# Macroevolutionary trends in the Dinosauria: Cope's rule

D. W. E. HONE,*† T. M. KEESEY,‡ D. PISANI§ & A. PURVIS*

*Department of Biological Sciences, Imperial College London, Ascot, UK
†Department of Entomology, The Natural History Museum, London, UK
‡834 Hyperion Avenue, Los Angeles, CA, USA
§Department of Zoology, The Natural History Museum, London, UK

## Abstract

Cope's rule is the tendency for body size to increase over time along a lineage. A set of 65 phylogenetically independent comparisons, between earlier and later genera, show that Cope's rule applied in dinosaurs: later genera were on average about 25% longer than the related earlier genera to which they were compared. The tendency for size to increase was not restricted to a particular clade within the group, nor to a particular time within its history. Small lineages were more likely to increase in size, and large lineages more likely to decrease: this pattern may indicate an intermediate optimum body size, but can also be explained as an artefact of data error. The rate of size increase estimated from the phylogenetic comparisons is significantly higher than the rate seen across the fauna as a whole. This difference could indicate that within-lineage selection for larger size was opposed by clade selection favouring smaller size, but data limitations mean that alternative explanations (which we discuss) cannot be excluded. We discuss ways of unlocking the full potential usefulness of phylogenies for studying the dynamics of evolutionary trends.

## Introduction

Cope's rule was first outlined by Edward Drinker Cope and states that organisms along an evolutionary lineage will tend to increase in size (Cope, 1887, 1896). Larger size usually confers a selective advantage on individuals within populations (Kingsolver & Pfennig, 2004). There are many, nonexclusive, reasons why larger size might confer selective advantage, notably enhanced defence, success in mating competition, extended longevity, increased intelligence (with increasing brain size), greater range of acceptable foods and predatory success, and the potential for thermal inertia (especially relevant to dinosaurs) (Bonner, 1968; Schmidt-Nielsen, 1984; Paul, 2000; Benton, 2002). However, large size is also believed to make taxa prone to extinction, because it is associated with long generation times and low abundance, making them slow to adapt to environmental pressures and recover from any significant drop in population size (McKinney, 1997).

*Correspondence:* David Hone, Department of Earth Sciences, University of Bristol, Wills Memorial Building, Queens Road, Bristol BS8 1RJ, UK.
Tel.: 0117 954 5244; fax: 0117 925 3385;
e-mail: david.hone@bristol.ac.uk

Cope's rule, where found, is a result of one or more of three processes: (i) necessary increase in average size when founders of clades are small organisms (Cope himself suggested that the trends he observed were the result of ancestors being small: Stanley, 1973); (ii) genuine natural selection where the advantages of larger size act to produce progressively larger descendants; or (iii) clade selection where larger taxa within a lineage tend to survive or proliferate in preference to smaller taxa (Cope, 1896; Vrba & Gould, 1986; McShea, 2000; Carroll, 2001; Benton, 2002; Kingsolver & Pfennig, 2004). Gould (1997) proposed that demonstration of Cope's rule would require rigorous testing across whole lineages, and that any assessment of generality would require many such analyses.

Here, we aim to perform a rigorous test for Cope's rule in Dinosauria and to partly discriminate among the possible processes giving rise to it, using a combination of phylogenetic and nonphylogenetic comparative analysis across the whole group. Alroy (1998, 2000) noted that insights into the dynamics behind evolutionary trends can be obtained by comparing the rate of size increase across a fauna as a whole with the rate of size increase seen along phylogenetic lineages. Both rates include

components of among-lineage and within-lineage selection, but in different amounts: the fauna-wide rate reflects among-lineage selection to a greater degree than the rate obtained from phylogenetic analyses. If the rates do not differ significantly (as Alroy, 1998 found for mammals), then within-lineage selection could be responsible for any overall trend. If the fauna-wide rate is higher, clade selection is implicated alongside natural selection; conversely, if the fauna-wide rate is significantly lower, clade selection among lineages is inferred to have opposed natural selection within them. Is there any evidence of clade selection in dinosaurs?

