Originally posted 31 July 2014; corrected 1 August 2014

www.sciencemag.org/content/345/6196/562/suppl/DC1

Supplementary Materials for

Sustained miniaturization and anatomical innovation in the dinosaurian ancestors of bird

Michael S. Y. Lee,* Andrea Cau, Darren Naish, Gareth J. Dyke

*Corresponding author. E-mail: mike.lee@samuseum.sa.gov.au

Published 1 August 2014, *Science* **345**, 562 (2014)
DOI: 10.1126/science.1252243

This PDF file includes

Other supplementary material for this manuscript includes the following data archived on Dryad Digital Repository (doi:10.5061/dryad.jm6pj)

1. Character list for data set 1 in Word format [data set 2 is available from (*22*)].
2. Taxon list and sources of anatomical information for data set 1 in Word format [data set 2 is available from (*22*)].
3. Stratigraphic and size data for data sets 1 and 2 in Excel format.
4. Nexus file for data set 1 as a plain text executable for PAUP [data set 2 is available from *22*)].
5. xml file for data set 1, a plain text executable for BEAST 1.7 or 1.8.
6. xml file for data set 2, a plain text executable for BEAST 1.7 or 1.8.

Correction: The revised SM (submitted in time by the author) includes updates to the references.

Supplementary Materials (SM)

Materials and Methods

A. Phylogenetic Data sets

Two of the largest current data sets encompassing all theropods, with very different taxon and character sets, were analysed: patterns found in both data sets are thus likely to be general across all theropods. (1) **Data set 1** is new to this study and consisted of 120 theropod taxa and 1549 characters. It is the most character-rich phylogenetic matrix for theropods compiled, expanded from (*21*) though the addition of 49 new characters and 28 taxa (with 9 fragmentary or immature taxa removed). Bayesian methods ideally require sampling of all characters including those which are invariant (same across all taxa) and autapomorphic (unique specialisations of single taxa, typically excluded due to being parsimony-uninformative). This data set is the only theropod data matrix to date which considers such characters: of the 1549 characters, 114 were invariant and 184 were autapomorphies. The character list, sources of anatomical information, and the character-by-taxon matrix are presented in Section G (files 1-4). (2) **Data set 2** is a recent published data set consisting of 100 theropod taxa and 421 parsimony-informative characters (*22*; expanded from *8,23*). As with nearly all previously published phylogenetic data sets, autapomorphies were not sampled.

B. Stratigraphic, Body Size and Body Mass Data

Stratigraphic data for each taxon was obtained from the primary literature, with the most recently published, well-corroborated age used (Section G, file 3). Where published ages were given in stratigraphic units (e.g. stage or epoch), the dates for the relevant unit were taken from the ICS/IUGS International Stratigraphic Chart (www.stratigraphy.org/column.php?id=Chart/Time%20Scale).

Body size data for each taxon was obtained from the primary literature (Section G, file 3). Femur length is the most commonly used proxy for body size in non-avian theropods (e.g. *11, 15,28*); it was the single measurement most predictive of body size (r=0.995: *23*) and is also a commonly-preserved feature measurable in many taxa. It exhibits a relatively constant relationship across non-avialan theropods including non-avialan Paravians (*6*; see below). Only measurements from typical adult specimens were used; measurements from specimens which were identified as definitely or likely juvenile were not used (those taxa were scored as missing data if no other adult measurements were available). To reduce heteroscedasticity, all femur length measurements were $\log_{10}$ transformed (*24*); thus, a doubling of femur length along a branch always results in the same increase in this quantitative trait ($\log_{10} 2 = 0.301$), regardless of absolute ancestral and descendant size. Log_{10} Femur Length (hereafter FL_{10}) for all ancestral nodes was co-estimated simultaneously with phylogeny (in the Bayesian analyses), or optimized on the most-parsimonious tree (in the parsimony analyses). Conversion of femur length to body size was based on the tight (r=0.995) empirical relationship (*24*):

$$\text{Log}_{10}\ \text{BodyMass (kg)} = -6.288 + 3.222\ \text{Log}_{10}\ \text{Femur length (mm)}$$

We apply this formula to infer ancestral body sizes within theropods only up to Avialae (Fig. 2), where it is highly predictive of "size", whether measured as mass (*24*) or snout-vent length (*6*). We do not use it to infer size on branches within Avialae, due to changes in femur morphology and allometry in these taxa due to adaptations for flight and/or diving, and femoral reorientation (e.g. *6,13*). Application of bird-specific regressions might produce accurate estimates. Hence, our discussion focuses on body size evolution along the avian stem lineage leading up to Avialae, but not beyond.

It should be noted that use of this formula still returns reasonable weights even for basal Avialans such as *Archaeopteryx*: its femur length and the above regression yields a weight of ~453g, consistent with most studies (*e.g. 32, 33*). However, it contradicts the anomalously low weight of 132g presented in Table S4 in ref (*14*), which is almost certainly erroneous. This discrepancy might be related to some skeletal measurements of *Archaeopteryx* in Table S3 in ref (*14*) being considerably smaller than values used here (Section G, file 3) and elsewhere (*34*).

C. Bayesian Analyses (simultaneous inference of phylogeny and character evolution)

The BEAST package (*25*), which implements Markov-Chain Monte Carlo Bayesian methods for estimating phylogeny and associated traits, has 5 capabilities that make it uniquely applicable to this data set. In particular, BEAST models for inferring dated phylogenies using DNA from real-time virus samples are fundamentally analogous to models required for inferring dated phylogenies using morphological traits from fossils sampled across deep time. Similarly, "diffusion" models for inferring the geographic spread of viruses (in two dimensions) are broadly similar to Brownian motion models of body size evolution (in a single dimension). The other potentially relevant package, MrBayes (*35*) cannot implement methods 1, 2 or 5 from the list below, and was consequently not used.

(1) BEAST can simultaneously infer tree topology, divergence dates (lineage durations), and ancestral states for both discrete and continuous traits. All variables are co-estimated: for instance, all discrete and continuous traits directly contribute to the estimated phylogeny and divergence dates (continuous traits are not “mapped onto” a pre-determined phylogeny). However, in the current analysis, tree topology and branch lengths are largely determined by the discrete characters, due to the large number of discrete traits (1549 or 421) versus the single continuous trait (body size).

(2) It can implement likelihood-based models of evolution for both discrete and continuous morphological traits. Discrete characters are modeled using the Lewis (*35*) Markov model which allows ordering of multistate characters (if desired), and also accommodates variability in rates of evolution among characters (using the gamma distribution) and across lineages (using relaxed clocks, *26*). Continuous traits are modeled using a Brownian motion process (*37*); as with discrete characters, the “rate” of Brownian motion can be constant throughout the tree (strict clock) or can vary across lineages (relaxed clock).

(3) It assesses uncertainty for each parameter, taking into account the uncertanties for every other estimated parameter. For instance, uncertainty in body size reconstruction takes

into consideration not only uncertainty inherent in the chosen reconstruction model (e.g. rate-constant Brownian motion), but also uncertainty in tree topology and divergence dates.

(4) It can directly infer dated phylogenies where the terminal taxa differ in stratigraphic age, i.e. it estimates the optimal phylogeny and lineage durations that best explain the stratigraphic distribution and characters exhibited by the terminal taxa (*35,38*). Traditional phylogenetic methods generally first infer tree topology (branching sequence), and then estimates divergence dates, e.g. to match the stratigraphic dates and minimise ghost lineages. In contrast, BEAST (like MrBayes) simultaneously infers phylogeny and divergence dates which best fit the combined character and stratigraphic data (*35,38*)

(5) In addition to calibrating trees via tip ages (see point 4 above), it can also enforce traditional node calibrations, where the ancestor of a particular set of taxa is constrained to be a certain age (or age distribution), without topological constraints. Unlike all other dating programs, BEAST does not require enforcing the monophyly of calibrated groups. Thus, it is possible to calibrate a tree yet leave phylogenetic relationships totally unconstrained (to determined by the character data); the calibration applies to the most recent common ancestor of a given set of taxa, regardless of whether or not they form an exclusive clade. In contrast, other dating programs (including MrBayes) enforces monophyly of calibrated taxon sets.

Simultaneous estimation of evolutionary rates, topology and divergence dates is now a standard practice in molecular phylogenetics and has been argued to superior at identifying global optima that best fit all relevant parameters (e.g. *26*). In the current context, it should be noted that this approach yields conservative estimates of rate changes, by attempting to dampen extreme rates via changes in branch lengths or topology. In particular, the need to infer implausibly fast rates in sections of the tree could be removed either by minor stretching of very short branches (which barely affects overall tree shape), or by accepting a marginally inferior tree topology that is much more stratigraphically consistent (and which cannot be rejected by topology tests). In contrast, fixing topology and divergence dates before calculating rates will often retrieve extremely short branches with implausibly fast rates (at an extreme, zero-length branches with infinitely fast rates: e.g. *39*). Hence, simultaneous analysis of rates and tree shape results in lower (ie conservative) estimates of rate variability than sequential analysis (e.g. *40*).

