Early Vertebrate Evolution

Dynamic biogeographic models and dinosaur origins

Michael S. Y. LEE[1,2*], Matthew G. BARON[3,4], David B. NORMAN[3] and Paul M. BARRETT[4]

[1] Biological Sciences Building, College of Science and Engineering, Flinders University, GPO Box 2100, Adelaide, SA 5001, Australia.
  Email: mike.lee@flinders.edu.au
[2] Earth Sciences Section, South Australian Museum, North Terrace, Adelaide, SA 5000, Australia.
[3] Department of Earth Science, University of Cambridge, Downing Street, Cambridge CB2 3EQ, UK.
[4] Department of Earth Sciences, Natural History Museum, London, Cromwell Road, London SW7 5BD, UK.
*Corresponding author

ABSTRACT: A comprehensive analysis of early dinosaur relationships raised the possibility that the group may have originated in Laurasia (Northern Hemisphere), rather than Gondwana (Southern Hemisphere) as often thought. However, that study focused solely on morphology and phylogenetic relationships and did not quantitatively evaluate this issue. Here, we investigate dinosaur origins using a novel Bayesian framework uniting tip-dated phylogenetics with dynamic, time-sliced biogeographic methods, which explicitly account for the age and locality of fossils and the changing interconnections of areas through time due to tectonic and eustatic change. Our analysis finds strong support for a Gondwanan origin of Dinosauria, with 99 % probability for South America (83 % for southern South America). Parsimony analysis gives concordant results. Inclusion of time-sliced biogeographic information affects ancestral state reconstructions (e.g., high connectivity between two regions increases uncertainty over which is the ancestral area) and influences tree topology (disfavouring uniting fossil taxa from localities that were widely separated during the relevant time slice). Our approach directly integrates plate tectonics with phylogenetics and divergence dating, and in doing so reaffirms southern South America as the most likely area for the geographic origin of Dinosauria.

KEY WORDS: Dinosauria, mesozoic, palaeobiogeography, plate tectonics, tip dating, Triassic.

Dinosaurs are one of the most charismatic and heavily studied groups of vertebrates (e.g., Sereno 1999), but despite ongoing work and many recent discoveries, certain aspects of early dinosaurian evolution remain unclear. A recent study of early dinosaur relationships presented a new phylogeny that departs radically from existing models, but forms a broader framework for inferring dinosaur evolutionary history than those that were previously available (Baron *et al*. 2017a). While that study was focused primarily on early dinosaurian anatomy and relationships, a general observation was made: based upon the geographic locations of key taxa in their phylogeny, a Laurasian setting for dinosaur origin may have been plausible. This proposal ran contrary to an established consensus based on earlier phylogenies that favoured a Gondwanan and, more specifically, South American setting for this event (Nesbitt *et al*. 2009; Brusatte *et al*. 2010; Langer *et al*. 2010; Sues *et al*. 2011; Boyd 2015). In particular, the European presence of silesaurids, a close outgroup to dinosaurs, and some early dinosaur taxa (e.g., *Saltopus*) hinted that northern Pangaea may have played an important role in the earliest stages of dinosaur diversification. However, given the space limitations associated with the original article, no quantitative analyses were carried out to evaluate this alternative hypothesis rigorously.

Here, we test the alternative hypothesis implicit in Baron *et al*. (2017a) by analysing their early dinosaur dataset using a novel integration of time-sliced Bayesian biogeographic models (Landis 2016) and tip-dated Bayesian phylogenetics (Drummond *et al*. 2012; Bielejec *et al*. 2014; Gavryushkina *et al*. 2017). The current method simultaneously estimates phylogenetic relationships, divergence dates, evolutionary rates and biogeographic patterns. Biogeographical changes are inferred using models that account for the changing relationships of continents (and thus dispersal abilities) over time – a 'chronobiogeographical' paradigm (Hunn & Upchurch 2001) – and this information provides potentially critical information for improving estimates of tree topology and divergence dates. For comparison, we also perform analogous analyses using traditional biogeographic parsimony-based methods.

Previous studies have advocated a similar approach to integrating phylogeny, stratigraphy, morphology and biogeography in a parsimony framework (reviewed in Rossie & Seiffert 2006). However, that implementation was problematic: as parsimony trees lack an explicit time element, only very simple time-sliced biogeographic models were possible (e.g., three areas and three time slices) and tree topology and divergence dates had to be assessed manually. The analyses attempted here tackle this complex multidimensional problem more efficiently, using explicit Bayesian computational approaches.

1. Material and methods

1.1. Morphological, stratigraphic and biogeographic data

Morphological data were derived from the matrix of Baron *et al.* (2017a), with 457 characters coded for 73 taxa (supplementary File 1 available at https://doi.org/10.1017/S1755691018000920). *Euparkeria* was designated as the outgroup to all other taxa in all Bayesian and parsimony analyses, as previously. The fragmentary early dinosauriform *Nyasasaurus* (Nesbit *et al.* 2013) was not included as it is being re-studied and may be a chimera (M.G.B., P.M.B. and S. Nesbitt, unpub. data). Changes unique to terminal branches (autapomorphies) were explicitly sampled in this dataset, comprising >10 % of the characters, making this matrix especially appropriate for tip-dated Bayesian approaches (Lee & Palci 2015).

Stratigraphic and biogeographic data were obtained for each terminal taxon (supplementary File 2). These data are compiled and presented in full in Baron (2018, Appendix 1). The stratigraphic data consisted of ranges (minimum and maximum ages) rather than point ages, thus allowing the dated analyses to account for stratigraphic uncertainty. Biogeographic data consisted of locality information categorised into the 25 broad tectonic regions employed in Landis (2016), which explicitly models the changing connectivities of these regions through 26 epochs of geological time spanning the entire Phanerozoic.

