaut A., 2012 Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **29:**1969-1973 (DOI:10.1093/molbev/mss075) (2012).

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

Harshman J., Huddleston C.J., Bollback J.P., Parsons T.J., Braun M.J., 2003 True and false gharials: a nuclear gene phylogeny of Crocodylia. *Syst Biol.* 52:386-402

Hoyal Cuthill, J.F. The morphological state space revisited: what do phylogenetic patterns in homoplasy tell us about the number of possible character states? *Interface Focus* DOI: 10.1098/rsfs.2015.0049 (2016).

Nylander J.A.A., Wilgenbusch J.C., Warren D.L., Swofford D.L., 2008 AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics* 24:581-583 (DOI:https://doi.org/10.1093/bioinformatics/btm388)

Kass, R.E., Raftery, A.E. Bayes factors. *J. Amer. Statistical Assoc.* 90: 773-795 (1995).

King, B., Lee, M.S.Y. Ancestral state reconstruction, rate heterogeneity, and the evolution of reptile viviparity. *Syst. Biol.* 64: 532-544 (2015).

Ksepka, D.T., Fordyce, R.E., Ando, T., Jones, C.M.. 2012. New fossil penguins (Aves, Sphenisciformes) from the Oligocene of New Zealand reveal the skeletal plan of stem penguins. *J. Vert. Paleontol.* 32: 235–254 (2012).

Lewis P.O, 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. *Syst Biol.* 50:913-25.

Lanfear R., Calcott B., Ho S.Y.W., Guindon S. 2012. PartitionFinder: Combined Selection of Partitioning Schemes and Substitution Models for Phylogenetic Analyses. *Mol. Biol. Evol.* 6:1695-1701

Oaks J.R., 2011. A time-calibrated species tree of Crocodylia reveals a recent radiation of the true crocodiles. *Evolution* 65:3285-97

O'Reilly, J.E., Puttick, M.N., Parry, L., Tanner, A.R., Tarver, J.E., Fleming, J., Pisani, D., Donoghue, P.C. Bayesian methods outperform parsimony but at the expense of precision in the estimation of phylogeny from discrete morphological data. *Biol. Lett.* 12(4): 20160081. doi: 10.1098/rsbl.2016.0081 (2016).

Rambaut A., Suchard M.A., Xie D. & Drummond A.J. (2014) Tracer v1.6, Available from http://beast.bio.ed.ac.uk/Tracer

Rambaut A. 2016. Figtree v1.4.3. Available from http://tree.bio.ed.ac.uk/software/figtree/

Ronquist F., Teslenko M., van der Mark P., Ayres D.L., Darling A., Höhna S., Larget B., Liu L., Suchard M.A., Huelsenbeck J.P., 2012 MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61:539-542 (doi: 10.1093/sysbio/sys029)

Swofford D.L., 2003 *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods).* Version 4. Sinauer Associates, Sunderland, MA.

Turner, A.H., Pritchard, A.C. Observations on the morphology of *Isisfordia duncani* and its phylogenetic placement in Neosuchia. PeerJ 3:e759 (2015).

Wright A.M. Hillis D.M. Bayesian analysis using a simple likelihood model outperforms parsimony for estimation of phylogeny from discrete morphological data. *PLoS ONE* 9, e109210 (2014).