Data sets 1 and 2 were analysed in BEAST, using the Lewis (*36*) Markov model for the discrete characters; characters which formed morphoclines were ordered (see Section G, file 4). All (ordered and unordered) discrete characters were treated as a single partition for estimating relevant parameters (e.g. mean rate, gamma). The most appropriate model for each data set was chosen using Bayes Factors (BF) *sensu* (*41*), i.e. twice the difference in marginal $\log_n$ likelihoods. The latter were estimated in Tracer (*42*), which implements the refinement by Suchard et al. (*43*). For both data sets 1 and 2, BF strongly favoured inclusion of the gamma parameter for among-character rate heterogeneity (BF 1302 & 125 respectively), and a relaxed (uncorrelated lognormal) clock over a strict clock for among-lineage rate heterogeneity (BF 2928 & 920 respectively). The relaxed clock analysis employed (see below) also returned very high variation in evolutionary rates across branches, again inconsistent with a strict clock (standard deviation of branch rates exceeding the mean,

and 95% HPD not abutting 0). The overall rate across the tree was given a very wide (conservative) uniform prior spanning 0 to 1000 changes per Ma (ie no change to extraordinarily fast rates). All characters were treated as independent. Character independence is a central assumption of all standard methods for phylogenetic inference (likelihood, parsimony, phenetics, and Bayesian). However, as organisms are integrated entities, this assumption is almost certainly violated in all real data sets (especially morphological ones), leading to potential errors such as over-confidence of related parameters, such as over-estimated clade probabilities.

The continuous trait (FL_{10}) was analysed using a Brownian model, with the tree-wide evolutionary rate/variance empirically estimated from the data using BEAST, using the relatively uninformative default prior (*37*). Analyses with relaxed clocks (branch-specific evolutionary rates) proved over-parameterized, with meaninglessly wide confidence intervals for rates on most branches. Additional (directed) models were tested with Bayestraits (*27*), which concluded the undirected model (as implemented in BEAST) was adequate.

Lineage durations (branch lengths) are integral to Brownian motion models, since large changes are less likely on short branches. Thus, the reconstructed ancestral value for a node will be most influenced by fossil taxa separated from that node by short branch lengths. For instance, the node representing the ancestral tyrannosauroid is reconstructed as small (FL_{10}=2.49, ~54kg; see Fig. S1), consistent with previous proposals (*44-46*). Even though two sampled tyrannosauroids (*Yutyrannus*, *Tyrannosaurus*) are huge, the small *Guanlong* (FL_{10}=2.54, ~81kg) is closest to the ancestral tyrannosauroid node in terms of branch lengths, and exerts a stronger influence on the reconstructed state.

Analyses were conducted using (1) only a single root age constraint, or (2) two additional internal constraints, on Paraves and Neotetanurae. Both approaches yielded qualitatively similar size and rate trends. The age of each constrained clade was given a uniform prior, between the maximum age (see below) and 0 Ma. In practice, clade ages younger than the oldest included taxon are not sampled; however, as clade content for varies across MCMC samples for internal nodes (because monophyly is not enforced), this effective younger bound varies for Paraves and Neotetanurae.

(1) The root age constraint consisted of the maximum age for Theropoda and was set at 246Ma, as this substantially pre-dates the earliest robust record of dinosaurs (~230Ma: *47-49*), and even the earliest potential dinosaurs (~243Ma: *50*). There is a rich global archosaur record in the Lower Triassic (~246-251Ma) which does not contain any unequivocal dinosaurs. (2) The first internal constraint consisted of a plausible upper limit (168.3 Ma) on the age of Paraves. This substantially pre-dates the earliest unequivocal paravians *Anchiornis*, *Aurornis* and *Xiaotingia* at ~159Ma (see Section G, file 3), and is the same age as the oldest potential evidence for paravian-like taxa, consisting of footprint evidence acknowledged to be of questionable stratigraphy and taxonomic affinity (*51*). (3) The second internal constraint consisted of a plausible upper limit (175Ma) on the age of Neotetanurae. This substantially pre-dates the earliest unequivocal neotetanurans, the allosauroids *Xuanhanosaurus qilixiaensis*, *Yangchuanosaurus zigongensis* and *Shidaisaurus jinae*, all of undetermined age within the Middle Jurassic (*52*), and the coelurosaur *Proceratosaurus bradleyi* from the Bathonian (*53*). There is a rich Lower Jurassic (175-200Ma) theropod fossil record that does

not include any neotetanurans or even undisputed tetanurans. The oldest undisputed tetanurans are ~175Ma (*54*), though there are potential tetanurans at about 196Ma (e.g. *Eshanosaurus*: *55*).

The BEAST files for data sets 1 and 2 (with size and age data) are in Section G (files 5 and 6). All analyses were performed in BEAST 1.7 and 1.8 (*25*), on the e-research SA (erSA) computer grid. Each BEAST analysis involved 6 replicate runs (with different random starting trees and random number seeds). Each of the 6 replicate runs involved 30 million steps with sampling every 5000 generations, with a burnin of 5 million steps. Convergence (stationarity) in numerical parameters was identified using Tracer (*42*): broadly overlapping, non-trending traces across all replicate runs for every parameter, with effective sample sizes (ESS) of every parameter exceeding 100. Convergence for both data sets was reached before the relevant burnin, and the post-burnin parameter and tree samples were retained for analysis and concatenated using LogCombiner in the BEAST package. Estimates (mean and 95% highest posterior density) for all numerical parameters were generated using Tracer (*42*). Convergence in topology was assessed using AWTY (*56*), with posterior probabilities of splits of post-burnin trees always highly correlated across the replicate runs. The maximum clade credibility (MCC) consensus tree using mean branch lengths was obtained using TreeAnotator in the BEAST package (*25*), together with estimates (mean and 95% highest posterior density) of tree-based parameters, including posterior probabilities, divergence dates, lineage durations (branch lengths), rates of morphological evolution (discrete characters), and ancestral state reconstructions for femur length / inferred body size (continuous character). The final summary trees with node values were generated using TreeAnotator and visualized via FigTree (*57*).

Data set 1. The dated maximum clade credibility (MCC) tree (with branch lengths in Ma) is shown in Figs 1 & S1 with inferred ancestral states for femur length/body size, in Fig. S2 with posterior probabilities for each clade, and in Figs 3 & S3 with rates of evolution on each branch (numerical values, in % divergence per Ma). These results are discussed in the main text. The dated MCC with branch lengths in terms of amount of morphological evolution is shown in Fig. S4; despite the elevated rates along the bird stem lineage, the absolute amounts of change along the branches near the origin of Avialae are rather small, emphasising the morphological continuum between birds and non-avian dinosaurs (e.g. *1,2,12,14,21*).

Data set 2. The dated MCC tree (with branch lengths in Ma) is shown in Fig. S5 branches coloured according to reconstructed body size; absolute values are also given. A pattern of consistent, unreversed size reduction along most of the avian stem (from Tetanurae upwards) is again found. Fig. S6 shows the tree branches coloured according to rates of change, along with rates for each branch. Because this data set did not sample autapomorphies, rates on terminal branches are underestimated. Even so, the pattern is similar to data set 1: the Avian stem exhibits consistently faster rates of evolution than the rest of the tree, even when only internal (non-terminal) branches are considered.

D. Testing for Trends: PGLS

The presence or absence of an overall trend towards size decrease (or increase) in the data set was determined using PGLS, one of the most powerful methods for detecting temporal evolutionary trends (*58*). BayesTraits (*27*) implements a range of models of continuous trait evolution, in a fully Bayesian framework integrating across different sampled tree topologies and branch lengths. An undirected Brownian motion model was tested against a model which also included a trend parameter (beta); in the latter model, the size of each descendant node is predicted jointly by the ancestral node, the estimated rate of Brownian motion ("variance"), and the trend parameter. Analyses employed 1000 primary (ie sampled) trees from each data set, and default BayesTraits priors. Both data sets exhibited no significant overall trend towards larger or smaller body size across all theropods (Fig. S7AB). Adding the trend parameter did not improve model fit (data set 1 - BF =0.14; data set 2 - BF = 0.07), and the estimated trend parameter was insignificant (Figure S7AB), with the 95% HPD interval broadly encompassing 0 (data set 1, mean=0.0015, HPD = -0.0012 to +0.0042; data set 2, mean = 0.0011, HPD = -0.0017 to +0.0037). Thus, both data sets do not exhibit a trend towards body size increase or decrease with time, when the overall tree is considered. This is consistent with recent results for three theropod subclades, which showed no directionality in these clades despite expected size trends due to herbivory (*28*).

To test for a significant trend on the bird stem lineage, we repeated the PGLS analysis (*27*) on the relevant subtree in each of the 1000 sampled trees. This subtree spanned the outgroup to the bird clade; all other taxa were pruned from each sampled tree (Fig. S7C). Only data set 1 sampled sufficient bird taxa to employ this test. This test is also conservative because the tested trend along the bird stem lineage will be diminished by any non-directionality on the other remaining branches (i.e. branches leading to the outgroup, and within birds). To reduce this effect, bird taxa on very long tip branches were also pruned (Fig. S7C). Despite the conservative nature of this test, addition of the trend parameter significantly improved model fit (BF =18.6) and accordingly the estimated trend parameter was significantly negative (mean = -0.0066, HPD = -0.0120 to -0.0007: Fig. S7D), indicating a significant trend of size reduction with time.