The dinosaurs are an ideal group for such an analysis for several reasons. First, they achieved colossal size and produced the biggest terrestrial herbivores and carnivores on record, including some genera thought to have approached the maximum possible size for a terrestrial organism. However, they are also famous for producing a classic example of phyletic dwarfism – the birds. This means that they include both size increase and decrease, and contain a large amount of variation. The dinosaurs also extend over a long and well-defined period of time. Because they represent almost all of the terrestrial macrofauna of that time, the possible effects of interclade competition can reasonably be ignored. A charismatic group, the dinosaurs benefit from a high level of interest and much data has been generated about them and their evolutionary relationships, including the recent phylogenetic supertree of dinosaurs by Pisani *et al.* (2002). Finally, as large vertebrates with a good fossil record in continental Europe and North America, they are well represented in terms of both numbers and completeness of specimens.

## Materials and methods

### Data

We used total adult body length in metres (data available at http://palaeo.gly.bris.ac.uk/hone/data.html) to indicate body size as most estimates of mass (which are usually used in studies of Cope's rule) use length to calculate volume and therefore mass and are often controversial at best (Colbert, 1962). Seebacher (1999) has also shown that length is a good substitute for mass in the Dinosauria, at least within taxa whose members are constructed on a similar bauplan. Other methods either show too little separation among taxa or involve lengthy calculations, which is impractical with so many species and genera (Colbert, 1962; Alexander, 1985; Henderson, 1999; Seebacher, 1999). Sexual dimorphism is limited in the Dinosauria (Fastovsky & Weishampel, 1996), and closure of the neurocentral suture makes adults relatively easy to distinguish from juveniles.

Species data were $log_{10}$-transformed, then averaged within genera: this deviates from Alroy's (1998) methods because there are relatively few dinosaur species per genus [estimated at 1.2 by Dodson, although this has probably fallen since then with the discovery of more monotypic genera] (Lewin, 1990). However it reduces reliance on single specimens as an estimate for the whole genus. Jablonski (1997) has suggested that testing at the generic level is appropriate as Cope's rule is perceived to be a large-scale trend.

Time data was taken as the midpoint of the age of earliest known appearance of each genus (e.g. mid-Maastrichtian, mid-Campanian, etc.); more precise data are seldom available so we did not use them.

### Analyses

The evolutionary dynamics underlying any trends are most clearly seen with an analysis that is based on phylogeny. Ideally, it would be possible to compare ancestors with descendants (Alroy, 2000), but the incompleteness of the fossil record generally precludes such a design. We therefore constructed phylogenetically independent comparisons between related taxa of different ages, using a recent supertree of 272 genera (Pisani *et al.*, 2002), the most inclusive dinosaur phylogeny available. A comparison in which the later taxon is larger than the earlier taxon supports Cope's rule. Under the null hypothesis of no body size trend, around half of such comparisons should show a size increase and the remainder a size decrease, with the mean change across all comparisons not differing significantly from zero.

The power to reject the null hypothesis increases with the temporal separation between the early and late taxa within comparisons, and with the number of comparisons. Comparisons between taxa of very similar age are likely to be uninformative, so we only compared taxa that differed in age by at least 10 Myr. (This constraint also reduces the risk that ghost ranges might invalidate the comparison.) Within this constraint, we maximized the number of independent comparisons and, within each comparison, maximized the temporal separation. Initial comparisons were identified by finding the cladistically most closely related taxa whose ages differed by at least 10 Myr and for which size data were available. Each taxon could be either a single genus or multiple genera of the same age. Lines were drawn through the branches of the phylogeny to link the taxa being compared; comparisons are phylogenetically independent if their lines neither meet nor cross (Burt, 1989). Further independent comparisons were then identified by proceeding through the phylogeny, from the most distal taxa not yet involved in comparisons towards the root; this approach maximizes sample size (Purvis & Rambaut, 1995). However, we did not make all possible comparisons near the root because length reflects mass accurately only among taxa of the same general shape and build (see above). We therefore only compared taxa having similar mode of locomotion (quadrupeds are heavier-set than bipeds) and diet (herbivores are heavier-set than carnivores)

(Seebacher, 1999). In comparisons where the bauplan is varied (e.g. within the Stegosauria), the lengths are very different and thus not likely to affect the proxy of length for mass. Having identified our maximal set of comparisons, we were sometimes able to increase the within-comparison temporal separation, by substituting the earlier taxon identified by the algorithm with a taxon having even greater stratigraphic age. Figure 1 shows the cladogram for the eumaniraptorans and demonstrates how the pairs were selected for comparison.