1.2. Bayesian tip-dated analyses using time-sliced biogeographic models

A novel recent study (Landis 2016) united statistical biogeography with divergence dating: inferred dispersal events on a phylogeny would be expected to exhibit maximum concordance with palaeobiogeography when that phylogeny is correctly time scaled. That study used biogeographic data and time-sliced mode-weighted dispersal matrices (see below) to estimate absolute divergence dates, in the relatively simple situation involving an undated ultrametric tree of fixed topology and fixed relative branch lengths (based on published molecular analyses). In this article we demonstrate that this approach can be integrated into a more complex situation that involves a non-ultrametric, tip-dated phylogeny in which both topology and relative branch lengths are being co-estimated using additional data. In this more complex situation, the biogeographic information does not solely dictate absolute dates; rather, this has to be co-estimated in combination with other information on tip dates, morphological data and clock model, tree and fossil sampling, priors, etc. Since tree topology and relative branch lengths are not fixed, the biogeographic information can, and will, influence both aspects, especially for parts of the tree where the signal from other data is weak.

The morphological, stratigraphic and biogeographic data were simultaneously analysed using tip-dated Bayesian analyses (Drummond & Suchard 2010; Ronquist *et al.* 2012; Gavryushkina *et al.* 2017). The optimal evolutionary histories that explain all three data sources were inferred using Markov chain Monte Carlo (MCMC) approaches as implemented in BEAST 1.8.4. BEAST (Suchard *et al.* 2018) was used because it is increasingly widely employed for tip-dated phylogenetic analysis of time-sampled data including fossils (e.g., Drummond & Suchard 2010; Gavryushkina *et al.* 2017), can implement time-sliced biogeographic models (currently BEAST 1 only: Bielejec *et al.* 2014) and can also implement random local clocks (Drummond & Suchard 2010), which are suitable for assessing clade-specific rates of change in single traits such as biogeography (Sanders *et al.* 2013; King & Lee 2015). The executable BEAST XML files, with annotations describing model, prior, MCMC and logging settings, are in supplementary Files 3–5. The description below summarises the major points, focusing on the biogeographic data and models.

In all 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 & Hillis 2014; O'Reilly *et al.* 2016). 47 characters were autapomorphies (variable but parsimony uninformative), and only six were invariant (thus, invariant characters were essentially unsampled). Polymorphic morphological data (0 and 1) could be treated exactly as scored (e.g., 0 or 1 but *not* 2) not as total uncertainty (0 or 1 or 2): an improvement over earlier versions of BEAST. 418 morphological characters were treated as unordered and 39 characters forming morphoclines were treated as ordered (see list in Baron *et al.* 2017a).

We tested both an unpartitioned model and a state-partitioned model for the morphological data (Gavryushkina *et al.* 2017): both yielded similar results. We present both results but focus on the unpartitioned model, as employing a single five-state substitution Lewis model for all traits ensures all state transitions were weighted identically, even for characters with different numbers of observed states. Partitioning characters according to the number of observed states can result in better model fit, but has minimal effect on phylogeny, dates or ancestral state reconstruction on datasets tested thus far (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); and (2) unless homoplasy is exceptionally rampant, the observed states of a character might not represent all possible states (Hoyall Cuthill 2015). Indeed, in related outgroups (Ksepka *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 four states (amino acid or DNA alignments are not typically analysed by assuming that the observed states at each variable position are the *only* possible states).

Bayes factors (BF) (*sensu* Kass & Raftery 1995), calculated using stepping-stone sampling (Xie *et al.* 2011), found significant variation in rates of morphological evolution across different characters, favouring the inclusion of the gamma parameter (Yang 1994) for among-character rate variation (BF ∼130). We used a gamma rather than log-normal distribution (Harrison & Larsson 2015) because the gamma distribution can be parameterised to resemble any log-normal distribution (but not the reverse). There was also significant variation in rates of morphological evolution across different lineages, with a relaxed (uncorrelated log-normal; Drummond *et al.* 2006) clock preferred over a strict clock (BF ∼220). However, there was no evidence that different dinosauromorph lineages had different rates of dispersal: for the biogeographic data, a relaxed (random local; Drummond & Suchard 2010) clock did not fit better than a strict clock (strict clock slightly favoured, BF <5). Thus, all Bayesian analyses employed the gamma parameter and relaxed clock for the morphological data, but a strict clock for the biogeographic character.