E. Testing for Trends: Parametric simulations

Parametric simulations ("bootstraps") were performed to test if the observed consecutive run of size reductions along the bird stem lineage could have been generated stochastically, under a null model where size increases or decreases randomly across the entire phylogeny.

Missing tip data can inflate inferred trends, because inferred ancestors at certain nodes are directly inferred from nodes above and below (rather than via tip data). To remove this bias, the phylogeny in Fig. S1 was pruned down to the 87 taxa which all had size (FL_{10}) information. We then inferred the evolution of size along this phylogeny using the Bayesian analyses discussed above, on this pruned, fixed topology. These analyses retrieved a pattern identical to that depicted in Figs 1 and 2: the longest "run" of consecutive size reductions in this 87-taxon tree was 27, and spanned all of the bird stem lineage.

We then simulated the evolution of size along this phylogeny, using an undirected Brownian motion model in Mesquite (*59*); the root value (2.679) and rate of change / variance

(0.045 per Ma) matched that inferred by BEAST from the actual data (BEAST's retrieved rate of 7.7 scaled by tree height of 173). We also confirmed that the simulations using a rate/variance of 0.045 yielded variation in tip values very similar to the actual values (e.g. similar range between largest and smallest sizes). Because the actual ancestral values for the real tree are not known (only tip values are known, ancestral values are inferred via the Bayesian analyses described above), we treated the simulated data in exactly the same way. We used the tip values at the simulations and inferred nodal values using the same methods used for the real data. Only 20 such simulations could be completed due to high computational burden. Across these 20 simulations, longest run of size increases was 14 (mean = 8.5), and longest run of decreases was 20 (mean = 10.7); the observed value of 27 decreases exceeds both these values substantially. These results suggest that the trend observed in the real data is too persistent to have been generated by chance, and is significant at least at to P=0.05 (the smallest value that can be retrieved given 20 simulations).

F. Parsimony Analyses and Character Optimisation

Parsimony analysis and optimization was also employed, to test the robustness of the above trends to alternative methodologies. These methods are very different to the Bayesian likelihood methods above: for instance, in a parsimony framework, temporal duration (length) of branches is irrelevant to both phylogenetic inference or character optimization.

Each data set was analysed in PAUP* (*60*), using search settings aimed at sampling as many tree islands as possible [HSEARCH addseq=random nreps=1000 nchuck=1000 chuckscore=1]. Both data sets 1 and 2 resulted in >100000 most parsimonious trees (MPTs); many more presumably exist but could not be retained due to memory constraints. However, the strict consensus tree obtained for data set 2 matched that from a previous parsimony analysis of the same data (*22*), suggesting the correct overall consensus topology was retrievable from the pools of sampled trees in both analyses. The strict consensus for both data sets contained large polytomies, and characters should not be optimized on such consensus trees, as they are not optimal trees and thus imply more homoplasy that any of the individual MPTs (e.g. *61*). Hence, a majority-rule consensus was obtained from the sampled MPTs (Figs S8, S9), which was fully resolved.

Body size was optimized onto the trees from data sets 1 and 2 using both linear and square-change parsimony in Mesquite (*59*). Both methods produced similar results, so only linear parsimony results are presented (Figs. S8, S9). Both data sets 1 and 2 indicate that body size consistently decreased, or remained unchanged, along every branch of the bird stem lineage from neotetanurans upward. Similar trends have been obtained with parsimony analysis of a data set of tetanurans (*7*). Such parsimony analyses also reveal size increases and decreases are otherwise scattered across theropods, even in the nearest relatives of birds [(*8*) see also (*62*)]. Thus, the parsimony analyses are consistent with the more parameterized Bayesian models, and demonstrate that the trends observed here (sustained miniaturisation confined to the bird stem lineage) are robust to different models and assumptions.

G. List of Data Files Archived on Dryad Digital Repository (doi:10.5061/dryad.jm6pj)

1. Character list for data set 1 in Word format [data set 2 is available from (*22*)].
2. Taxon list and sources of anatomical information for data set 1 in Word format [data set 2 is available from (*22*)].
3. Stratigraphic and size data for data sets 1 and 2 in Excel format.
4. Nexus file for data set 1 as a plain text executable for PAUP (data set 2 is available from *22*).
5. xml file for data set 1, a plain text executable for BEAST 1.7 or 1.8.
6. xml file for data set 2, a plain text executable for BEAST 1.7 or 1.8.

SUPPLEMENTARY FIGURES S1-S9

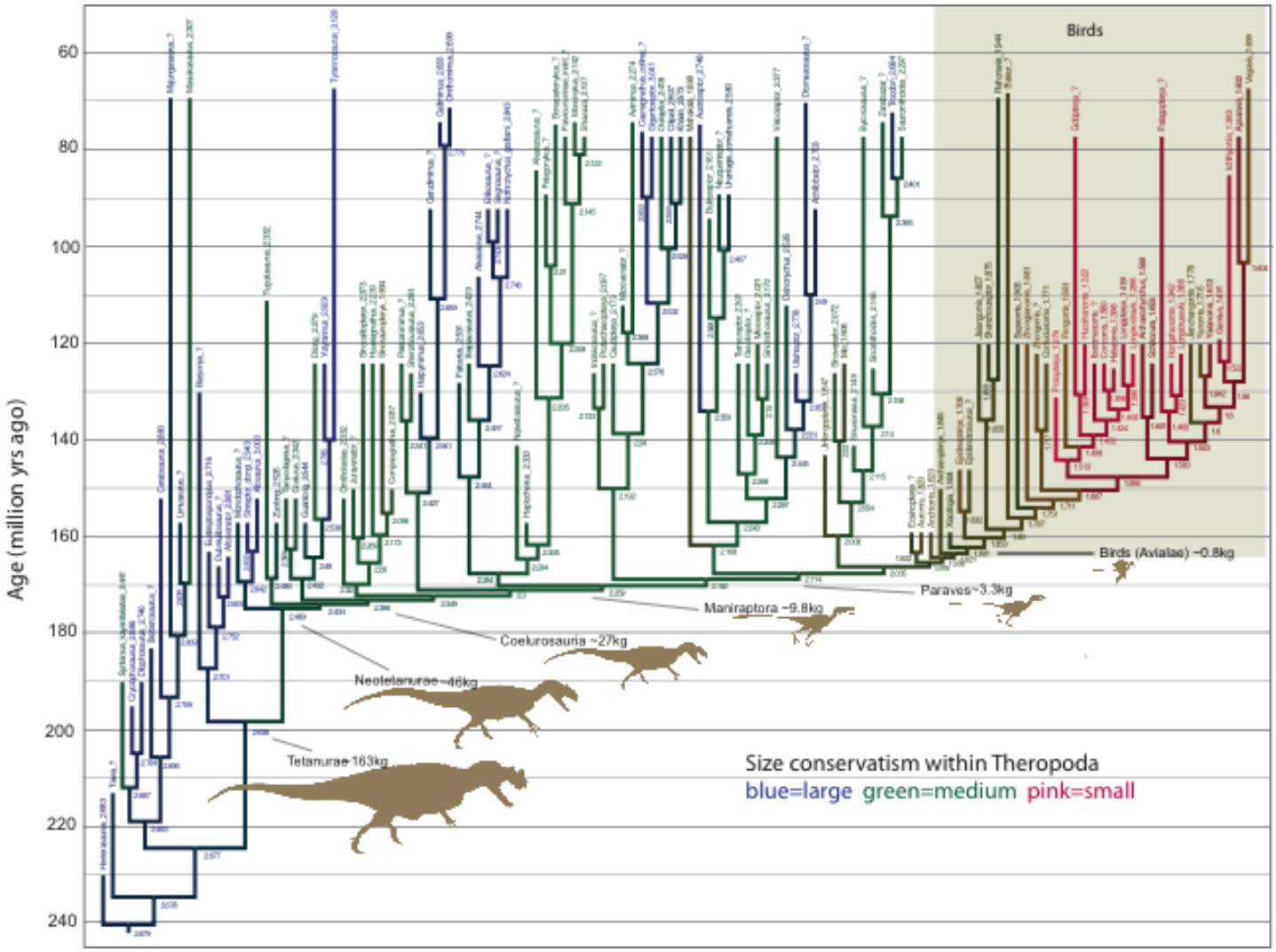

Fig. S1. **Size evolution across the theropod-bird transition.** This is more detailed version of Fig. 1, and represents the Bayesian maximum clade credibility (MCC) tree from data set 1, with size information superimposed. All taxon names are shown, and size (indexed by $\log_{10}$ femur length) is shown for all tip taxa (observed values) and all nodes (reconstructed values). Silhouettes used in all figures from phylopic.org.