For each comparison, the average log(size) of earlier genera was subtracted from the average log(size) of later genera to compute a size increase (which may be negative); the pooled set of size increases were analysed with *t*-tests (to test whether the mean differed from zero) and Wilcoxon signed-ranks tests (to test whether the median differed from zero). All tests were two-tailed. An overall rate of within-lineage size increase was estimated by least-squares regression through the origin (Garland *et al.*, 1992) of size increase on temporal separation.
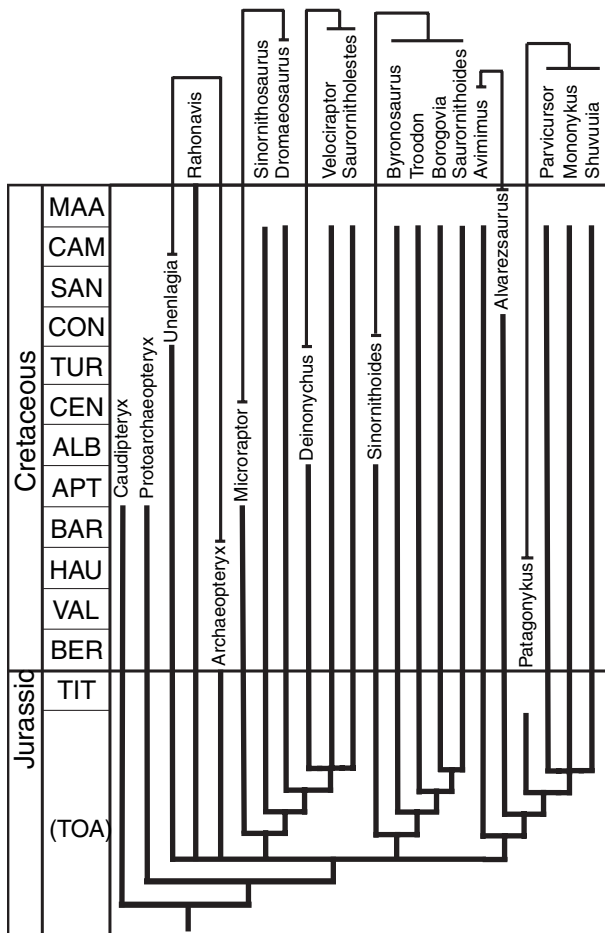


**Fig. 1** Cladogram for the eumaniraptorans (redrawn from Pisani *et al.*, 2002), demonstrating how the comparisons between earlier and later genera were constructed.

More complex patterns are possible than a simple size increase or decrease, and these can be detected by analysing the size increases together with the sizes of the taxa being compared (Alroy, 2000). We regressed size change on the average size of taxa being compared; a significant result would indicate a nonlinear trend, such as might be obtained if a lineage was evolving up against a boundary minimal or maximal size. A regression of size change on the size of the earlier taxon being compared was used to test for the existence of an attractor – a body size towards which lineages tended to evolve from above and below (Alroy, 2000), although the interpretation of a significant result is hampered by a 'regression to the mean' artefact (Alroy, 1998, 2000; see also Discussion below). To test the temporal consistency of any trend, we regressed size change on the average date of the taxa being compared. We used logistic regression to test whether the sign of the size change was predicted by the (log transformed) temporal separation of the taxa being compared: if there has been a general trend in body size, but the size data contain random error, then the trend should be most apparent in comparisons spanning a long interval.

Nonphylogenetic analysis is, by itself, uninformative about the evolutionary process responsible for any trend (Alroy, 2000) but it does provide an estimate of the overall fauna-wide rate of size change that can be compared with the estimate obtained from the matched-pairs comparisons. The rate was therefore also estimated as the slope of the least-squares regression of log(size) on date, across all genera in the dataset. This rate was compared with that obtained from the matched-pairs comparisons by an approximate *t*-test analogous to Welch's *t*-test (Brown & Rothery, 1993).