The most appropriate available tree prior in BEAST 1.8 was used (birth–death serial sampling of Stadler 2010; sampled ancestor tree priors only available in BEAST2). Tip calibrations were employed using the full stratigraphic range for each taxon, with each taxon's age being represented by a uniform prior spanning minimum–maximum ages. In the focal analysis, a hard constraint on maximum root age was employed for the
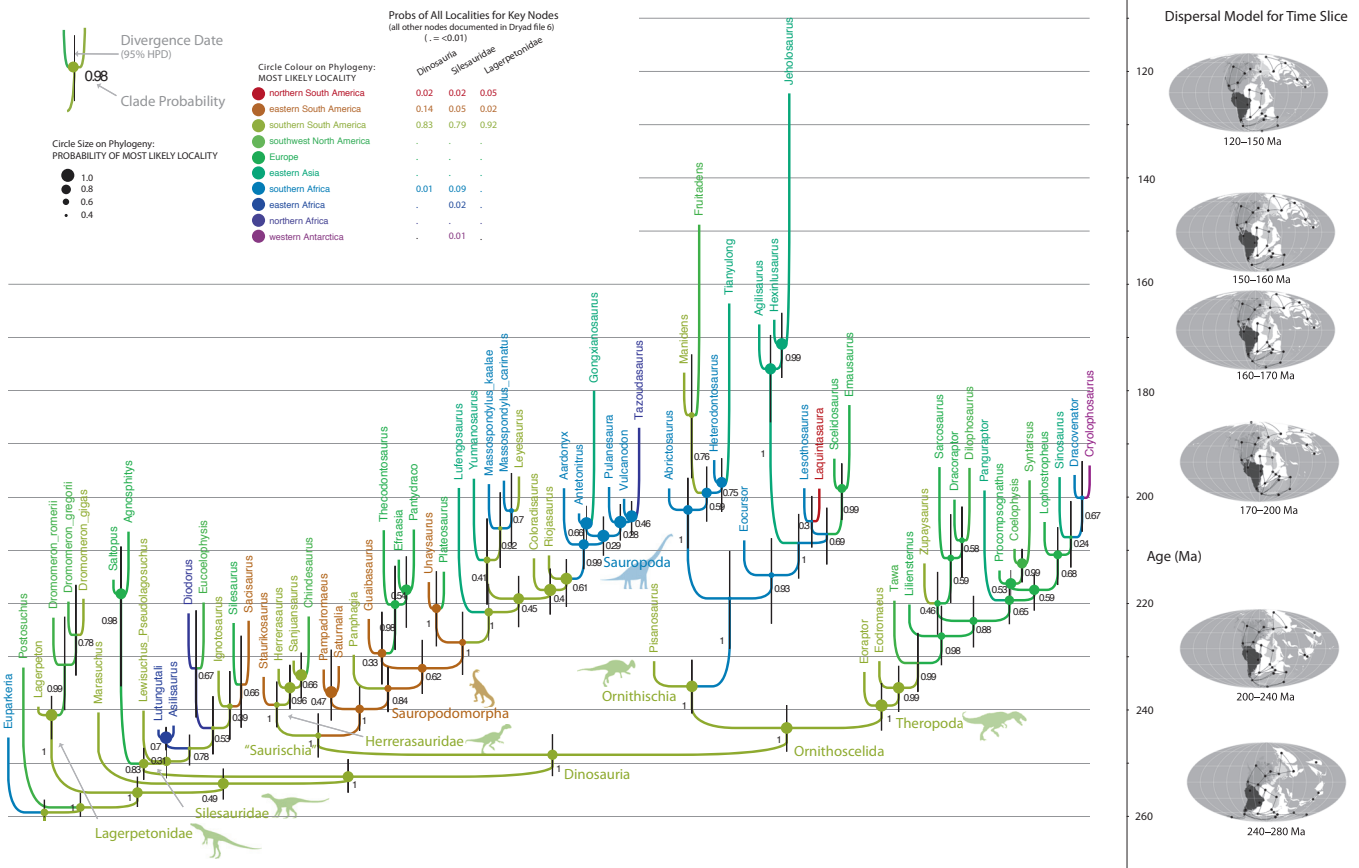
Figure 1 'Dynamic' time-sliced biogeographic model of early dinosaur evolution based on Bayesian tip-dated analysis of the data matrix in Baron *et al.* (2017a). Ancestral areas were inferred accounting for changing continental connections; dispersals between adjacent areas are favoured, but connectivities of areas change as per plate tectonics. The model recovers a South American origin of Dinosauria but with non-trivial probability for closely adjacent areas. (Lagerpetid silhouette by Nobu Tamara from phylopic.org; maps after Landis 2016; all other images the authors.)

main analyses described: 259.9 Ma. This represents the oldest *potential* fossil for the much larger clade Archosauriformes (Bernardi *et al.* 2015), and substantially exceeds the age of the oldest fossils analysed here. However, additional analyses, where the root age constraint was changed to a looser (soft exponential) constraint, did not change the phylogenetic and biogeographic conclusions. There was no explicit minimum root age constraint; this was dictated by the oldest included fossil (∼124 Ma).

The biogeographic data were treated as a discrete character in a separate partition, with 25 states corresponding to the tectonic provinces listed in Landis (2016), a necessary simplification due to the available functionality in BEAST (see Discussion). Two main Bayesian analyses were performed, in which the biogeographic data scored for the terminal taxa was treated in very different ways, to test the robustness of the results generated by different models.

1.2.1. Dynamic, time-sliced biogeographic model (Landis 2016). This model recognises that the rate of dispersal between the 25 areas increases with greater connectivity, i.e., decreasing distance and absence of water barriers (as dinosaurs are essentially terrestrial, e.g., Poropat *et al.* 2016). Twenty-five (25) areas have 300 possible connections between area pairs, assuming reversibility (above-diagonal elements of a 25×25 matrix). These 300 connections were classified as consistent with short-range, medium-range or long-range dispersal. Distances between area pairs, however, vary through time due to tectonic motion, eustatic fluctuations and changes to the areal extent of shelf seas, resulting in a dynamic or time-heterogeneous dispersal

model. Landis (2016) defined connectivities for 26 time slices or 'epochs': epochs 7–12 span the time interval relevant to this study (280–120 Ma) and are used here. They are here renumbered as epochs 1–6 respectively; the age boundaries of all epochs are shown in Figure 1 and their connectivities in supplementary Figure 1.