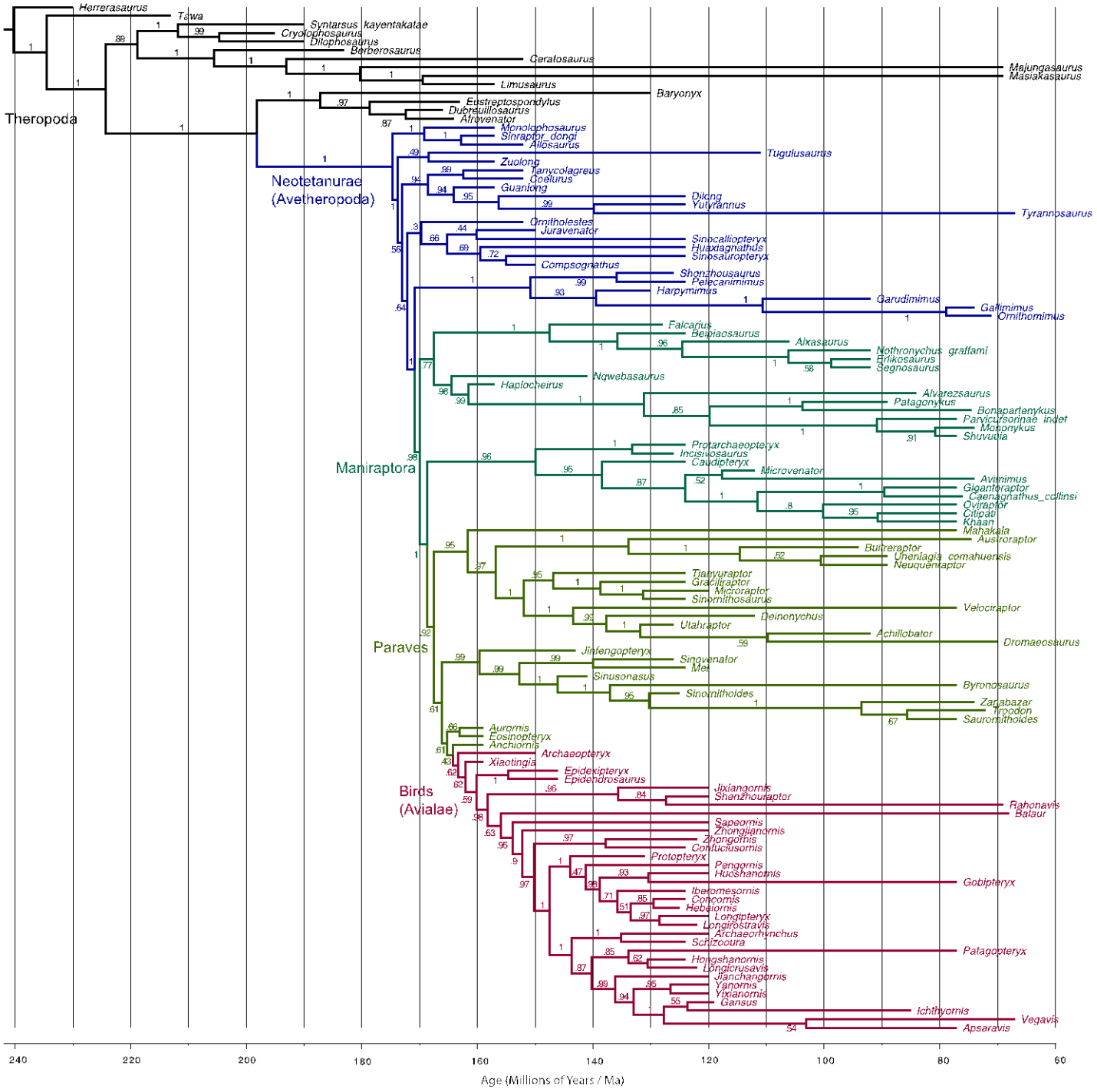

Fig. S2. **Theropod phylogeny, based on the new morphological data matrix (data set 1).** This tree is the Bayesian MCC tree with posterior probabilities shown at nodes. Tree is colour-coded by clade: Black (and upwards) = Theropoda, Blue (and upwards) = Neotetanurae, Green (and upwards) = Maniraptora, Olive (and upwards) = Paraves, Pink = birds (Avialae / Aves *sensu lato*).

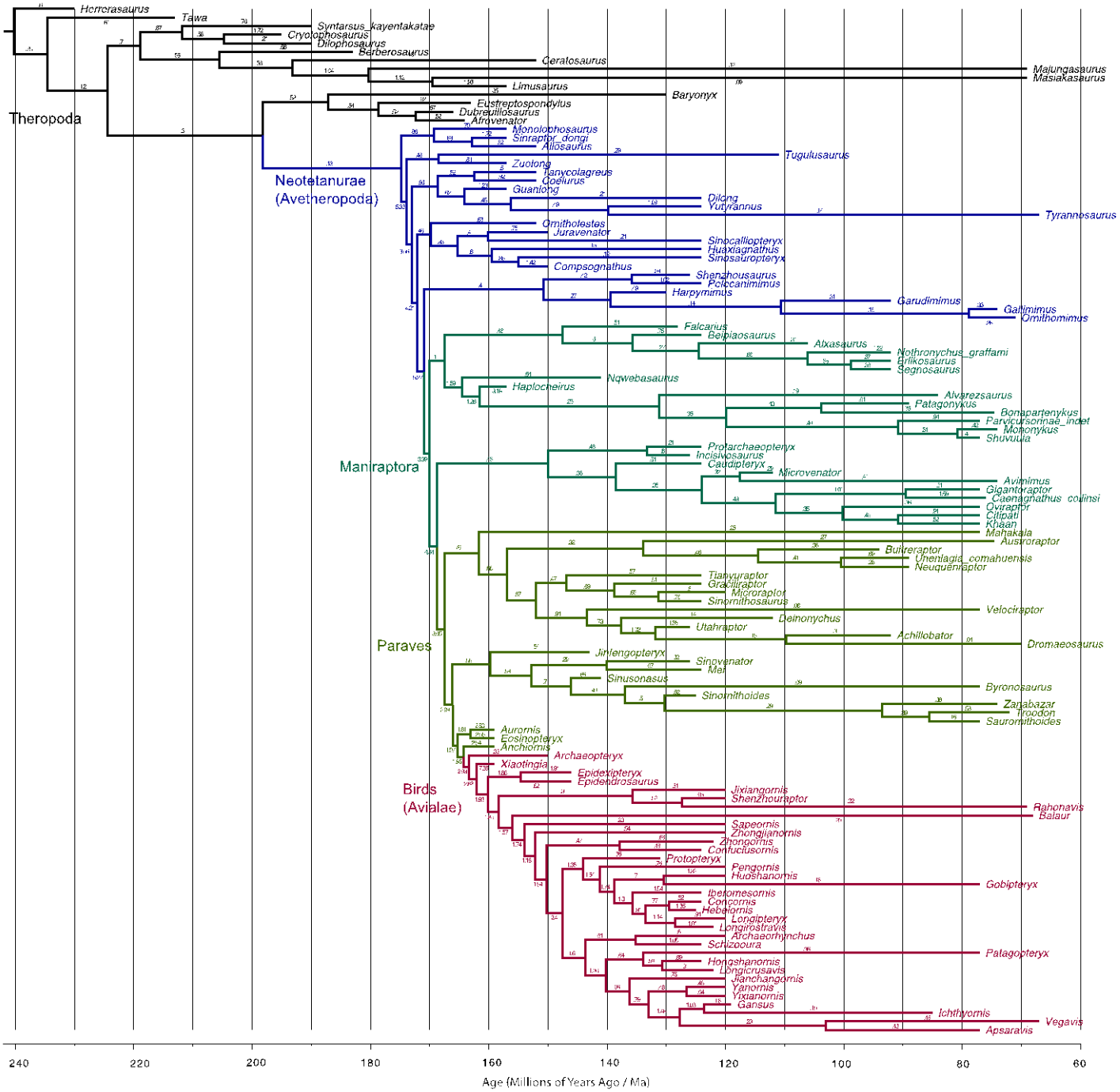

Fig. S3. **Rates of morphological evolution in theropods, based on data set 1.** Bayesian MCC tree with mean evolutionary rates shown on each branch (percentage divergence per million years, across all 1549 discrete characters; a rate of 1% equates to a 0.01 probability of change per character per lineage per million years). The bird stem lineage is consistently faster than the rest of the tree, with the fastest rates occurring between Neotetanurans and Paraves (see also Fig. 3, where branches are colour-coded according to rate). Tree is colour-coded by clade (see Fig. S2).

Fig. S4. **The dinosaur-bird continuum**: **Amounts of morphological evolution in theropods, based on data set 1.** Bayesian MCC tree, with branch lengths scaled to the absolute amount (rather than rate) of evolutionary divergence across all discrete characters. The branch leading to "birds" (Avialae, Aves *sensu lato*) does not undergo exceptional amounts of evolution. Tree is colour-coded by clade (see Fig. S2).

