

Body size evolution in Mesozoic birds: little evidence for Cope's rule

R. J. BUTLER* & A. GOSWAMI†

*Department of Palaeontology, The Natural History Museum, London, UK

†Department of Earth Sciences, University of Cambridge, Cambridge, UK

Keywords:

Aves;
body size;
Cope's rule;
evolution;
flight;
Mesozoic;
phylogeny.

Abstract

Cope's rule, the tendency towards evolutionary increases in body size, is a long-standing macroevolutionary generalization that has the potential to provide insights into directionality in evolution; however, both the definition and identification of Cope's rule are controversial and problematic. A recent study [*J. Evol. Biol.* 21 (2008) 618] examined body size evolution in Mesozoic birds, and claimed to have identified evidence of Cope's rule occurring as a result of among-lineage species sorting. We here reassess the results of this study, and additionally carry out novel analyses testing for within-lineage patterns in body size evolution in Mesozoic birds. We demonstrate that the nonphylogenetic methods used by this previous study cannot distinguish between among- and within-lineage processes, and that statistical support for their results and conclusions is extremely weak. Our ancestor–descendant within-lineage analyses explicitly incorporate recent phylogenetic hypotheses and find little compelling evidence for Cope's rule. Cope's rule is not supported in Mesozoic birds by the available data, and body size evolution currently provides no insights into avian survivorship through the Cretaceous–Paleogene mass extinction.

Introduction

Cope's rule, the tendency towards evolutionary increases in body size, has been referred to as both a 'psychological artefact' (Gould, 1997: 199) and 'among the most pervasive patterns in the history of life' (Kingsolver & Pfennig, 2004: 1608). Although recent work has begun to provide a framework for understanding how Cope's rule might operate (e.g. Kingsolver & Pfennig, 2004; Brown & Sibly, 2006), there is little consensus as to how often, and in which taxonomic groups, Cope's rule can be identified to have occurred (see Table 1 of Moen, 2006 for a review of examples and counterexamples of Cope's rule). Moreover, the definition of Cope's rule is itself problematic: some workers appear to consider any increase in relative size within a clade to represent Cope's rule, regardless of whether it is the result of within- or among-lineage

processes (e.g. Hone *et al.*, 2005, 2008), whereas others focus primarily on within-lineage (i.e. ancestor–descendant) changes (e.g. Alroy, 1998, 2000). Finally, it has been suggested that, although Cope's rule might be genuine, it is merely an artefact of the fact that founders of clades are often small organisms (Stanley, 1973), as opposed to being the result of strong directional selection for larger body size across all size ranges.

Birds are the most speciose (around 10 000 species) major clade of extant terrestrial vertebrates, and are descended from small-bodied theropod dinosaurs (e.g. Norell & Xu, 2005). The earliest bird (Turner *et al.*, 2007), the magpie-sized *Archaeopteryx*, is known from the Late Jurassic, with the earliest unambiguous crown-group birds (Neornithes) known from the Late Cretaceous (e.g. Clarke *et al.*, 2005; Fountaine *et al.*, 2005). During the approximately 90 Myr between *Archaeopteryx* and the Cretaceous–Paleogene (KPg) mass extinction, Mesozoic birds evolved into a wide diversity of forms, including wholly extinct, stem-group clades such as Enantiornithes and Hesperornithiformes (Chiappe & Dyke, 2002; Zhou, 2004). Understanding the morphology,

Correspondence: Richard J. Butler, Department of Palaeontology, The Natural History Museum, Cromwell Road, London SW7 5BD, UK.
Tel.: +44 207 942 5582; fax: +44 207 942 5546;
e-mail: r.butler@nhm.ac.uk

inter-relationships and sequence of character acquisition amongst basal Mesozoic birds is crucial to understanding the evolution of modern birds from their nonavian dinosaur ancestors.

Recently, Hone *et al.* (2008) provided the first analysis of body size change in Mesozoic birds. Specifically, they carried out nonphylogenetic analyses that plotted body size proxies against stratigraphic age, and proposed that their analyses indicated the occurrence of 'Cope's rule *sensu stricto*' in Mesozoic birds as a result of 'reduction in variance with the apparent loss of smaller forms' (Hone *et al.*, 2008). In addition, Hone *et al.* suggested that body size decreases in the single Mesozoic bird lineage to cross the KPg boundary (Ornithuromorpha) might account for bird survivorship through the KPg mass extinction event. Identification of Cope's rule in Mesozoic birds is interesting and perhaps surprising, particularly in the light of the relatively large size of the earliest known bird (*Archaeopteryx*) and the right-skewed mass distribution of extant birds (Blackburn & Gaston, 1994). Moreover, any study that claims to have identified the operation of Cope's rule is significant, given the potential for insights into major macroevolutionary trends. Because of the potential importance of the results of Hone *et al.* (2008), we have reassessed their analyses. We find significant problems with both their nonphylogenetic methodological approach and with the characterization and interpretation of their results. Perhaps most strikingly, Hone *et al.* (2008) failed to provide measures of statistical significance for their analysis of body size evolution, requiring, at the very least, reanalysis following their exact methodology. Here, we re-examine the evidence for Cope's rule in Mesozoic birds using the body size data of Hone *et al.* (2008).

Hone *et al.* (2005) defined Cope's rule as the result of one or more of three processes: (i) necessary increase in average size when the founders of clades are small organisms; (ii) genuine within-lineage natural selection where the advantages of larger size act to produce progressively larger descendants; and (iii) among-lineage sorting of species, where larger species within a higher level clade tend to survive or proliferate in preference to smaller species. Process (i) is probably invalid for Mesozoic birds, as the earliest known members of Aves (e.g. *Archaeopteryx*) are relatively large (Hone *et al.*, 2008). Furthermore, apparent increase in average size as a result of increasing variation, rather than directional selection, is not considered Cope's rule by most workers (Jablonski, 1997). Particularly within endothermic organisms, increasing average size, without increasing minimum size, may well reflect increased variability (difference between minimal and maximal sizes in a clade) coupled with physiological barriers or other hindrances to decreasing size, and it is important to distinguish this kind of passive trend from the active trend generally understood as Cope's rule (McShea, 1994).