Short-, medium- or long-range dispersals are assigned separate relative rates; areas linked by short-range dispersal are *also* linked by medium- and long-range dispersal, areas linked by medium-range dispersal are *also* linked by long-range dispersal (Landis 2016, Fig. 4). Thus, under all possible parameterisations, this yields the intuitive pattern that dispersal rates between areas separated by short distances exceed the dispersal rates of those areas separated by medium distances, which in turn exceed the dispersal rates between areas separated by long distances. The summed (but still relative) dispersal rates between all 600 area pairs are thus encoded into a mode-weighted dispersal matrix for each of the 26 time slices (resulting in 26 dispersal matrices, of which matrices 7–12 are relevant here). In Landis (2016), the relative rates of short-, medium- or long-range dispersals were empirically estimated using RevBayes; when implemented in BEAST using the epoch substitution model (Bielejec *et al.* 2014) on a 25×25 reversible general substitution matrix (Lemey *et al.* 2009), these have to be fixed. We thus employ the empirical rates estimated in Landis (2016) that were for non-marine Testudines, a reptile group found in many of the same areas and time slices: short-range and medium-range dispersals have similar rates (average 0.49), and long-range dispersals
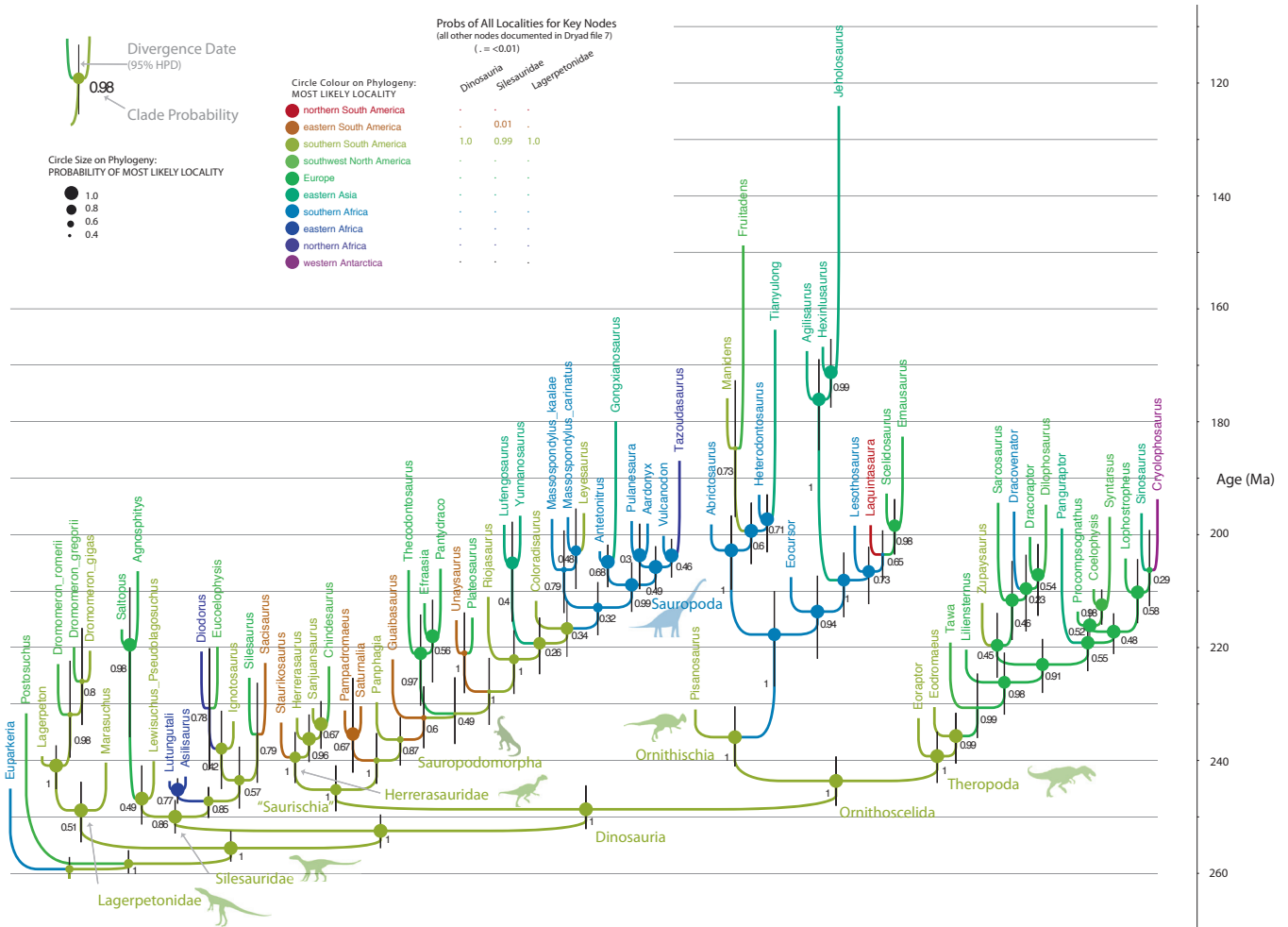
Figure 2 'Static' biogeographic model of early dinosaur evolution based on Bayesian tip-dated analysis of the data matrix in (Baron *et al.* 2017a). Ancestral areas were inferred without accounting for changing biogeographic connections, so taxa move directly between all areas with equal probability in all time slices. Southern South America is recovered as the ancestral area for Dinosauria with unrealistically high certainty. (Lagerpertid silhouette by Nobu Tamara from phylopic.org; all other images the authors.)

have a much lower rate (0.02). This results in the following summed relative rates between pairs of areas in each time-sliced mode-weighted dispersal matrix:

- area pairs separated by short distances (short + med + long range dispersal) = 1
- area pairs separated by medium distances (med + long range dispersal) = 0.51
- area pairs separated by long distances (long range dispersal only) = 0.02.

The mode-weighted dispersal matrices for the six epochs considered here (supplementary Fig. 1) are presented and annotated in the BEAST xml files.

As non-marine Testudines have traits that might be argued to make them more efficient overwater dispersers in comparison with non-avian dinosaurs (such as low metabolic rate, natural buoyancy), the analyses were repeated with lower values for medium and long range dispersal (0.1 and 0.01), with very similar results regarding topology, dates and ancestral areas.

1.2.2. Static, unordered model. This simple biogeographic model treats locality data in the same way as an unordered morphological character: every state (locality) can change to directly every other state (locality), and these changes happen at the same rate for all pairs of states. Thus, relative rates between all pairs of areas in the dispersal matrix are identical (arbitrarily set to unity), and this matrix also remains the same

across all time slices. This is an extremely simple biogeographic model that has been argued to be far from realistic (e.g., Hunn & Upchurch 2001): for instance, dispersals between contiguous land areas are modelled to occur at the same rate as transoceanic dispersals. However, it continues to be regularly applied in biogeographic studies including those using Bayesian methods (e.g., Schaefer *et al.* 2012). Furthermore, as the most 'distance-unweighted' and 'time-homogeneous' possible dispersal model, it represents an extreme alternative to the distance-weighted, time-heterogeneous model implemented above.

In the two analyses described above (which considered biogeography under very different models), ancestral states were logged for all sampled trees; thus, in the consensus trees, the relative probabilities for each of the 25 areas was calculated for each ancestral node. These probabilities are shown for three key nodes (Dinosauria, Silesauridae, Lagerpetidae) in Figures 1 and 2; the values for all other nodes are in the consensus tree files (supplementary Files 7–10), which are best viewed using FigTree. Finally, to test the influence of incorporating biogeographic data at all (regardless of model), the Bayesian analysis was also repeated with the morphological and stratigraphic data only.

All Bayesian analyses were run with a MCMC chain length of 50 million generations, with a burn-in of 20 % confirmed as adequate using Tracer (Rambaut *et al.* 2018) and AWTY (Nylander *et al.* 2008). Analyses were replicated four times to

confirm convergence, and the post-burn-in samples of all four runs concatenated for consensus trees and summary statistics with the aid of LogCombiner and TreeAnnotator (Drummond *et al*. 2012) and FigTree (Rambaut 2016). Executables to perform the above analyses are available as supplementary Files 3–6.