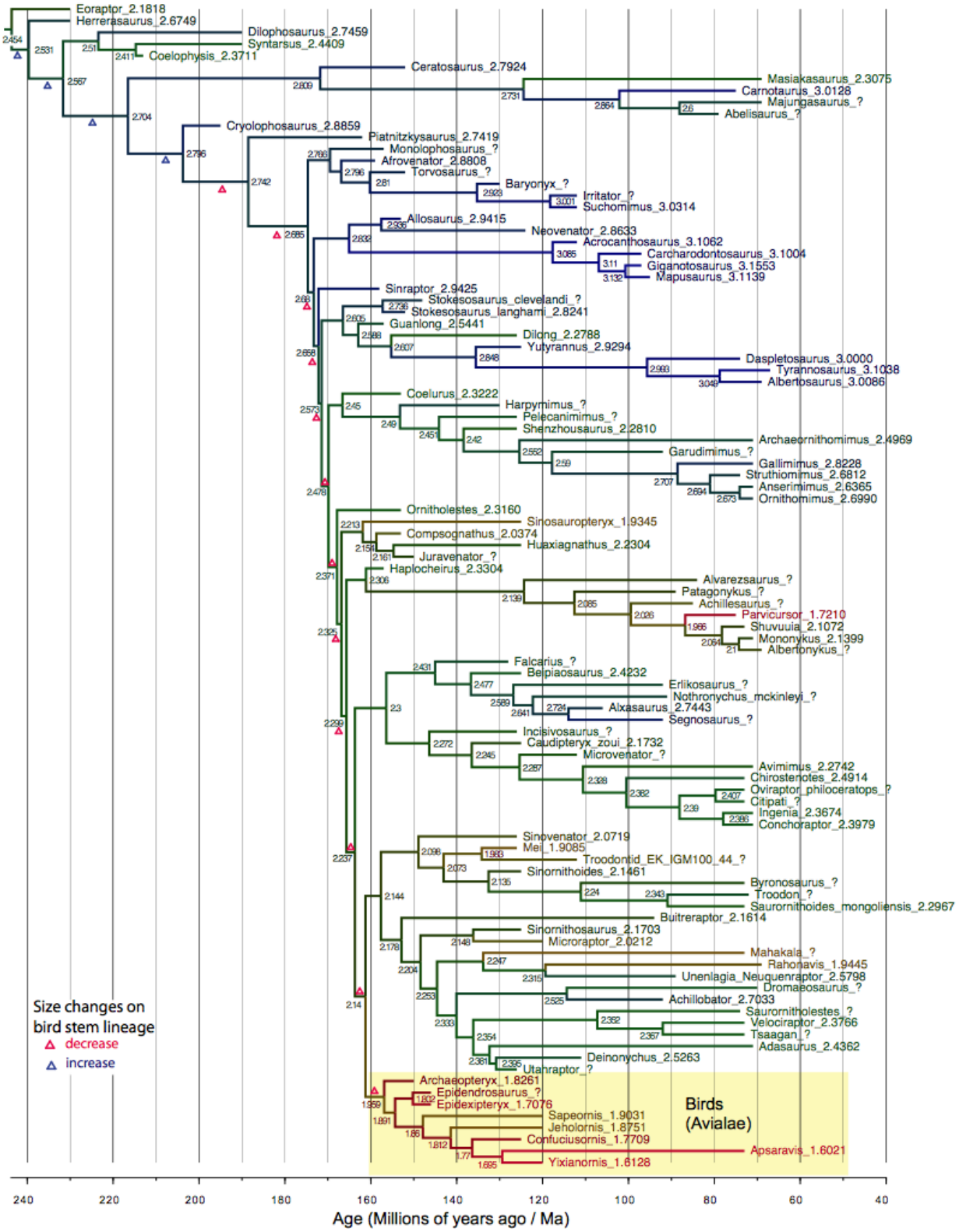

Fig. S5. **Theropod phylogeny and body size evolution, based on data set 2**. The same trend of continuous, unreversed size reduction along most of the bird stem lineage is found as in data set 1. Bayesian MCC tree with branches colour-coded according to body size as indexed by $\log_{10}$ femur length (compare with Figs. 1 and S1); numbers denote observed values at tips, or inferred ancestral values at nodes. Triangles denote size increases or decreases along the bird stem lineage (compare with Fig 2a), and size trends in this lineage are plotted in Fig. 2b.

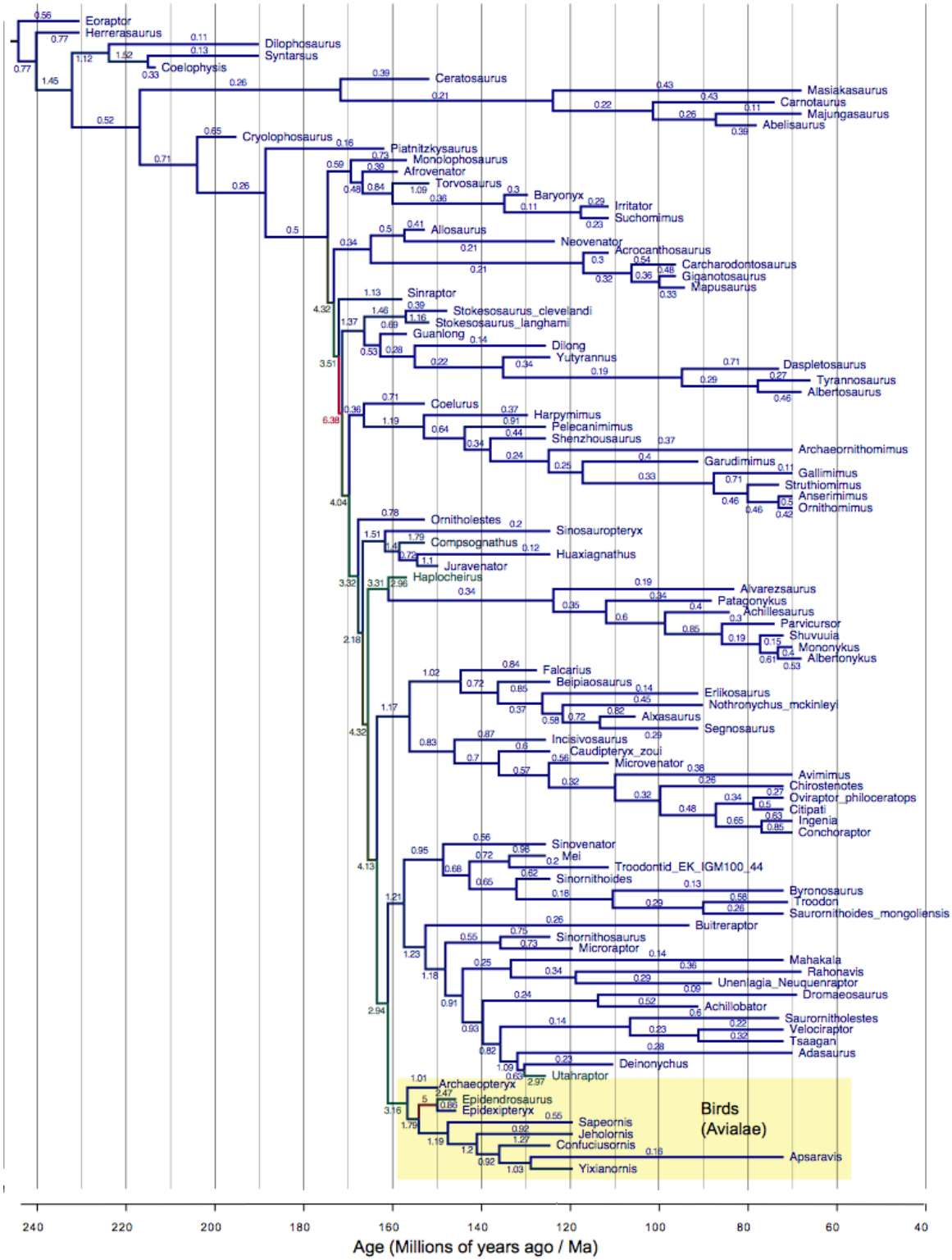

Fig. S6. **Theropod evolutionary rates, based on data set 2**. Fast rates characterise the bird stem lineage, as in data set 1 (compare with Fig. 3). Bayesian MCC tree with branches colour-coded according to inferred rate of evolution (percentage divergence per million years, across all 1549 discrete characters; a rate of 1% equates to a 0.01 probability of change per character per lineage per million years).