We explored patterns of size change in two other ways. First, Jablonski (1997) provided a graphical method of representing Cope's rule or other trends in size over time. This uses deformed cylinders to represent the variance of size changing over time within taxa. We have adapted this approach to demonstrate Cope's rule with a 'Jablonski polygon'. Simply, the sizes of the taxa under study are plotted against time (or stratigraphy). A line is then drawn between the largest and smallest taxa at the earliest and latest time. If Cope's rule has been operating, this should provide a right-leaning rhombus, as both the smallest and largest taxa at a later evolutionary time should be larger than their respective counterparts at an earlier time. The scatter for the other taxa that appear within the rhombus will give an indication of how much variation lies outside this trend (ideally all will lie within the borders). The polygons may also expand or contract, indicating an increase or decrease in variance respectively. This is a very simple method and should only be used as an indication of trends and a demonstration of variance within a group.

Finally, we made comparisons within the 20 families delineated in Pisani *et al.* (2002) supertree. The families

contain an average of about 11 genera. We assessed how often the biggest genus appeared after, rather than before, the smallest; how often the last genus was larger, rather than smaller, than the earliest; and how often the largest genus was among the last known in the family.

Phylogenetic analyses of the fossil data require that both the phylogeny and the fossil record are adequate for the purpose. We used the method of Benton (1995) to test the goodness of fit between the phylogeny and the stratigraphy. Briefly, the phylogeny is collapsed to a Hennigian comb, giving an unambiguous order of origination of all nodes; the nodes are also ranked by the time of earliest appearance of any taxon; the two sets of ranks are then compared using Spearman rank correlation. To reduce the loss of information entailed by reducing the phylogeny to a Hennigian comb, we tested both orders and each family and superfamily separately

## Results

We were able to test for congruence between stragriphy and phylogeny using Spearman rank correlation for a total of 18 clades. Of these 18 analyses, 13 were significant ($P < 0.05$), including both the orders and both superfamilies. Of those which were not significant, two were very small clades, meaning that the result was based on very little data, one was very poorly resolved taxonomically, and one covered only two stratigraphic periods, making ranking difficult. Overall therefore, we found a close match between phylogeny and stratigraphy indicating that both the phylogeny and the fossil record are probably of sufficient quality to underpin our analyses. The families of Protoceratopsidae, Iguanodontidae and Spinosauridae could not be tested as they did not have the minimum required our nodes for testing (Benton, 1995), and the Diplodocoidea could no be tested as they are all from the same stratigraphic period.

The 65 phylogenetically independent matched-pairs comparisons that were constructed are listed in Table 1. In them, the later taxa are on average 25.7% longer than the earlier (mean change in $\log_{10}$(length) = 0.0992, $t_{64} = 2.908$, $P < 0.01$; median change = 0.0969, $W = 1393.5$, $P < 0.01$). A histogram of the size changes is shown in Fig. 2. Across all 65 comparisons, regression through the origin of size increase on temporal separation gives a rate of 0.00283 $\log_{10}$ m/Myr (SE = 0.00085). Size changes are independent of the average size of the taxa being compared ($t_{63} = -1.66$, $P = 0.10$), indicating a constant directional bias towards larger size (Alroy, 2000). Size changes are predicted by the size of the earlier taxon ($t_{63} = -4.579$, $P < 0.001$): the regression indicates a stable equilibrium body length of around 7.8 m (where the regression crosses the *x*-axis, i.e. with earlier taxa of this size, the later taxa will on average be neither smaller nor larger). However, although this regression is highly significant, it is not strongly explan-

atory ($r^2 = 0.24$); the 95% confidence interval on the equilibrium body length is 4.9–17.5 m and the 95% inverse prediction intervals are extremely wide (0.19–462 m). Size changes are independent of the average ages of taxa being compared ($t_{63} = -0.335$, n.s.), indicating that the trend was consistent over time. The logistic regression showed a near-significant tendency for more comparisons to be positive as the temporal separation between taxa increases ($\chi_1^2 = 3.134$, $P = 0.08$).

The distribution of genus size against time (midpoint of stratigraphic stage of earliest appearance) demonstrating the variation in size through the Mesozoic and the slow trend towards greater size can be seen in Fig. 3. Linear regression of these data gives a fauna-wide rate of 0.00081 $\log_{10}$ m/Myr (SE = 0.00051). This rate is significantly lower than that seen within comparisons ($t = 2.032$, d.f. = 85.5, $P < 0.05$).