1.3. Parsimony analyses

Parsimony analysis and optimisation were also performed to compare the quantitative, probabilistic models with traditional cladistic approaches. We employed the standard sequential approach of parsimony analysis of phylogeny using morphological data alone, followed by optimisation of biogeography (e.g., Boyd 2015); integrated, time-sliced biogeographic methods developed in a parsimony framework (Rossie & Seiffert 2006) could not be implemented here due to the number of areas and time slices (which would have resulted in a step matrix with dimensions 150 × 150, containing 22,500 cells requiring manual calculation).

Parsimony analysis of the morphological dataset was performed using PAUP* (Swofford 2003), with all morphological character changes assigned unit weight (i.e., 'unweighted') and treated as ordered or unordered as per Baron *et al*. (2017a). For rooting, *Euparkeria* was the outgroup to all other taxa. Heuristic searches with multiple (1000) random starting points were used, saving only 2000 trees per tree island to avoid memory overflow on large tree islands. The executable PAUP file with all search settings is available in supplementary File 11.

Biogeography was optimised on all most-parsimonious trees, using parsimony and treating the character as a static trait with 25 unordered states (a dynamic, time-sliced dispersal cost matrix cannot be implemented; see above). Character optimisation used Mesquite (Maddison & Maddison 2017), via the function 'Analysis:Tree/Trace Characters Over Trees/ Parsimony'.

2. Results

The Bayesian analysis that incorporates biogeographic data under a dynamic model (informed by plate tectonics) recovers a tree topology (Fig. 1, supplementary File 7) very similar to that of the parsimony-based tree of Baron *et al*. (2017a, b; see also Parry *et al*. 2017). Notably, there is strong posterior probability (PP) ∼ 1 support for the two key clades: the theropod–ornithischian clade (Ornithoscelida) and herrerasaurid–sauropodomorph clade (a redefined 'Saurischia'). *Agnosphitys* is again recovered as a member of Silesauridae (see Baron *et al*. 2017a), but so is *Saltopus* (*contra* Baron *et al*. 2017b; *contra* Baron & Williams 2018; also note that Baron *et al*. (2017b) recover *Saltopus* as the sister taxon of *Marasuchus*, highlighting the unstable relationships of this taxon). The interesting position of *Marasuchus*, as an early diverging member of Lagerpetidae brings Lagerpetidae into Dinosauriformes, as currently defined (Nesbitt *et al*. 2017).

This analysis provides further strong support for a southern hemisphere origin of Dinosauria (Nesbitt *et al*. 2009; Langer *et al*. 2010, 2017; Brusatte *et al*. 2010; Sues *et al*. 2011; Baron *et al*. 2017b). The posterior probabilities of each locality being the ancestral area of dinosaurs are as follows: southern South America 0.83, eastern South America 0.14, northern South America 0.02 and southern Africa 0.01 (total South America 0.99). Despite the recovery of a handful of Northern Hemisphere taxa as early diverging members of Silesauridae and Dinosauria, only Southern Hemisphere areas are favoured as the point of origin for both of these dinosauriform clades

(Fig. 1). Performing the same analysis, but partitioning morphological characters according to the number of discrete states, gives very similar topological and biogeographic results (supplementary File 8).

The Bayesian analysis that incorporates biogeographic data but uses a static, unordered biogeographic model (Fig. 2; supplementary File 9) recovers a tree topology that is broadly similar to that in the dynamic analysis, notably robustly (PP ∼ 1) retrieving the ornithoscelidan and herrerasaurid–sauropodomorph clades. The position of *Marasuchus* and *Lewisuchus* varies slightly across analyses. However, the position of *Dracovenator* differs drastically between analyses. This static biogeographic model retrieves even more precise inferences for the origin of Dinosauria, with southern South America being virtually the only region sampled (PP > 0.999). Notable differences in topology (*Dracovenator*), and in ancestral states (higher confidence in South America), are likely due to the simplistic biogeographic model (see Discussion).

The Bayesian analysis that excluded biogeographic data also resulted in a tree (supplementary Fig. 2; supplementary File 10) very similar to both above analyses (that used biogeographic data); however, in terms of the affinities of *Dracovenator*, the tree was most similar to that obtained in the analysis using the static, unordered biogeographic model.

In the parsimony analyses, 128,000 most-parsimonious trees of length 1734 steps were found, and the strict and majority rule consensus trees are in supplementary Figure 3. The majority rule consensus was the same as one of the most-parsimonious trees. As not all trees from each optimal island could be saved (see above), the number of most-parsimonious trees varied slightly when this search was repeated, but the same optimal tree length and same consensus trees were found. The TNT analysis in Baron *et al*. (2017a) found the same shortest tree length (1734) but saved only 94 trees; their strict consensus was very similar to the strict consensus here, but the smaller sample of saved primary trees resulted in slightly (but artefactually) better resolution among early sauropodomorph taxa. The strict consensus trees from both analyses are otherwise identical.

In all 128,000 most-parsimonious trees, the following ancestral areas were reconstructed on the three key nodes (supplementary Fig. 3):

- Dinosauria: southern South America
- Silesauridae: southern South America
- Lagerpetidae: southern South America or southwest North America (equally parsimonious).

However, as parsimony approaches lack an explicit temporal component and cannot incorporate complex plate tectonic movements (see above), or quantify probabilities in inferred ancestral areas, our discussion focuses on the novel Bayesian analyses.