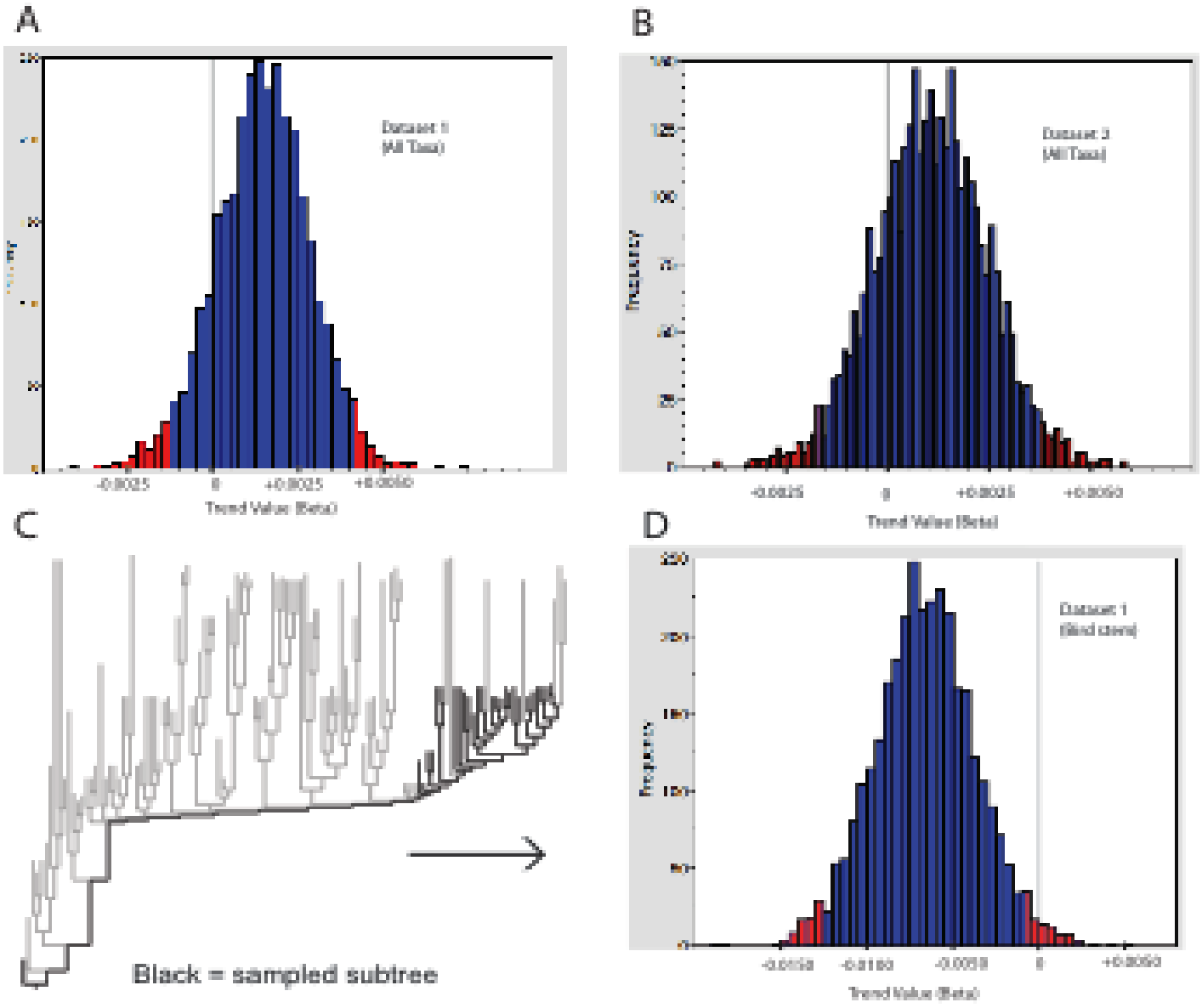

Fig. S7. **No overall trend in body size evolution across theropods, but a significant trend in the bird lineage**. The estimated trend parameter (beta) from a BayesTraits (*27*) analysis across 1000 sampled trees in (A) data set 1 and (B) data set 2. In both data sets, the mean estimate of the trend parameter is close to 0 (~0.0012) and the 95% highest posterior density (blue) encompasses 0. Adding this parameter also does not improve model fit (see Section D). (C) Pruned subtree retaining the outgroup and basal bird taxa; this is the maximum MCC consensus of the pruned subtrees (1000 sampled subtrees representing relationships among the retained taxa were used for actual analysis). (D) The estimated trend parameter (beta) from a Bayestraits analysis of the 1000 pruned subtrees; the mean estimate is ~5x times the absolute magnitude than that retrieved for the full tree (-0.0066), and the 95% highest posterior density (blue) excludes zero.

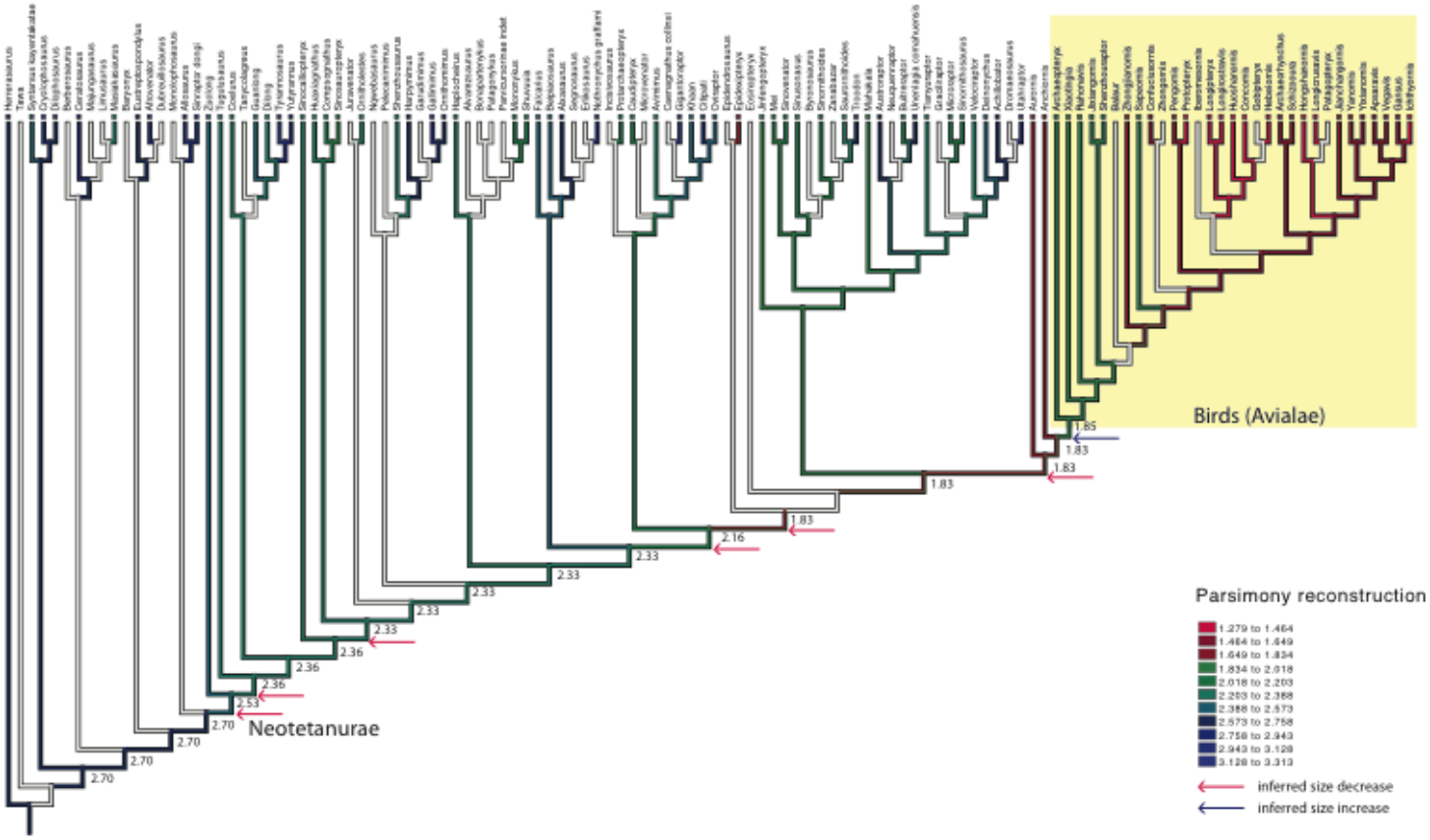

Fig. S8. **Theropod phylogeny and size evolution based on parsimony (cladistic) analysis of data set 1.** Majority-rule consensus of >100 000MPTs, with ancestral node reconstructions based on linear parsimony. Branches are colour-coded according to size, as indexed by $\log_{10}$ femur length (blue=large, green=medium, pink=small); numbers at nodes are inferred values for size along the bird stem lineage. Where there is a range of equally-parsimonious values for a node, Mesquite (*59*) by default prints the lower value (shown here). Using the mean value or upper value does not change the retrieved pattern.