As discussed by Alroy (2000: 320), the presence of nonrandom within-lineage trends (process ii) can only be tested adequately using ancestor–descendant comparisons: such comparisons can identify nonrandom evolution within lineages, but cannot quantify the relative importance of within- vs. among-lineage trends. Ancestor–descendent comparisons were not carried out by Hone *et al.* (2008); those authors suggested that our understanding of Mesozoic bird taxonomy is insufficient to allow such ancestor–descendent comparisons at present. We here make the first attempt to use such body size comparisons for Mesozoic birds. The claim of Hone *et al.* (2008) that they had identified Cope's rule in Mesozoic birds occurring due to 'the apparent loss of smaller forms' suggests that they believed they had identified the operation of process (iii) with larger species surviving/proliferating at the expense of smaller species; however, the nonphylogenetic approach utilized by these authors cannot distinguish between within- and among-lineage processes (Alroy, 2000). Nevertheless, we reanalyse their data using the same nonphylogenetic approaches: if ancestor–descendent comparisons find no evidence for nonrandom evolution within lineages but body size trends are recognized by nonphylogenetic approaches, then it is a plausible hypothesis that those latter trends are the result of among-lineage sorting.

Methods

Reanalysis of the data of Hone *et al.* (2008)

Hone *et al.* (2008; hereafter HEA) used a series of least-squares regressions of log femur length against stratigraphic age to provide evidence for body size trends in Mesozoic birds. There are significant problems with this approach to trend analysis; most importantly, this approach increases type I error rate (a fact acknowledged by HEA; Martins *et al.*, 2002) and it cannot resolve whether any identified increase in body size is a result of selection within or among lineages (see in particular Alroy, 2000). Ancestor–descendent comparisons (see below) are capable of identifying the presence of within-lineage trends (see above), although they cannot quantify the relative importance of within- versus among-lineage trends (Alroy, 2000). However, similar methods to those used by HEA have been used in many studies of macroevolutionary trends and, as discussed above, may be informative when paired with ancestor–descendent comparisons. HEA did not report significance values for any of their regressions, but did suggest that they were 'significant'. However, Shapiro–Wilk's test demonstrates that the data presented by HEA are not normally distributed ($W = 0.841$, $P = 0.00$) and, therefore, the parametric methods they used cannot assess statistical significance (they can still be used to compute correlation coefficient). Here, we present the significance

values for their original analyses and additionally reanalyse the data with more appropriate nonparametric tests.

It is important to note in the caption to Table S1 of HEA that the authors excluded 32 of the 117 specimens included in their full data set for lack of humeral or femoral data. Additionally, HEA excluded all but the largest specimens for each species, resulting in a final data set of 43 specimens, with each species represented by a single specimen, the largest one measured. Other authors have noted (in extant vertebrate populations) that maxima may not be good indicators of body size within populations, and that selection acts on adult body size through all reproductively active individuals, not just the largest fraction (e.g. Meiri, 2007). Moreover, using only the largest fossil individual potentially results in bias: a species represented by a large number of specimens will generally appear larger than a species that in reality has the same average size but is represented in the fossil record by fewer specimens – as a result some other studies of body size evolution in fossil taxa have chosen to use average sizes rather than maxima (M. Laurin, personal communication). For these reasons, we therefore reran the analyses of HEA with the same data selection (largest specimens only), as well as two additional analyses that use: (1) all of the specimen data (with missing femoral data extrapolated from humeral data according to the equation provided by HEA); and (2) the mean femoral length (FL) for each species. We present the significance values for the Pearson's product-moment correlation coefficient used by HEA and the 95% confidence intervals for the slopes of the least-squares regression lines that HEA presented as evidence of the significant positive relationship between stratigraphic age and body size in Mesozoic birds. Because the residuals of the least-squares regression were not normally distributed (Shapiro–Wilk's test for all Mesozoic Aves, $W = 0.972$, $P = 0.037$), bootstrapped 95% confidence intervals (2000 replicates) were used instead of SE. In addition, the nonparametric Spearman's rank correlation analysis (Table 1) was used to determine the relationship between stratigraphic age and log FL for the same five clades (Aves, Pygostylia, Ornithothoraces, Enantiornithes, and Ornithuromorpha) examined in

HEA. Analyses were carried out in PAST (Hammer *et al.*, 2001).

Lastly, to test the claim by HEA that Cope's rule acted in Mesozoic birds through a reduction in variance coupled with a loss of smaller forms, we calculated variance in FL (mean values for species were used) in four equal time bins of minimum age and maximum age (to account for large variation in stratigraphic ranges). The time bins used were: (a) 148–128; (b) 128–108; (c) 108–88; and (d) 88–68 Ma. It should be noted that the oldest time bin is poorly sampled in both the minimum age and maximum age analyses (two and three species respectively).

Testing for within-lineage directional changes in body size using phylogenetic data

We combined the results of two recent large-scale phylogenetic studies (Clarke *et al.*, 2006; Chiappe *et al.*, 2007) to produce an informal supertree of Mesozoic birds that contains 27 terminal taxa. These phylogenetic studies differ in their respective taxonomic scope: Clarke *et al.* (2006) focused on broad-scale patterns of early avian phylogeny, whereas Chiappe *et al.* (2007) focused on relationships within the clade Enantiornithes. The results of these studies complement one another, and can be combined into a single tree without conflict. Moreover, the topology of the supertree is largely in agreement with other recent analyses of early avian phylogeny (e.g. You *et al.*, 2006), differing mainly in the choice of operational taxonomic units. This maximally taxonomically inclusive supertree was not used directly in the body size analyses because it contains multiple taxa for which FL, the body size proxy utilized by HEA, is not available, as well as secondarily flightless taxa (excluded by HEA: see below). As a result, we pruned this supertree to include 19 volant taxa for which FL measurements are available, creating a pruned topology that we refer to as 'supertree A' [the general topology of this supertree is shown in Fig. 1; however, note that the figure actually shows the slightly more resolved 'supertree B' (see below) and thus contains only 18 taxa]. The taxonomic sampling present in supertree A is considerably smaller than that used by HEA, but allows us to incorporate detailed phylogenetic data and explicitly test for within-lineage evidence of Cope's rule. The majority of the FL measurements were taken from the data set of HEA (see their Table S1); however, we also included measurements for two taxa absent in the data set of HEA: *Eoenantiornis* (FL = 26.5 mm; Hou *et al.*, 1999) and *Longirostravis* (FL = 20 mm; Hou *et al.*, 2004). Although not explicitly discussed by HEA, those authors excluded secondarily flightless taxa (e.g. *Patagopteryx*, *Hesperornis*, *Baptornis* and *Elsornis*) because the loss of flight would potentially expose those taxa to different selection pressures to those experienced by volant forms (Hone and Benton, 2007); likewise, we pruned secondarily

Table 1 Results of Spearman's rank correlation analysis using species means (see text), and the largest specimens only (method employed by Hone *et al.*, 2008).