3. Discussion

Many of the oldest dinosaurs (primarily herrerasaurids and sauropodomorphs) occur in southern and eastern South America (Martinez and Alcober 2009; Brusatte *et al*. 2010; Ezcurra 2010; Langer *et al*. 2010; Martinez *et al*. 2011; Cabreira *et al*. 2011, 2016); it is unsurprising, therefore, that our dynamic and static Bayesian biogeographic analyses, and our parsimony analyses, strongly support a Gondwanan origin for Dinosauria. In addition, many of the earliest diverging ornithischian lineages (including heterodontosaurids, *Laquintasaura* and *Lesothosaurus*) are known from Gondwana (southern Africa and northern South America). However, as

these taxa date from the Early Jurassic they do not strongly affect predictions of the area in which dinosaurs originated. Although early dinosaur and dinosauromorph material has been reported from Late Triassic Northern Hemisphere localities (e.g., Yates 2003; Irmis *et al.* 2007; Niedźwiedzki *et al.* 2014), the inclusion of some of these taxa in our analyses had minimal influence on the predicted origin of Dinosauria itself (*contra* Baron *et al.* 2017a). However, the fragmentary nature of many early dinosaurs from the Northern Hemisphere precluded their inclusion in this analysis, so the potential effects of these taxa on ancestral area reconstruction remain untested. In addition, the oldest Gondwanan dinosaurs (e.g., *Eoraptor, Herrerasaurus*) are stratigraphically closer to the dinosaurian common ancestor than are their younger Laurasian counterparts (e.g., *Tawa, Chindesaurus*), and thus more strongly influence the inferred biogeographic area inhabited by this ancestor. A southern ancestry for Dinosauria is thus reconstructed, and northern forms are inferred to have dispersed into the areas where they were found. This was also the conclusion reached by Nesbitt *et al.* (2009) following discovery of the primitive North American Triassic theropod *Tawa*. While the tree topology in Baron *et al.* (2017a) is novel, the broad biogeographic implications remain consistent with inferences from other phylogenies (Nesbitt *et al.* 2009; Langer *et al.* 2010; Brusatte *et al.* 2010; Sues *et al.* 2011; Boyd 2015; Baron *et al.* 2017b; Langer *et al.* 2017).

Future work should include the problematic taxon *Nyasasaurus* in order to test its influence on early dinosaur biogeography, given its potentially Anisian age, east African occurrence and its status as either the earliest 'true' dinosaur or a very close dinosaur relative (Nesbitt *et al.* 2013; Baron *et al.* 2017a, b). Intuitively, it seems likely that *Nyasasaurus* would increase the probability of dinosaur origin occurring in eastern or southern Africa, but this hypothesis requires further testing following revision of the available material (M.G.B., P.M.B. and S. Nesbitt, unpub. data).

The close dinosaurian outgroups Silesauridae and Lagerpetidae are also found to have a Southern Hemisphere origin in the dynamic and static Bayesian analyses, and are inferred to have a similar biogeographic pattern to Dinosauria (Figs 1, 2). The positioning of *Saltopus* within Silesauridae is novel and, along with the recovery of *Agnosphitys* as another silesaurid (Baron *et al.* 2017a), provides further evidence for the presence of this clade in Europe and their likely global distribution in the Late Triassic. Our results furthermore suggest that the entire clade Dinosauromorpha has a southern origin, and that Northern Hemisphere forms are all (ultimately) descended from Southern Hemisphere ancestors, with lagerpetids, silesaurids, ornithischians, theropods and sauropodomorphs each spreading northward independently from areas that now form parts of South America and possibly Africa. However, it must also be noted that, according to a recent avemetatarsalian phylogeny (Nesbitt *et al.* 2017), many Laurasian taxa occupy key positions outside of Dinosauromorpha on the avemetatarsalian tree (almost all Triassic pterosaurs and one aphanosaurian). As the origin of pterosaurs would almost certainly optimise as Laurasian, this creates potential ambiguity around the place of origin for Ornithodira, given our results for Dinosauromorpha.

Incorporation of plate tectonics into biogeographic ancestral state reconstruction (Landis 2016) discernibly influences the probability of inferred ancestral areas, as could be predicted (e.g., Hunn & Upchurch 2001). Under the dynamic model, southern South America, eastern South America, northern South America and southern Africa have maximum connectivities in all time slices, resulting in all four areas having non-trivial probabilities of being the geographic origin of dinosaurs

(0.83, 0.14, 0.02, 0.01; total 0.982). Even though most early dinosaur fossils are found in southern South America, the high connectivities with the other three areas means an origin in one of the adjacent areas cannot be excluded. In contrast, the static biogeographic model does not allow for increased connectivities between particular areas and, as a result, reconstructs southern South America as the area of origin with (unreasonably) high certainty (>0.999).

Inclusion of biogeographic data and plate dynamics into tip-dated phylogenetics also appears to influence the retrieved tree topology, as found in similar studies using parsimony (Rossie & Seiffert 2006). The dynamic plate model disfavours uniting fossil taxa from localities that were remotely separated (at the time the organisms existed), because it implies either long-range dispersal during that time slice, or long ghost lineages extending into some earlier time slice where the localities were more adjacent. Such effects would be expected to be greatest where morphological data alone cannot robustly resolve affinities (Rossie & Seiffert 2006). Although the biogeographic trait is only a single character, the low rate for long-range dispersals allows it to exert relatively strong influence against certain tree topologies. Accordingly, the fragmentary South African theropod *Dracovenator* is recovered as the sister taxon of the (palaeogeographically proximal) Antarctic *Cryolophosaurus* in the dynamic biogeographical Bayesian analyses (Fig. 1), but as the sister taxon of the (palaeogeographically remote) English *Sarcosaurus* in the Bayesian analyses that either employ a static biogeographic model which does not penalise long-range dispersal (Fig. 2), or that ignores biogeography altogether (supplementary Fig. 2).

In our analyses, due to the current functionality of BEAST, biogeography had to be treated as a standard discrete character. This entails the simplifying assumption that each species-level lineage can only occupy one of the 25 broad geographic areas at any given time, with the concomitant assumption that when the lineage invades a new area, the former area is vacated by extinction. While neither of these assumptions can by logic be strictly accurate, the empirical data do seem to suggest that it might be suitable as a coarse approximation: all of the dinosaur species in this analysis are restricted to single areas, and even if one looks across all other dinosaurs, and considers genera (rather than species), very few dinosaur lower-level taxa occur across multiple areas: formerly wide-ranging morphotaxa are usually found to consist of multiple geographically localised, morphologically coherent taxa (e.g., Norman 1998; Rowe *et al.* 2011). Incorporation of models of character change that better capture range evolution (e.g., fusion of areas, vicariance) would further improve the analyses attempted here: such models are available in other packages (e.g., BioGeoBEARS: Matzke 2013), but these packages can currently only take prespecified trees as input, and optimise biogeography onto these trees. Co-estimation of trees and biogeography, as implemented here, is the only way to investigate whether biogeography can influence tree topology, and also more directly integrates phylogenetic uncertainty into biogeographic inferences.