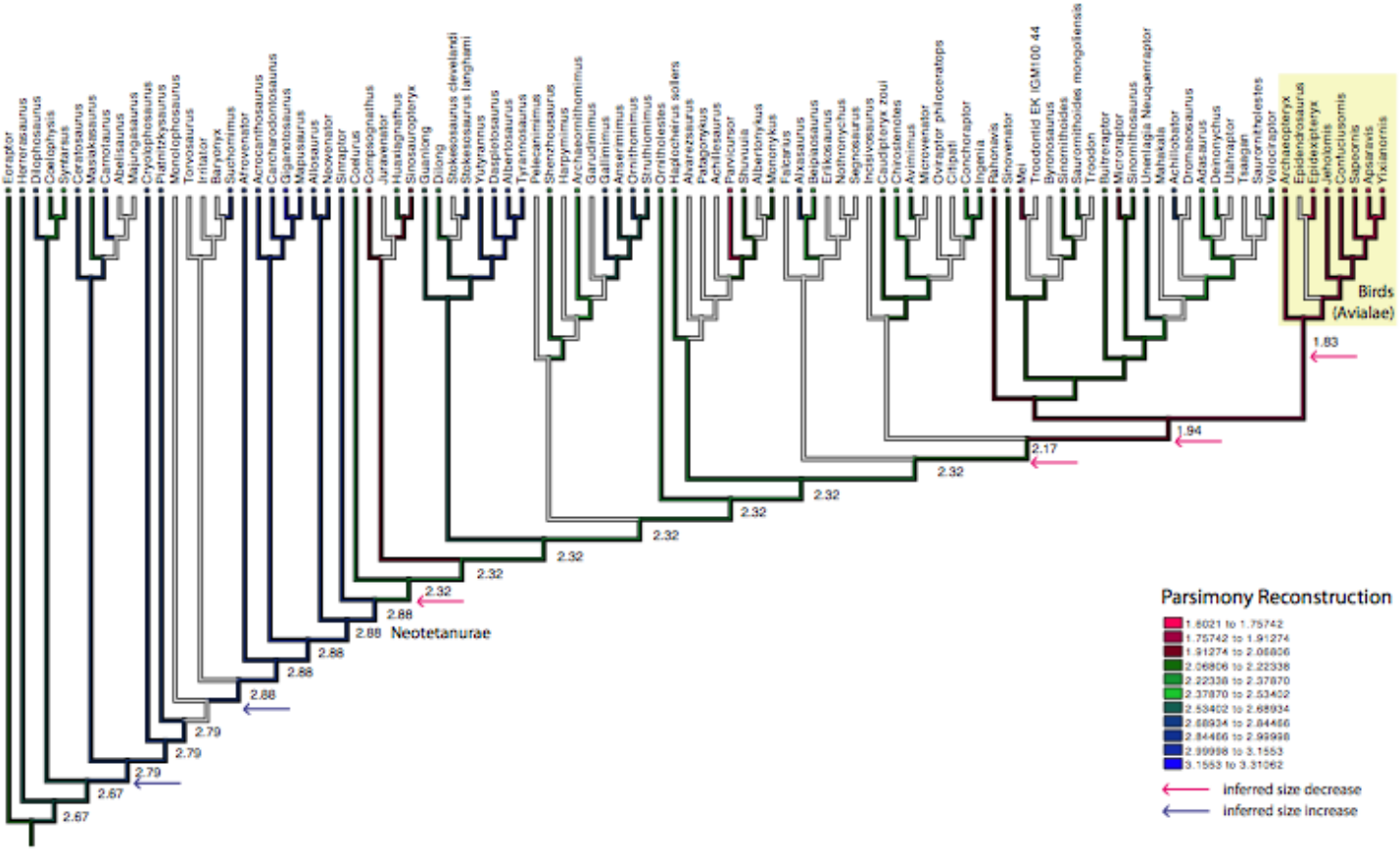

Fig. S9. **Theropod phylogeny and size evolution based on parsimony (cladistic) analysis of data set 2.** Majority-rule consensus of >100 000MPTs, with ancestral node reconstructions based on linear parsimony. Branches are colour-coded according to size, as indexed by $\log_{10}$ femur length (blue=large, green=medium, pink=small); numbers at nodes are inferred values for size along the bird stem lineage. Where there is a range of equally-parsimonious values for a node, Mesquite (*59*) by default prints the lower value (shown here). Using the mean value or upper value does not change the retrieved pattern.

References and Notes

1. J. Gauthier, L. F. Gall, Eds., *New Perspectives on the Origin and Early Evolution of Birds* (Peabody Museum of Natural History, New Haven, CT, 2001).

2. G. Dyke, G. Kaiser, Eds., *Living Dinosaurs: the Evolutionary History of Modern Birds* (Wiley, Chichester, 2011).

3. J. Gauthier, Saurischian monophyly and the origin of birds. *Mem. Calif. Acad. Sci.* **8**, 1–55 (1986).

4. P. C. Sereno, The origin and evolution of dinosaurs. *Annu. Rev. Earth Planet. Sci.* **25**, 435–489 (1997). doi:10.1146/annurev.earth.25.1.435

5. K. Padian, A. J. de Ricqlès, J. R. Horner, Dinosaurian growth rates and bird origins. *Nature* **412**, 405–408 (2001). Medline doi:10.1038/35086500

6. T. A. Dececchi, H. C. E. Larsson, Body and limb size dissociation at the origin of birds: Uncoupling allometric constraints across a macroevolutionary transition. *Evolution* **67**, 2741–2752 (2013). Medline doi:10.1111/evo.12150

7. F. E. Novas, M. D. Ezcurra, F. L. Agnolín, D. Pol, R. Ortíz, New Patagonian Cretaceous theropod sheds light about the early radiation of Coelurosauria. *Rev. Mus. Argentino de Cienc. Nat. n.s.* **14**, 57–81 (2012).

8. A. H. Turner, D. Pol, J. A. Clarke, G. M. Erickson, M. A. Norell, A basal dromaeosaurid and size evolution preceding avian flight. *Science* **317**, 1378–1381 (2007). Medline doi:10.1126/science.1144066

9. M. T. Carrano, in *Amniote Paleobiology*, M. T. Carrano, T. J. Gaudin, R. W. Blob, J. R. Wible, Eds. (Univ. of Chicago Press, Chicago, 2006), chap. 8.

10. D. W. E. Hone, T. M. Keesey, D. Pisani, A. Purvis, Macroevolutionary trends in the Dinosauria: Cope's rule. *J. Evol. Biol.* **18**, 587–595 (2005). Medline doi:10.1111/j.1420-9101.2004.00870.x

11. R. B. Sookias, R. J. Butler, R. B. J. Benson, Rise of dinosaurs reveals major body-size transitions are driven by passive processes of trait evolution. *Proc. Biol. Sci.* **279**, 2180–2187 (2012). Medline doi:10.1098/rspb.2011.2441

12. B.-A. S. Bhullar, J. Marugán-Lobón, F. Racimo, G. S. Bever, T. B. Rowe, M. A. Norell, A. Abzhanov, Birds have paedomorphic dinosaur skulls. *Nature* **487**, 223–226 (2012). Medline doi:10.1038/nature11146

13. A. M. Heers, K. P. Dial, From extant to extinct: Locomotor ontogeny and the evolution of avian flight. *Trends Ecol. Evol.* **27**, 296–305 (2012). Medline doi:10.1016/j.tree.2011.12.003

14. V. Allen, K. T. Bates, Z. Li, J. R. Hutchinson, Linking the evolution of body shape and locomotor biomechanics in bird-line archosaurs. *Nature* **497**, 104–107 (2013). Medline doi:10.1038/nature12059

15. M. N. Puttick, G. H. Thomas, M. J. Benton, High rates of evolution preceded the evolution of birds. *Evolution* 10.1111/evo.12363 (2014). doi: 10.1111/evo.12363

16. R. B. J. Benson, N. E. Campione, M. T. Carrano, P. D. Mannion, C. Sullivan, P. Upchurch, D. C. Evans, Rates of dinosaur body mass evolution indicate 170 million years of sustained ecological innovation on the avian stem lineage. *PLoS Biol.* **12**, e1001853 (2014). Medline doi:10.1371/journal.pbio.1001853

17. D. K. Zelenitsky, F. Therrien, G. M. Erickson, C. L. DeBuhr, Y. Kobayashi, D. A. Eberth, F. Hadfield, Feathered non-avian dinosaurs from North America provide insight into wing origins. *Science* **338**, 510–514 (2012). Medline doi:10.1126/science.1225376

18. X. Zheng, Z. Zhou, X. Wang, F. Zhang, X. Zhang, Y. Wang, G. Wei, S. Wang, X. Xu, Hind wings in basal birds and the evolution of leg feathers. *Science* **339**, 1309–1312 (2013). Medline doi:10.1126/science.1228753

19. R. B. J. Benson, J. N. Choiniere, Rates of dinosaur limb evolution provide evidence for exceptional radiation in Mesozoic birds. *Proc. Biol. Sci.* **280**, 20131780 (2013).doi:10.1098/rspb.2013.1780 Medline

20. Materials and methods are available as supplementary materials on *Science* Online. Data files are archived on Dryad Digital Repository (doi:10.5061/dryad.jm6pj).

21. P. Godefroit, A. Cau, H. Dong-Yu, F. Escuillié, W. Wenhao, G. Dyke, A Jurassic avialan dinosaur from China resolves the early phylogenetic history of birds. *Nature* **498**, 359–362 (2013). Medline doi:10.1038/nature12168

22. X. Xu, K. Wang, K. Zhang, Q. Ma, L. Xing, C. Sullivan, D. Hu, S. Cheng, S. Wang, A gigantic feathered dinosaur from the lower cretaceous of China. *Nature* **484**, 92–95 (2012). Medline doi:10.1038/nature10906

23. J. N. Choiniere, X. Xu, J. M. Clark, C. A. Forster, Y. Guo, F. Han, A basal alvarezsauroid theropod from the early Late Jurassic of Xinjiang, China. *Science* **327**, 571–574 (2010). Medline doi:10.1126/science.1182143

24. P. Christiansen, R. A. Farina, Mass prediction in theropod dinosaurs. *Hist. Biol.* **16**, 85 (2004). doi:10.1080/08912960412331284313

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

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

27. M. Pagel, A. Meade, *BayesTraits: Software and Documentation* (2013); www.evolution.reading.ac.uk/BayesTraitsV2Beta.html.