	Species means (<i>P</i> -value)	Largest only (<i>P</i> -value)
Aves	-0.3079 (0.0446*)	-0.224 (0.1487)
Pygostylia	-0.3044 (0.0711)	-0.2451 (0.1438)
Ornithothoraces	-0.3433 (0.0505*)	-0.3721 (0.0329*)
Enantiornithes	-0.2809 (0.1644)	-0.3294 (0.100)
Ornithuromorpha	0.6736 (0.971)	0.4865 (0.2682)

0.01 < **P* > 0.05. No results are significant at the *P* < 0.01 level.

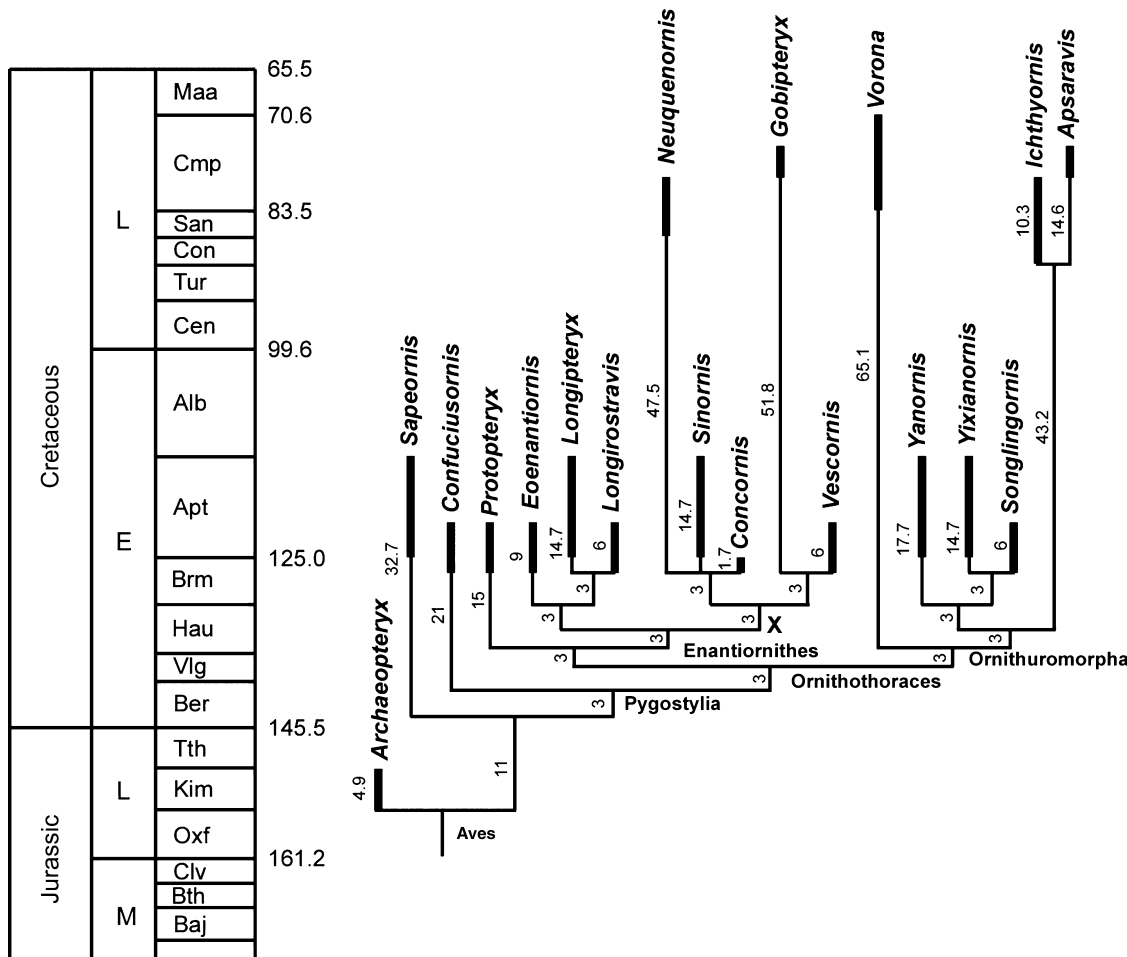


Fig. 1 ‘Supertree B’ of Mesozoic birds demonstrating inter-relationships, the five major clades (Aves, Pygostylia, Ornithothoraces, Enantiornithes, and Ornithuromorpha) used in the analyses, and inferred branch lengths. Note that the majority (~60%) of taxa included in the supertree are from just two formations: the Yixian Formation (late Barremian–early Aptian) and the Jiufotang Formation (Aptian) of China. The same bias is present in the nonphylogenetic data set of Hone *et al.* (2008), see Fig. 2. ‘Supertree A’ (19 taxa) differs in that there is no resolution of relationships within the clade labelled ‘X’ (which also includes the taxon *Eoalulavis* from the Yixian Formation).

flightless birds as part of the process of compiling our supertree A.

One problem with supertree A is that it contains a major polytomy (six taxa) within Enantiornithes, which could potentially affect our results. However, Chiappe *et al.* (2007) noted that this polytomy was the result of instability of two taxa (*Enantiornis* and *Eoalulavis*), whereas the groupings of *Gobipteryx* and *Vescornis* on the one hand, and *Sinornis*, *Neuquenornis* and *Concornis* on the other, were supported by all three of their most parsimonious trees. We therefore generated a reduced consensus tree (Wilkinson, 1994, 1995) for the data of Chiappe *et al.* (2007) that excluded the unstable taxon *Eoalulavis* (*Enantiornis* had already been excluded due to lack of data on FL) and combined this with the phylogeny of Clarke *et al.* (2006) to create ‘supertree B’ (Fig. 1) with increased resolution (the remaining polytomy

contains only three taxa), but lower taxonomic sampling ($n = 18$). Using supertree B, we carried out a second set of analyses of body size change, following the full set of procedures discussed below. Results of all within-lineage analyses are presented in Tables 2–3 and Mesquite datafiles are presented as Supporting Information (S1 and S2).