Our novel analysis shows that tip-dated phylogenetic analysis can incorporate biogeography using time-sliced models accurately informed by continental drift. By disfavouring dispersals between remote areas, these models have the potential to improve both estimates of ancestral areas and refine phylogenetic relationships (Rossie & Seiffert 2006). Evolutionary history is thus simultaneously reconstructed using morphology, stratigraphy and biogeography.

4. Acknowledgements

5. Supplementary material

Supplementary material is available online at https://doi.org/10.1017/S1755691018000920.

6. References

Alekseyenko, A. V., Lee, C. & Suchard, M. A. 2008. Wagner and Dollo: a stochastic duet by composing two parsimonious solos. *Systematic Biology* **57**, 772–84.

Baron, M. G. 2018. The origin and early evolution of the Dinosauria. PhD Dissertation, University of Cambridge, UK. https://doi.org/10.17863/CAM.18898

Baron, M. G., Norman, D. B. & Barrett, P. M. 2017a. A new hypothesis of dinosaur relationships and early dinosaur evolution. *Nature* **543**, 501–06.

Baron, M. G., Norman, D. B. & Barrett, P. M. 2017b. Baron et al. Reply. *Nature* **551**, E4–E5.

Baron, M. G. & Williams, M. E. 2018. A re-evaluation of the enigmatic dinosauriform *Caseosaurus crosbyensis* from the Late Triassic of Texas, USA and its implications for early dinosaur evolution. *Acta Palaeontologica Polonica* **63**, 129–45.

Bernardi, M., Klein, H., Petti, F. M. & Ezcurra, M. D. 2015. The origin and early radiation of archosauriforms: integrating the skeletal and footprint record. *PLoS ONE* **10**, e0128449.

Bielejec, F., Lemey, P., Baele, G., Rambaut, A. & Suchard, M. A. 2014. Inferring heterogeneous evolutionary processes through time: from sequence substitution to phylogeography. *Systematic Biology* **63**, 493–504.

Boyd, C. A. 2015. The systematic relationships and biogeographic history of ornithischian dinosaurs. *PeerJ* **3**, e1523.

Brusatte, S. L., Nesbitt, S. J., Irmis, R. B., Butler, R. J., Benton, M. J. & Norell, M. A. 2010. The origin and early radiation of dinosaurs. *Earth-Science Reviews* **101**, 68–100.

Cabreira, S. F., Schultz, C. L., Bittencourt, J. & Langer, M. C. 2011. New stem-sauropodomorph (Dinosauria, Saurischia) from the Triassic of Brazil. *Naturwissenschaften* **98**, 1035–40.

Cabreira, S. F., Kellner, A. W. A., Dias-da-Silva, S., da Silva, L. R., Bronzati, M., Marsola, J. C. d. A., Müller, R. T., Bittencourt, J. d. S., Batista, B. J. A., Raugust, T., Carrilho, R., Brodt, A. & Langer, M. C. 2016. A unique Late Triassic dinosauromorph assemblage reveals dinosaur ancestral anatomy and diet. *Current Biology* **26**, 1–6.

Drummond, A. J., Ho, S. Y. W., Phillips, M. J. & Rambaut, A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biology* **4**, e88.

Drummond, A. J., Suchard, M. A., Xie, D. & Rambaut, A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* **29**, 1969–73.

Drummond, A. J & Suchard, M. A. 2010. Bayesian random local clocks, or one rate to rule them all. *BMC Biology* **8**, 114.

Ezcurra, M. D. 2010. A new early dinosaur (Saurischia: Sauropodomorpha) from the Late Triassic of Argentina: a reassessment of dinosaur origin and phylogeny. *Journal of Systematic Palaeontology* **8**, 371–425.

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

Harrison, L. B. & Larsson, H. C. E. 2015. Among-character rate variation distributions in phylogenetic analysis of discrete morphological characters. *Systematic Biology* **64**, 307–24.

Hoyal Cuthill, J. F. 2015. The morphological state space revisited: what do phylogenetic patterns in homoplasy tell us about the number of possible character states? *Interface Focus* **5**, 20150049.

Hunn, C. A. & Upchurch, P. 2001. The importance of time/space in diagnosing the causality of phylogenetic events: towards a 'chronobiogeographical' paradigm? *Systematic Biology* **50**, 391–407.

Irmis, R. B., Nesbitt, S. J., Padian, K., Smith, N. D., Turner, A. H., Woody, D. & Downs, A. 2007. A Late Triassic dinosauromorph assemblage from New Mexico and the rise of dinosaurs. *Science* **317**, 358–61.

Kass, R. E. & Raftery, A. E. 1995. Bayes factors. *Journal of the American Statistical Association* **90**, 773–95.

King, B., Qiao, T., Lee, M. S. Y., Min, Z. & Long, J. A. 2017. Bayesian morphological clock methods resurrect placoderm monophyly and reveal rapid early evolution in jawed vertebrates. *Systematic Biology* **66**, 499–516.

King, B. & Lee, M. S. Y. 2015. Ancestral state reconstruction, rate heterogeneity, and the evolution of reptile viviparity. *Systematic Biology* **64**, 532–44.

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. *Journal of Vertebrate Paleontology* **32**, 235–54.

Landis, M. J. 2016. Biogeographic dating of speciation times using paleogeographically informed processes. *Systematic Biology* **66**, 128–44.

Langer, M. C., Ezcurra, M. D., Bittencourt, J. S. & Novas, F. E. 2010. The origin and early evolution of dinosaurs. *Biological Review* **85**, 55–110.

Langer, M. C., Ezcurra, M. D., Rauhut, O. W. M., Benton, M. J., Knoll, F., McPhee, B. W., Novas, F. E., Pol, D. & Brusatte, S. L. 2017. Untangling the dinosaur family tree. *Nature* **551**, E1–E3.

Lee, M. S. Y. & Palci, A. 2015. Morphological phylogenetics in the genomic age. *Current Biology* **25**, R922–29.