The creation of Jablonski polygons allows for a more visual representation of the change in size among several dinosaur groups (Fig. 4). Here polygons are shown for the sauropods (black, dashes) and theropods (grey, crosses) [with the eumaniraptorans (pale grey, circles)]. The polygons contain most of the data points (80% of the theropods and 67% of the sauropods). This is not in itself significant but does demonstrate that the polygon is an accurate reflection of the data – the points do not just connect some outliers or remain restricted to the centre of the clusters. In each case the variance and maximum size increase throughout each period with extinction events at the end of the Triassic and Jurassic reducing them.

Within the 18 families, the largest genus appeared after the smallest in 12 of 15 (three were all the same age; sign test: $P < 0.01$). The last genus was larger than the earliest in 12 of 15 families (sign test: $P < 0.018$). The largest genus was among the last in 10 of 15 families (sign test: n.s.).

## Discussion

Our phylogeny-based analysis shows a clear trend towards larger body size across the Dinosauria. The trend is not limited to any particular subclade and is found in predators and herbivores, in quadrupeds and bipeds. Notably, the size increases occur at various starting sizes and so the size change is not just a result of beginning with small genera. Curiously, the eumaniraptorans, the most direct bird ancestors, show no marked trend, despite the fact that they are generally quoted as showing a decrease in size as a flight adaptation (Sereno, 1997). However, this may be a result of their very unusual stratigraphic distribution, with *Archaeopteryx* appearing in the mid Jurassic, but many nonflying genera from the group appearing in the late Cretaceous. Previous examples of Cope's rule include North American fossil mammals (Cope, 1887; Alroy, 1998), parasitic isopods (Poulin, 1995), fossil foraminifera (Arnold *et al.*, 1995; Webster &

**Table 1** Sixty-five phylogenetically independent matched-pairs comparisons selected for analysis based on the phylogeny.

| Comparison no. | Ancestors | Descendants | Time difference (Myr) | Size difference (log) |
|---|---|---|---|---|
| 1 | Patagonykus | Parvicursor | 107.5 | −0.30 |
| | | Mononykus | | |
| | | Shuvuula | | |
| 2 | Sinornithoides | Byronosaurus | 39.2 | 0.30 |
| | | Troodon | | |
| | | Borogovia | | |
| | | Suarornithoides | | |
| 3 | Deinonychus | Saurornitholestes | 39.2 | −0.21 |
| | | Velociraptor | | |
| 4 | Microraptor | Dromaeosaurus | 46.6 | 0.60 |
| 5 | Alvarezsaurus | Avimimus | 10.5 | −0.30 |
| 6 | Caudipteryx | Elmisaurus | 51.2 | 0.52 |
| 7 | Aralosaurus | Gryposaurus | 14.3 | 0.15 |
| 8 | Erlikosaurus | Theriznosaurus | 18.8 | 0.20 |
| | Segnosaurus | | | |
| 9 | Beipiasaurus | Alaxasaurus | 18.5 | 0.23 |
| 10 | Archaeopteryx | Unenlagia | 55.8 | 0.74 |
| 11 | Archaeornithomimus | Gallimimus | 28.6 | 0.15 |
| 12 | Garudimimus | Deinocheirus | 19.7 | 0.40 |
| 13 | Pelecanimimus | Harpymimus | 12.9 | 0.00 |
| 14 | Compsognathus | Sinosauropteryx | 28.4 | 0.11 |
| 15 | Proceratosaurus | Ornitholestes | 19.4 | −0.18 |
| 16 | Coelurus | Alioramus | 79.3 | 0.62 |
| | | Tarbosaurus | | |
| | | Tyrannosaurus | | |
| 17 | Acrocanthosaurus | Carcharodontosaurus | 11.1 | 0.09 |
| | | Giganotosaurus | | |
| 18 | Allosaurus | Neovenator | 28.4 | −0.14 |
| 19 | Cryolophosaurus | Sinraptor | 40.1 | 0.07 |
| 20 | Gasosaurus | Marshosaurus | 19.4 | 0.15 |
| 21 | Piatnitkysaurus | Afrovenator | 32.4 | 0.15 |
| 22 | Eustreptospondylus | Torvosaurus | 14.5 | 0.14 |
| 23 | Baryonyx | Irritator | 18.5 | −0.09 |
| | | Spinosaurus | | |
| 24 | Carnotaurus | Majungatholus | 28.2 | 0.03 |
| 25 | Ligabueino | Masiakasaurus | 61.4 | 0.63 |
| | | Noasaurus | | |
| | | Abelisaurus | | |
| 26 | Sarcosaurus | Caratosaurus | 46.2 | 0.18 |
| 27 | Coelophysis | Syntarsus | 20.3 | −0.14 |
| 28 | Lilensternus | Dilophosaurus | 15.9 | −0.07 |
| 29 | Procompsognathus | Megalosaurus | 36.9 | 0.90 |
| 30 | Argyrolosaurus | Aeolosauruss | 28.1 | −0.28 |
| | | Alamosaurus | | |
| | | Nequensaurus | | |
| 31 | Argentinosaurus | Opisthocoelocaudia | 28.2 | −0.52 |
| 32 | Brachiosaurus | Sauroposeidon | 46.9 | 0.06 |
| 33 | Dicraeosaurus | Amargasaurus | 22.9 | −0.25 |
| 34 | Apatosaurus | Antarctosaurus | 79.6 | 0.20 |
| | Barosaurus | | | |
| | Diplodocus | | | |
| 35 | Euhelops | Nemegtosaurus | 75.0 | −0.02 |
| 36 | Shunosaurus | Camarasauruss | 14.4 | 0.17 |
| | | Haplocanthosaurus | | |
| 37 | Vulcanodon | Barapasaurus | 19.0 | 0.44 |
| 38 | Euskelosaurus | Yunnanosaurus | 20.3 | −0.28 |
| | | Massospondylus | | |
| | | Lufengosaurus | | |