Our methodologies for reconstructing body size change follow those of Laurin (2004) and Carrano (2005) – see also Finarelli & Flynn (2006). Supertree A (19 taxa) and supertree B (18 taxa) were both fitted to the stratigraphic record, and branch lengths representing millions of years were assigned to each, with a minimal length of 3 Myr assigned to internal branches (Laurin, 2004; Fig. 1). Taxon ranges were obtained from Padian (2004), and calculations assume that the body size data for the taxon represent its size at the youngest end of the stratigraphic

Table 2 Results of ancestor–descendent pairwise comparisons for Mesozoic birds using the phylogenetic data in supertree A (19 terminal taxa).

Clade	Mean	Sum	Skew	Median	<i>n</i>	Positive changes	Negative changes	χ^2	<i>P</i>
Aves	0.029	0.9363	0.782	−0.036	32	12	20	2	0.157
Pygostylia	0.029	0.8319	0.781	−0.036	28	10	18	2.286	0.131
Ornithothoraces	0.031	0.8148	0.744	−0.036	26	9	17	2.462	0.117
Enantiornithes	0.022	0.3133	1.589	−0.044	15	4	11	3.267	0.071
Ornithuromorpha	0.056	0.5571	0.039	0.028	10	5	5	0	1

n, number of ancestor–descendent pairwise comparisons. *n* is higher than the number of taxa in supertree A because our ancestor–descendent test comparisons include changes occurring between internal nodes as well as changes occurring between internal nodes and terminal taxa: Mean: mean difference in log femoral length between ancestor and descendent pairs; sum, skew and median: sum, skew and median of all such differences; Positive changes: number of positive changes between ancestor and descendent pairs; negative changes: number of negative changes between ancestor and descendent pairs; *P*, significant value for the chi-squared test. No results are significant at the 95% level. The chi-square test determines whether the number of ancestor–descendant negative/positive body size changes differs from a null hypothesis in which body size increase and body size decrease are equally likely.

Table 3 Results of ancestor–descendent pairwise comparisons for Mesozoic birds using the phylogenetic data in supertree B (18 terminal taxa) (for definitions see Table 2).

Clade	Mean	Sum	Skew	Median	<i>n</i>	Positive changes	Negative changes	χ^2	<i>P</i>
Aves	0.0285	0.9401	0.8102	−0.0255	33	12	21	2.4545	0.164
Pygostylia	0.0288	0.8354	0.8143	−0.0255	29	10	19	2.7931	0.138
Ornithothoraces	0.0303	0.8174	0.8169	−0.0281	27	9	18	3	0.123
Enantiornithes	0.0208	0.3114	1.7017	−0.0255	15	4	11	3.2667	0.121
Ornithuromorpha	0.0562	0.5621	0.0427	0.0212	10	5	5	0	1

None of the chi-squared results are significant at the 95% level.

range to which it is assigned (Laurin, 2004). Following HEA for comparability of results, we used the largest known FL for each taxon in ancestral state reconstruction [results of analyses using species means (see above) do not differ significantly from those using the largest known individual of each species; RJB, unpublished data]. For both analyses, we used Mesquite 2.01 (Maddison & Maddison, 2007) to reconstruct ancestral states for nodes in each supertree using weighted squared-change parsimony (SCP). SCP minimizes the sum of squared change along all branches of the tree, weighting branches by their length. Unresolved nodes were treated as hard polytomies, as required of this option.

Using reconstructed ancestral values, we then analysed ancestor–descendant changes (i.e. changes between internal nodes, and between internal nodes and terminal taxa) for each of the clades (see above) considered by HEA, in each case assessing whether the mean change, median change, sum change and total number of changes were positive or negative. Nonparametric chi-squared goodness-of-fit tests were used to determine whether the number of ancestor–descendant negative/positive body size changes differed from the null hypothesis in which body size increase and body size decrease are equally likely.

This approach has similarities to independent contrast (FIC) analysis (Felsenstein, 1985) of size vs. time (as performed by Laurin, 2004), but it differs in three

important respects (Laurin, personal communication): first, by using squared-change parsimony to compute nodal values, which means that sister taxa and more basal nodes influence ancestral values (in FIC, only descendants are used in the calculation of nodal values); second, by not regressing change in size against difference in geological age; third, by not standardizing differences. Thus, the statistical properties of this method are less well known than those of FIC. The test used here is not parametric; so, it may have lower power than standard FIC.

We additionally used Mesquite 2.01 to assess whether or not a phylogenetic signal is present in the body size character, following, in part, the methodology of Laurin (2004). As noted by Laurin (2004), the value of optimizations of characters that show no phylogenetic signal is dubious. We carried out a simulation, generating 10 000 trees in which the position of terminal taxa on the tree was randomly permuted while the topology of the tree (and branch lengths, if specified) was held constant. We then calculated the squared length of the body size character across each supertree and all of the simulated trees, and compared the former with the latter – if the squared length of the body size character is less on the supertree than in at least 95% of the randomly generated trees then it is possible to conclude that the evolution of this character is associated with this tree (i.e. that there is a phylogenetic signal in this character) (Laurin, 2004).

To assess the quality of the Mesozoic bird fossil and the fit of the phylogeny to stratigraphy, we assessed the correlation between stratigraphic age of taxa (derived from data in HEA) and patristic distance (PD) for taxa included within supertree A using Spearman's rank correlation. PD represents the number of branching points (nodes) passed from the root of the cladogram to the relevant terminal taxon. This analysis was carried out in PAST, using a permutation test with 1000 random replicates.