28. L. E. Zanno, P. J. Makovicky, No evidence for directional evolution of body mass in herbivorous theropod dinosaurs. *Proc. Biol. Sci.* **280**, 20122526 (2013). Medline doi:10.1098/rspb.2012.2526

29. J. Hanken, D. B. Wake, Miniaturization of body size: Organismal consequences and evolutionary significance. *Annu. Rev. Ecol. Syst.* **24**, 501–519 (1993). doi:10.1146/annurev.es.24.110193.002441

30. O. W. M. Rauhut, C. Foth, H. Tischlinger, M. A. Norell, Exceptionally preserved juvenile megalosauroid theropod dinosaur with filamentous integument from the Late Jurassic of Germany. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 11746–11751 (2012). Medline doi:10.1073/pnas.1203238109

31. T. S. Kemp, The concept of correlated progression as the basis of a model for the evolutionary origin of major new taxa. *Proc. Biol. Sci.* **274**, 1667–1673 (2007). Medline doi:10.1098/rspb.2007.0288

32. D. W. E. Hone, G. J. Dyke, M. Haden, M. J. Benton, Body size evolution in Mesozoic birds. *J. Evol. Biol.* **21**, 618–624 (2008). Medline doi:10.1111/j.1420-9101.2007.01483.x

33. A. Ekzanowski, in *Mesozoic Birds: Above the Heads of Dinosaurs*, L. Chiappe and L. Witmer, Eds. (Univ. of California Press, Berkeley, 2002), pp. 129–159.

34. P. Wellnhofer, *Archaeopteryx: der Urvogel von Solnhofen* (Verlag Dr. Friedrich Pfeil, Munich, 2008).

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

36. P. O. Lewis, A likelihood approach to estimating phylogeny from discrete morphological character data. *Syst. Biol.* **50**, 913–925 (2001). Medline doi:10.1080/106351501753462876

37. P. Lemey, A. Rambaut, J. J. Welch, M. A. Suchard, Phylogeography takes a relaxed random walk in continuous space and time. *Mol. Biol. Evol.* **27**, 1877–1885 (2010). Medline doi:10.1093/molbev/msq067

38. M. S. Y. Lee, A. Cau, D. Naish, G. J. Dyke, Morphological clocks in paleontology, and a mid-Cretaceous origin of crown Aves. *Syst. Biol.* **63**, 442–449 (2014). Medline doi:10.1093/sysbio/syt110

39. G. Hunt, M. T. Carrano, in *Short Course on Quantitative Methods in Paleobiology*, J. Alroy, G. Hunt, Eds. (Paleontological Society, Denver, CO, 2010), pp. 245–269.

40. M. S. Y. Lee, J. Soubrier, G. D. Edgecombe, Rates of phenotypic and genomic evolution during the Cambrian explosion. *Curr. Biol.* **23**, 1889–1895 (2013). Medline doi:10.1016/j.cub.2013.07.055

41. R. E. Kass, A. E. Raftery, Bayes factors. *J. Am. Stat. Assoc.* **90**, 773–795 (1995). doi:10.1080/01621459.1995.10476572

42. A. Rambaut, A. J. Drummond, *Tracer: MCMC Trace Analysis Tool*, v1.5. http://beast.bio.ed.ac.uk/ (2009).

43. M. A. Suchard, R. E. Weiss, J. S. Sinsheimer, Bayesian selection of continuous-time Markov chain evolutionary models. *Mol. Biol. Evol.* **18**, 1001–1013 (2001). Medline doi:10.1093/oxfordjournals.molbev.a003872

44. T. R. Holtz Jr., The phylogenetic position of the Tyrannosauridae: Implications for theropod systematics. *J. Palaeontol.* **68**, 1100–1117 (1994).

45. O. W. M. Rauhut, A tyrannosauroid dinosaur from the Upper Jurassic of Portugal. *Palaeontology* **46**, 903–910 (2003). doi:10.1111/1475-4983.00325

46. S. L. Brusatte, M. A. Norell, T. D. Carr, G. M. Erickson, J. R. Hutchinson, A. M. Balanoff, G. S. Bever, J. N. Choiniere, P. J. Makovicky, X. Xu, Tyrannosaur paleobiology: New research on ancient exemplar organisms. *Science* **329**, 1481–1485 (2010). Medline doi:10.1126/science.1193304

47. S. L. Brusatte, S. J. Nesbitt, R. B. Irmis, R. J. Butler, M. J. Benton, M. A. Norell, The origin and early radiation of dinosaurs. *Earth Sci. Rev.* **101**, 68–100 (2010). doi:10.1016/j.earscirev.2010.04.001

48. M. D. Ezcurra, A new early dinosaur (Saurischia: Sauropodomorpha) from the Late Triassic of Argentina: a reassessment of dinosaur origin and phylogeny. *J. Syst. Palaeontol.* **8**, 371–425 (2010). doi:10.1080/14772019.2010.484650

49. M. C. Langer, M. D. Ezcurra, J. S. Bittencourt, F. E. Novas, The origin and early evolution of dinosaurs. *Biol. Rev. Camb. Philos. Soc.* **85**, 55–110 (2010). Medline doi:10.1111/j.1469-185X.2009.00094.x

50. S. J. Nesbitt, P. M. Barrett, S. Werning, C. A. Sidor, A. J. Charig, The oldest dinosaur? A Middle Triassic dinosauriform from Tanzania. *Biol. Lett.* **9**, 20120949 (2013). Medline doi:10.1098/rsbl.2012.0949

51. A. Mudroch, U. Richter, U. Joger, R. Kosma, O. Idé, A. Maga, Didactyl tracks of paravian theropods (Maniraptora) from the ?Middle Jurassic of Africa. *PLOS ONE* **6**, e14642 (2011). 10.1371/journal.pone.0014642 Medline doi:10.1371/journal.pone.0014642

52. M. T. Carrano, R. B. J. Benson, S. D. Sampson, The phylogeny of Tetanurae (Dinosauria: Theropoda). *J. Syst. Palaeontol.* **10**, 211–300 (2012). doi:10.1080/14772019.2011.630927

53. O. W. M. Rauhut, A. C. Milner, S. Moore-Fay, Cranial osteology and phylogenetic position of the theropod dinosaur *Proceratosaurus bradleyi* (Woodward, 1910) from the Middle Jurassic of England. *Zool. J. Linn. Soc.* **158**, 155–195 (2010). doi:10.1111/j.1096-3642.2009.00591.x

54. R. B. J. Benson, The osteology of *Magnosaurus nethercombensis* (Dinosauria, Theropoda) from the Bajocian (Middle Jurassic) of the United Kingdom and a re-examination of the oldest records of tetanurans. *J. Syst. Palaeontol.* **8**, 131–146 (2010). doi:10.1080/14772011003603515

55. P. M. Barrett, The affinities of the enigmatic dinosaur *Eshanosaurus deguchiianus* from the Early Jurassic of Yunnan Province, People's Republic of China. *Palaeontology* **52**, 681–688 (2009). doi:10.1111/j.1475-4983.2009.00887.x

56. J. C. Wilgenbusch, D. L. Warren, D. L. Swofford, AWTY: A system for graphical exploration of MCMC convergence in Bayesian phylogenetic inference (2004); http://ceb.csit.fsu.edu/awty.

57. A. Rambaut, *Figtree: Tree Figure Drawing Tool* (Institute of Evolutionary Biology, University of Edinburgh, 2006); http://tree.bio.ed.ac.uk/software/figtree/.

58. M. Laurin, Assessment of the relative merits of a few methods to detect evolutionary trends. *Syst. Biol.* **59**, 689–704 (2010). Medline doi:10.1093/sysbio/syq059

59. W. P. Maddison, D. R. Maddison, *Mesquite: A Modular System for Evolutionary Analysis*, Version 2.75 (2011); http://mesquiteproject.org/mesquite/mesquite.html.

60. D. L. Swofford, *PAUP*: Phylogenetic Analysis Using Parsimony* (*and Other Methods)*, Version 4.0. (Sinauer Associates, Sunderland, MA, 2003).

61. J. M. Carpenter, Choosing among multiple equally parsimonious cladograms. *Cladistics* **4**, 291–296 (1988). doi:10.1111/j.1096-0031.1988.tb00476.x

62. G. Han, L. Chiappe, S.-A. Ji, M. Habib, A. H. Turner, A. Chinsamy, X. Liu, L. Han, A new raptorial dinosaur with exceptionally long feathering provides insights into dromaeosaurid flight performance. *Nature Commun.* **5**, 4382 (2014). doi: 10.1038/ncomms5382