Lemey, P., Rambaut, A. & Drummond, A. J. 2009. Bayesian phylogeography finds its roots. *PLoS Computational Biology* **5**, e1000520.

Lewis, P. O. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology* **50**, 913–25.

Maddison, W. P. & Maddison, D. R. 2017. Mesquite: a modular system for evolutionary analysis. Version 3.2 http://mesquiteproject.org.

Martinez, R. N., Sereno, P. C., Alcober, O. A., Columbi, C. E., Renne, P. R., Montanez, I. P. & Curre, B. S. 2011. A basal dinosaur from the dawn of the dinosaur era in southwestern Pangaea. *Science* **331**, 206–10.

Martinez, R. N. & Alcober, O. A. 2009. A basal sauropodomorph (Dinosauria: Saurischia) from the Ischigualasto Formation (Triassic, Carnian) and the early evolution of Sauropodomorpha. *PLoS ONE* **4**, e4397.

Matzke, N. J. 2013. Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Frontiers of Biogeography* **5**, 242–48.

Nesbitt, S. J., Smith, N. A., Irmis, R. B., Turner, A. H., Downs, A. & Norell, M. A. 2009. A complete skeleton of a Late Triassic saurischian and the early evolution of dinosaurs. *Science* **326**, 1530–33.

Nesbitt, S. J., Barrett, P. M., Werning, S., Sidor, C. A. & Charig, A. J. 2013. The oldest dinosaur? A Middle Triassic dinosauriform from Tanzania. *Biology Letters* **9**, 20120949.

Nesbitt, S. J., Butler, R. J., Ezcurra, M. D., Barrett, P. M., Stocker, M. R., Angielczyk, K. D., Smith, R. M. H., Sidor, C. A., Niedzwiedzki, G., Sennikov, A. G. & Charig, A. J. 2017. The earliest bird-line archosaurs and the assembly of the dinosaur body plan. *Nature* **544**, 484–87.

Niedźwiedzki, G., Brusatte, S. L., Sulej, T. & Butler, R. J. 2014. Basal dinosauriform and theropod dinosaurs from the mid–late Norian (Late Triassic) of Poland: implications for Triassic dinosaur evolution and distribution. *Palaeontology* **57**, 1121–42.

Norman, D. B. 1998. On Asian ornithopods (Dinosauria: Ornithischia). 3. A new species of iguanodontid dinosaur. *Zoological Journal of the Linnean Society* **122**, 291–348.

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**, 581–83.

O'Reilly, J. E., Puttick, M. N., Parry, L., Tanner, A. R., Tarver, J. E., Fleming, J., Pisani, D. & Donoghue, P. C. 2016. Bayesian methods outperform parsimony but at the expense of precision in the estimation of phylogeny from discrete morphological data. *Biology Letters* **12**, 20160081.

Parry, L. A., Baron, M. G. & Vinther, J. 2017. Multiple optimality criteria support Ornithoscelida. *Royal Society Open Science* **4**, 170833.

Poropat, S. F., Mannion, P. D., Upchurch, P., Hocknull, S. A., Kear, B. P., Kundrat, M., Tischler, T. R., Sloan, T., Sinapius, G. H. K., Elliott, J. A & Elliott, D. A. 2016. New Australian sauropods shed light on cretaceous dinosaur palaeobiogeography. *Scientific Reports* **6**, 34467.

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

Rambaut, A., Drummond, A. J., Xie, D., Baele, G. & Suchard, M. A. 2018. Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* **67**, 901–04.

Ronquist, F., Klopfstein, S., Vilhelmsen, L., Schulmeister, S., Murray, L. & Rasnitsyn, A. P. 2012. A total-evidence approach to dating with fossils, applied to the early radiation of the Hymenoptera. *Systematic Biology* **61**, 973–99.

Rossie, J. B. & Seiffert, E. R. 2006. Continental paleobiogeography as phylogenetic evidence. *In* Lehman, S.M. & Fleagle, J.G. (eds) *Primate biogeography*. 469–522. New York: Springer.

Rowe, T. B., Sues, H.-D. & Reisz, R. R. 2011. Dispersal and diversity in the earliest North American sauropodomorph dinosaurs, with a description of a new taxon. *Proceedings of the Royal Society B: Biological Sciences* **27**, 1044–53.

Sanders, K. L., Rasmussen, A. R., Mumpuni, Elmberg, J., de Silva, A. Guinea, M., Lee, M. S. 2013. Recent rapid speciation and ecomorph divergence in Indo-Australian sea snakes. *Molecular Ecology* **22**, 2742–59.

Schaefer, H., Hechenleitner, P., Santos-Guerra, A., de Sequeira, M. M., Pennington, R. T., Kenicer, G. & Carine, M. A. 2012. Systematics, biogeography, and character evolution of the legume tribe Fabeae with special focus on the middle-Atlantic island lineages. *BMC Evolutionary Biology* **12**, 250.

Sereno, P. C. 1999. The evolution of dinosaurs. *Science* **284**, 2137–47.

Stadler, T. 2010. Sampling-through-time in birth-death trees. *Journal of Theoretical Biology* **267**, 396–404.

Suchard, M. A., Lemey, P., Baele, G., Ayres, D. L., Drummond, A. J. & Rambaut, A. 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution* **4**, vey016.

Sues, H. D., Nesbitt, S. J., Berman, D. S. & Henrici, A. C. 2011. A late-surviving basal theropod dinosaur from the latest Triassic of North America. *Proceedings of the Royal Society B: Biological Sciences* **278**, 3459–64.

Swofford, D. L. 2003. *PAUP*. phylogenetic analysis using parsimony (*and other methods). version 4*. Sunderland, MA: Sinauer Associates.

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

Xie, W., Lewis, P. O., Fan, Y., Kuo, L. & Chen, M.-H. 2011. Improving marginal likelihood estimation for Bayesian phylogenetic model selection. *Systematic Biology* **60**, 150–60.

Yang, Z. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *Journal of Molecular Evolution* **39**, 306–14.

Yates, A. M. 2003. A new species of the primitive dinosaur *Thecodontosaurus* (saurischia: Sauropodomorpha) and its implications for the systematics of early dinosaurs. *Journal of Systematic Palaeontology* **1**, 1–42.