Results

Reanalysis of data of Hone *et al.* (2008) with original methods

Hone *et al.* (2008) interpreted the results of their analysis of 43 specimens, across all Mesozoic birds, as showing a significant positive slope. However, they reported a regression line of $y = -0.0028x + 1.8525$ with an $r^2 = 0.0558$ (note that a negative slope using this approach translates into a positive relationship between time and body size – as age decreases, body size increases – and vice versa, because age is measured in millions of years before present day; this is true for all of the following nonphylogenetic analyses). An r^2 value of 0.0558 is not significant ($P = 0.1271$) and essentially means that the x -axis (time) explained only 5.58% of the variance in the y -axis (log femur length). Furthermore, the 95% confidence intervals on the slope of the regression line ranged from -0.0071 to 0.0012 ; that is, they included zero. Therefore, these results do not

support a significant positive relationship between age and body size across all Mesozoic birds. This lack of relationship was supported by the nonparametric Spearman's rank correlation analysis (Table 1; $r = -0.2240$, $P = 0.1487$).

The analysis using the full data set of 99 specimens (Fig. 2) for which temporal and femoral or humeral data were provided in HEA produced a weak but significant correlation (Pearson's $r = -0.2035$, $P = 0.0444$). However, the slope of the regression line ($y = -0.0018x + 1.8206$) again was not significantly different from zero (95% CI -0.0037 to 0.0001). Correspondingly, nonparametric tests found no significant relationship (Spearman's $r = -0.0159$, $P = 0.1189$). A third analysis across all Mesozoic birds used the mean, rather than the largest, body sizes for the seven species with multiple specimens. In this analysis ($n = 43$), there was a significant positive relationship (size increase; $y = -0.0038x + 1.9638$, Pearson's $r = -0.3334$, $P = 0.0289$) between age and body size. This relationship was upheld by the nonparametric tests for correlation (Table 1; Spearman's $r = -0.3079$, $P = 0.0446$) and slope (bootstrapped 95% CI -0.0077 to -0.0001). However, the r^2 value of 0.1111, while significant, means that stratigraphic age still explained very little of the variation in body size seen across Mesozoic birds.

When each of the four subclasses was considered individually, similar results were found. Because only the data set using species means returned a significant result in the analysis of all Mesozoic birds, we ran each analysis with species means for taxa with multiple specimens available, as well as with the largest-only data

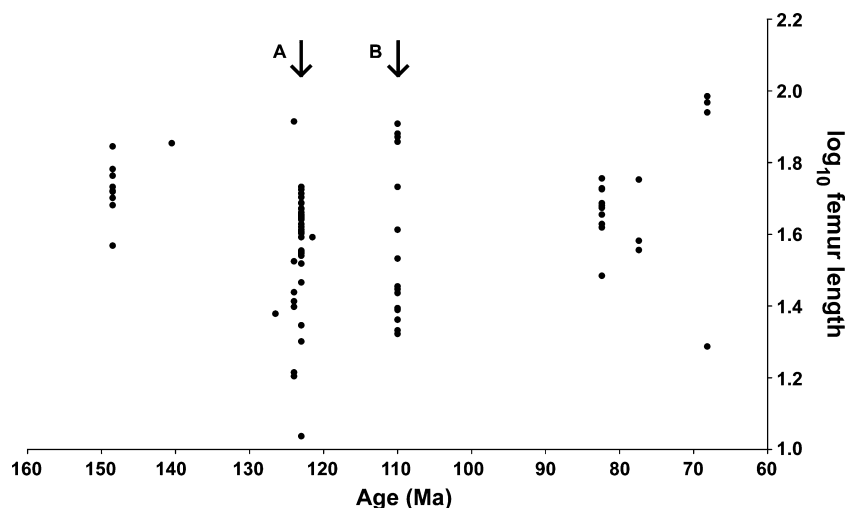


Fig. 2 Scatter plot comparing \log_{10} FL (y -axis, proxy for body mass) of Mesozoic birds against age (x -axis, Ma), for all specimens included in the data set of Hone *et al.* (2008). Timescale is from the Late Jurassic (160 Ma) to the end of the Cretaceous (65 Ma). Following Hone *et al.* (2008), the time of occurrence of each bird species is taken as the mid-point of its total stratigraphic range. Note that the earliest (Late Jurassic) birds are of relatively large size. Data points from two lagerstätten that have provided the majority of recent discoveries of Mesozoic Aves are arrowed, the Yixian (A) Jiufotang (B) formations of China. Equivalent lagerstätten are not known from the Late Cretaceous, which may partially account for the scarcity of small birds from this time period in the Hone *et al.* data set.

set used by HEA. In the original analyses of the largest-only data set, Enantiornithes showed a significant correlation ($n = 26$, $y = -0.0053x + 2.0541$, Pearson's $r = -0.3999$, $P = 0.0430$). Nonparametric tests failed to support the statistical significance of this correlation (Table 1; Spearman's $r = -0.3294$, $P = 0.1000$) or slope (bootstrapped 95% CI -0.0105 to 0.0049). Similarly, analyses using species means for Enantiornithes found no significant correlation between age and body size (Table 1; Spearman's $r = -0.2809$, $P = 0.1644$).

Within the poorly sampled Ornithuromorpha ($n = 7$), neither the correlation ($r = 0.5097$, $P = 0.2430$) nor the slope ($y = 0.0056x + 1.1233$) presented by HEA were significant (Table 1; Spearman's $r = 0.4865$, $P = 0.2682$; bootstrapped 95% CI of slope = -0.0027 to 0.03344). Use of species means, instead of largest-only data, did not change results (Table 1; Spearman's $r = 0.6736$, $P = 0.971$). Although HEA claimed that this clade showed the only negative relationship between body size and age (i.e. younger birds are smaller), the slope was not significantly different from zero.

Across Ornithothoraces (which includes Ornithuromorpha and Enantiornithes), the original analysis showed a weak, but significant, relationship between age and body size ($y = 0.0041x + 1.9342$; Pearson's $r = -0.3708$, $P = 0.0336$). Nonparametric tests support this correlation (Table 1; Spearman's $r = -0.3721$, $P = 0.0329$), but not the slope (bootstrapped 95% CI = -0.0088 to 0.0006). Analysis of species means returned a marginally significant correlation (Table 1; Spearman's $r = -0.3433$, $P = 0.0505$).

Finally, Pygostylia did not show a significant correlation in the original analysis ($y = -0.0041x + 1.9342$, Pearson's $r = -0.3708$, $P = 0.0733$) or in the reanalysis with nonparametric methods (Table 1; Spearman's $r = -0.2451$, $P = 0.1438$; bootstrapped 95% confidence intervals for the slope -0.0075 to 0.0015). Use of species means does not change this result (Table 1; Spearman's $r = -0.3044$, $P = 0.0711$).

Our analyses of variance in four time bins for minimum and maximum age also failed to support some of the conclusions of HEA. From the oldest to the youngest, the variances for the four time bins based on minimum age were: (a) 5.4; (b) 355.3; (c) 381.3; and (d) 622.6. Because of large variation in stratigraphic ranges, the results using maximum age differed: (a) 892.0; (b) 371.5; (c) 66.4; and (d) 962.5. Differences in magnitude and sampling in each time bin occur between the two analyses; however, neither shows that variance decreased markedly with time in Mesozoic birds. The analysis based upon minimum ages instead suggests an increase in variance over time.

Reanalyses incorporating phylogenetic data

A significant phylogenetic signal is present in the body size data ($P = 0.039$) using supertree A; however, a significant phylogenetic signal is absent ($P = 0.29$) when

supertree B is considered. This reduced phylogenetic signal may result from the partial resolution of the major polytomy present in the supertree A or alternatively from the decreased taxonomic sampling present in supertree B; in any case it suggests that body size evolution may not necessarily be closely associated with this phylogeny. The value of the ancestral state reconstructions may therefore be questionable (at least for the analyses utilizing the reduced supertree), and our interpretations of the results of the ancestral state reconstruction analyses (below) are necessarily conservative.

Results of ancestor–descendent comparisons are generally similar for both supertrees. Median body size changes are negative and the number of negative ancestor–descendent changes reconstructed using SCP exceeds the number of positive changes in all clades except Ornithuromorpha (negative and positive changes are equal for this clade and median body size changes are weakly positive). However, the sum and mean body size changes are weakly positive in all cases. The reason for this apparent contradiction (median body size changes are negative, mean body size changes are positive) is apparent when ancestral state reconstructions are plotted against changes occurring across branches (Fig. 3; Alroy, 2000). The majority of ancestor–descendent changes are negative, but there are a small number of large positive body size changes (commonly associated with long branch lengths): these large positive body size changes occur along branches that connect reconstructed nodal values to large-bodied terminal taxa. We tested whether the number of positive ancestor–descendent changes is significantly greater than the number of negative ancestor–descendent changes; as demonstrated by chi-squared goodness-of-fit tests, our results do not differ significantly from the null hypothesis that body size increase and body size decrease are equally likely.

There is no correlation ($P = 0.670$ – 0.972) between PD and stratigraphic age of taxa. This suggests that major biases or gaps exist in either the bird fossil record or the taxa selected for inclusion in the phylogeny, and that stratigraphically late taxa are not necessarily deeply nested: for example, *Vorona* is one of the latest occurring birds in the data set, but is phylogenetically basal, representing the most basal member of Ornithuromorpha included in our analysis, under the phylogenetic definition favoured by Hone *et al.* (2008).

Discussion

As demonstrated above, HEA inappropriately used parametric tests despite the non-normality of their data. Even with use of parametric methods, only two of the five results (Enantiornithes and Ornithothoraces) reported as significant by HEA were actually significant at the 95%

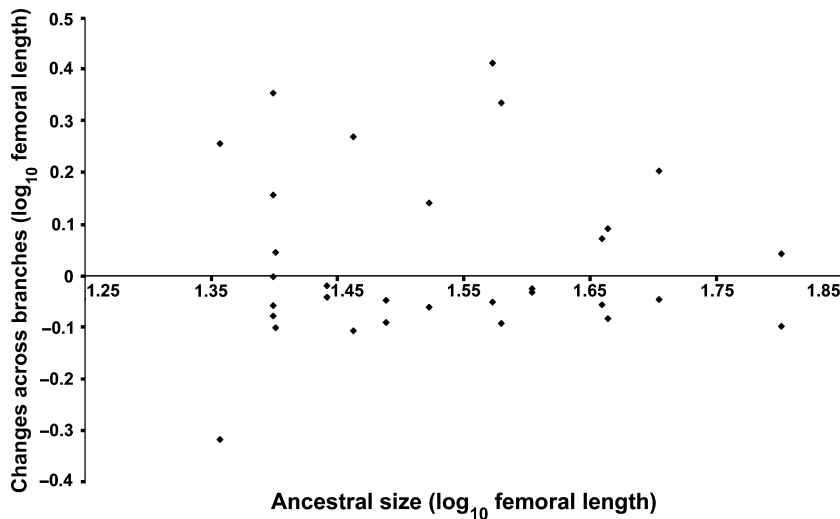


Fig. 3 Scatter plot comparing ancestral size at nodes in supertree A (inferred using SCP) in terms of \log_{10} FL (y -axis, proxy for body mass) against inferred changes occurring across branches (i.e. changes between ancestors and descendants: changes between internal nodes and between internal nodes and terminal taxa). The equivalent plot for supertree B (not shown) is extremely similar. Note that, overall, there are more negative changes than positive (see also Table 2), but that there are a number of large positive changes. For Cope's rule to be valid, the number of positive ancestor–descendent changes should be significantly (in statistical terms) higher than the number of negative changes.

level. When more suitable nonparametric tests for correlation between stratigraphic age and body size were used, only one clade, Ornithothoraces, showed a significant correlation. However, it is worth stressing that neither of the constituent clades of Ornithothoraces, Enantiornithes and Ornithuromorpha, displayed a significant relationship between age and body size using nonparametric tests. Similarly, the conclusion by HEA that Pygostylia showed strong positive trends, and that Ornithuromorpha exhibited a trend toward decreasing size, is not supported by the HEA data or by the current analysis. In summary, the analyses of HEA provide a mixed picture of body size evolution amongst Mesozoic birds: when appropriate statistical techniques are utilized most analyses find no evidence for significant trends in body size evolution, but there is some weak statistical support for a trend towards enlarged body size within Ornithothoraces. Overall, the results of HEA provide little convincing evidence for the operation of Cope's rule.

Our reanalysis of the HEA data using species means and their methods did return marginally significant correlations between age and body size across all Mesozoic Aves, and, again, within Ornithothoraces. As stated above, mean size for a species has been suggested to be a better measure for studying body size trends, and HEA provided no justification for their use of largest body size. In fact, analysis of species means provided stronger support for their interpretation of body size evolution (i.e. Cope's rule) in Mesozoic birds than did maximum body sizes, although significant results only occur in two of the five analyses, and are only weakly supported. Again, an overall pattern of increasing body size consistent with Cope's rule is absent. Moreover, and perhaps more importantly, there are a number of additional problems with the nonphylogenetic approach utilized by HEA.

First, although the fossil record of Mesozoic birds has improved dramatically in recent years (Chiappe & Dyke,

2002; Zhou, 2004), and despite claims for the record's completeness (Fountaine *et al.*, 2005), a number of major biases are evident. Recent discoveries have been dominated by numerous exceptionally preserved taxa from the lagerstätten of the Yixian (Early Cretaceous: latest Barremian–early Aptian) and Jiufotang (Early Cretaceous: Aptian) formations of China, and this bias is immediately evident in figures 3 and 4 of HEA, and, indeed, in our Figs 1 and 2. Equivalent lagerstätten are unknown from the Late Cretaceous and preservation of birds in this epoch is generally poor – figure 4 of Fountaine *et al.* (2005) demonstrates that although numerous avian taxa have been named from the Campanian and the Maastrichtian, nearly all have been named on the basis of very fragmentary material. The difference in preservation potential between the Early and Late Cretaceous may well account for the scarcity of Late Cretaceous birds (and in particular the scarcity of small taxa) in the data set of HEA. Indeed, some notably small Late Cretaceous taxa were excluded from the HEA data set due to their poor preservation: an obvious example is the sparrow-sized enantiornithine *Alexornis antecedens* from the Campanian of Mexico (Brodkorb, 1976). Furthermore, our comparisons of PD and stratigraphic age suggest low stratigraphic congruence for Mesozoic avian phylogeny, suggesting major gaps in the completeness of the bird record.

Nonphylogenetic analyses of body size evolution should, whenever possible be accompanied by analyses carried out in an explicit phylogenetic context that explore ancestor–descendant changes – this is the only approach capable of identifying nonrandom within-lineage trends (Alroy, 2000). Numerous approaches to phylogenetic reconstruction of body size evolution are available (e.g. Alroy, 1998, 2000; Laurin, 2004; Carrano, 2005; Hone *et al.*, 2005; Finarelli & Flynn, 2006; Moen, 2006), although some may be more reliable than others.

Although taxon-sampling remains low in comparison to other groups of Mesozoic archosaurs, moderately large phylogenies for Mesozoic birds are available (e.g. Clarke *et al.*, 2006; Chiappe *et al.*, 2007), particularly if combined with one another, and should be utilized wherever possible. Our explicitly phylogenetic body size analyses find that positive ancestor–descendent body size changes do not occur significantly more commonly than negative ancestor–descendent body size changes (in fact, negative body size changes seem to be more common, but not significantly), although positive body size changes (when they occur) appear to be generally of greater magnitude than negative body size changes. It is possible to envision a scenario where this pattern might result in an overall increase in body size within a clade. However, the infrequency of positive body size changes means that there is little convincing evidence for an active within-lineage trend towards increased body size amongst Mesozoic birds (*contra* HEA). That this should be the case is perhaps unsurprising: the earliest bird, *Archaeopteryx*, is relatively large in size, as are basal members of the clades Pygostylia (*Confuciusornis*), Ornithothoraces (*Protopteryx*, *Vorona*), Enantiornithes (*Protopteryx*) and Ornithuromorpha (*Vorona*). Figure 3 of HEA shows known FLs for Mesozoic birds plotted on a temporal scale – it is evident from their figure that the majority of known post-*Archaeopteryx* Mesozoic avian diversity is of taxa that are smaller in size than is *Archaeopteryx*, and that only a few species evolved larger body sizes (Fig. 2).

Contradictions are apparent in the conclusions drawn by HEA. On the one hand, the authors stated that Cope's rule occurs in Mesozoic birds as a result of a 'reduction in variance', with effects that are 'in any case different from the "small-ancestor problem" of Stanley' (Hone *et al.*, 2008). By contrast, in the next paragraph they stated that Mesozoic birds were undergoing an 'increase in variance' and that observed body size change 'does not require the assumption of a positive driving force [...] other than the observation that early members of the clade were small'. In any case, our analyses show that in fact variance did not decrease markedly over time in Mesozoic birds. Our reanalysis of their data, and new analyses carried out within an explicit phylogenetic framework, indicate that there is little compelling evidence for an overall trend toward increased body size, either within- or among-lineages – we thus believe that the identification of Cope's rule in Mesozoic birds by HEA was both premature and erroneous based upon the data and statistical results that were available to them. Our understanding of Mesozoic bird diversity and phylogeny, and of the utility of body size proxies such as FL, remains in its infancy.

Hone *et al.* (2008) claimed to have identified evidence of a trend towards body size decrease in Ornithuromorpha, the clade most closely related to modern birds, and speculated that this might help explain the survivorship of birds across the KPg boundary. As discussed

above, the results of HEA for Ornithuromorpha are not statistically significant, and the regression line shown in their figure 4d possesses a negative slope largely because of the presence of a single (small) outlier. We find no significant evidence to support the hypothesis that Ornithuromorpha were undergoing a trend towards decreased body size prior to the KPg extinction; the data of HEA can therefore contribute little to the discussion of avian survivorship across the KPg boundary.

Acknowledgments

The authors thank David Hone for providing us with a copy of the original body size data set. Thanks to Paul Barrett, Mark Bell, John Finarelli, David Marjanović, and Jason Moore for discussion and comments on previous versions of this manuscript. We thank Michel Laurin for his helpful review comments.

References

- Alroy, J. 1998. Cope's rule and the dynamics of body mass evolution in North American fossil mammals. *Science* **280**: 731–734.
- Alroy, J. 2000. Understanding the dynamics of trends within evolving lineages. *Paleobiology* **26**: 319–329.
- Blackburn, T.M. & Gaston, K.J. 1994. The distribution of body sizes of the world's bird species. *Oikos* **70**: 127–130.
- Brodtkorb, P. 1976. Discovery of a Cretaceous bird, apparently ancestral to the Orders Coraciiformes and Piciformes (Aves, Carinatae). *Smithson. Contrib. Paleobiol.* **27**: 67–73.
- Brown, J.H. & Sibly, R.M. 2006. Life-history evolution under a production constraint. *PNAS* **103**: 17595–17599.
- Carrano, M.T. 2005. Body-size evolution in the Dinosauria. In: *Amniote Paleobiology* (M.T. Carrano, T.J. Gaudin, R.W. Blob & J.R. Wible, eds), pp. 225–268. University of Chicago Press, Chicago, IL.
- Chiappe, L.M. & Dyke, G.J. 2002. The Mesozoic radiation of birds. *Annu. Rev. Ecol. Syst.* **33**: 91–124.
- Chiappe, L.M., Suzuki, S., Dyke, G.J., Watabe, M., Tsogtbaatar, K. & Barsbold, R. 2007. A new enantiornithine bird from the Late Cretaceous of the Gobi Desert. *J. Syst. Palaeontol.* **5**: 193–208.
- Clarke, J.A., Tambussi, C.P., Noriega, J.I., Erickson, G.M. & Ketchum, R.A. 2005. Definitive fossil evidence for the extant avian radiation in the Cretaceous. *Nature* **433**: 305–308.
- Clarke, J.A., Zhou, Z. & Zhang, F. 2006. Insight into the evolution of avian flight from a new clade of Early Cretaceous ornithurines from China and the morphology of *Yixianornis grabaui*. *J. Anat.* **208**: 287–308.
- Felsenstein, J. 1985. Phylogenies and the comparative method. *Am. Nat.* **125**: 1–15.
- Finarelli, J.A. & Flynn, J.J. 2006. Ancestral state reconstruction of body size in the Caniformia (Carnivora, Mammalia): the effects of incorporating data from the fossil record. *Syst. Biol.* **55**: 301–313.
- Fountaine, T.M.R., Benton, M.J., Dyke, G.J. & Nudds, R.L. 2005. The quality of the fossil record of Mesozoic birds. *Proc. R. Soc. Lond. B* **272**: 289–294.

- Gould, S.J. 1997. Cope's rule as psychological artifact. *Nature* **385**: 199–200.
- Hammer, Ø., Harper, D.A.T. & Ryan, P.D. 2001. **PASt**: paleontological statistics software package for education and data Analysis. *Palentol. Electronica* **4**: 9.
- Hone, D.W.E. & Benton, M.J. 2007. Cope's rule in the Pterosauria, and differing perceptions of Cope's rule at different taxonomic levels. *J. Evol. Biol.* **20**: 1164–1170.
- Hone, D.W.E., Keesey, T.M., Pisani, D. & Purvis, A. 2005. Macroevolutionary trends in the Dinosauria: Cope's rule. *J. Evol. Biol.* **18**: 587–595.
- Hone, D.W.E., Dyke, G.J., Haden, M. & Benton, M.J. 2008. Body size evolution in Mesozoic birds. *J. Evol. Biol.* **21**: 618–624.
- Hou, L., Martin, L.D., Zhou, Z. & Feduccia, A. 1999. *Archaeopteryx* to opposite birds – missing link from the Mesozoic of China. *Vertebr. Palasiat.* **4**: 88–95.
- Hou, L., Chiappe, L.M., Zhang, F. & Chuong, C.M. 2004. New Early Cretaceous fossil from China documents a novel trophic specialization for Mesozoic birds. *Naturwissenschaften* **91**: 22–25.
- Jablonski, D. 1997. Body-size evolution in Cretaceous molluscs and the status of Cope's rule. *Nature* **385**: 250–252.
- Kingsolver, J.G. & Pfennig, D.W. 2004. Individual-level selection as a cause of Cope's rule of phyletic size increase. *Evolution* **58**: 1608–1612.
- Laurin, M. 2004. The evolution of body size, Cope's rule and the origin of amniotes. *Syst. Biol.* **53**: 594–622.
- Maddison, W.P. & Maddison, D.R. 2007. *Mesquite: a modular system for evolutionary analysis*. Version 2.01. <http://mesquiteproject.org>.
- Martins, E.P., Diniz-Filho, J.A.F. & Housworth, E.A. 2002. Adaptive constraints and the phylogenetic comparative method: a computer simulation test. *Evolution* **56**: 1–13.
- McShea, D.W. 1994. Mechanisms of large-scale evolutionary trends. *Evolution* **48**: 1747–1763.
- Meiri, S. 2007. Size evolution in island lizards. *Global Ecol. Biogeogr.* **16**: 702–708.
- Moen, D.S. 2006. Cope's rule in cryptodiran turtles: do the body sizes of extant species reflect a trend of phyletic size increase? *J. Evol. Biol.* **19**: 1210–1221.
- Norell, M.A. & Xu, X. 2005. Feathered dinosaurs. *Annu. Rev. Earth Planet. Sci.* **33**: 277–299.
- Padian, K. 2004. Basal Avialae. In: *The Dinosauria*, 2nd edn (D.B. Weishampel, P. Dodson & H. Osmólska, eds), pp. 210–231. University of California Press, Berkeley, CA.
- Stanley, S.M. 1973. An explanation of Cope's rule. *Evolution* **27**: 1–26.
- Turner, A.H., Pol, D., Clarke, J.A., Erickson, G.M. & Norell, M.A. 2007. A basal dromaeosaurid and size evolution preceding avian flight. *Science* **317**: 1378–1381.
- Wilkinson, M. 1994. Common cladistic information and its consensus representation: reduced Adams and reduced cladistic consensus trees and profiles. *Syst. Biol.* **43**: 343–368.
- Wilkinson, M. 1995. More on reduced consensus methods. *Syst. Biol.* **44**: 435–439.
- You, H.-L., Lamanna, M.C., Harris, J.D., Chiappe, L.M., O'Connor, J., Ji, S.-A., Lü, J.-C., Yuan, C.-X., Li, D.-Q., Zhang, X., Lacovara, K.J., Dodson, P. & Ji, Q. 2006. A nearly modern amphibious bird from the Early Cretaceous of northwestern China. *Science* **312**: 1640–1643.
- Zhou, Z. 2004. The origin and early evolution of birds: discoveries, disputes and perspectives from the fossil record. *Naturwissenschaften* **19**: 455–471.

Received 29 February 2008; revised 30 May 2008; accepted 30 June 2008

Supporting information

Additional supporting information may be found in the online version of this article:

Table S1 Mesquite input file for supertree A.

Table S2 Mesquite input file for supertree B.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.