Table 1 Continued.

| Comparison no. | Ancestors | Descendants | Time difference (Myr) | Size difference (log) |
|---|---|---|---|---|
| 39 | Blikanasaurus | Camelotia | 16.4 | 0.35 |
| 40 | Thecodontosaurus | Anchisaurus | 22.7 | 0.06 |
| 41 | Brachylophosaurus | Shantungosauruss Anatotitan Edmontosaurus | 12.9 | 0.26 |
| 42 | Bactrosaurus | Pararabdon Charonosaurus Hypacrosaurus | 28.1 | 0.22 |
| 43 | Eolambia | Gilmoreosaurus | 11.6 | −0.05 |
| 44 | Probactrosaurus | Telmatosaurus | 50.9 | −0.08 |
| 45 | Iguanadon | Altirhinus Ouranosaurus | 11.9 | −0.04 |
| 46 | Camptosaurus | Muttaburrasaurus | 48.0 | 0.12 |
| 47 | Hypsilophodon | Parkosaurus | 55.9 | 0.08 |
| 48 | Zephyrosaurus | Orodromeus | 39.2 | 0.14 |
| 49 | Yandusaurus | Othnelia | 14.4 | 0.03 |
| 50 | Tenontosaurus | Thescelosaurus | 39.2 | −0.27 |
| 51 | Zuniceratops | Triceratops Arrinoceratops Torosaurus Pachyrhinosaurus | 23.6 | 0.33 |
| 52 | Breviceratops Protoceratops Graciliceratops | Montanoceratops | 16.5 | 0.22 |
| 53 | Udanoceratops | Leptoceratops | 16.5 | −0.30 |
| 54 | Yaverlandia | Stygimoloch | 55.9 | 0.48 |
| 55 | Stenopelix | Wannanosaurus Homalocephale | 46.6 | −0.05 |
| 56 | Pinacosaurus | Ankylosaurus | 16.5 | 0.28 |
| 57 | Mymoorapelta | Tsagantegia Talarurus | 56.2 | 0.14 |
| 58 | Gargoyleosaurus | Minmi Shamosaurus | 35.8 | 0.10 |
| 59 | Sauropelta Silvisaurus | Edmontonia Panoplosaurus | 39.2 | 0.09 |
| 60 | Hylaeosaurus | Struthiosaurus | 57.1 | −0.10 |
| 61 | Stegosaurus Kentrosaurus | Wuerhosaurus | 17.9 | −0.05 |
| 62 | Huayangosaurus | Toujiangosaurus | 14.4 | 0.24 |
| 63 | Scelidosaurus | Emausaurus | 13.8 | −0.24 |
| 64 | Agilisaurus | Gongbusaurus | 10.1 | −0.08 |
| 65 | Pisanosaurus | Lesothosaurus | 20.3 | 0.00 |

This gives the ancestor–descendant pairs, time difference and difference in log size. The data for this table were used to construct Figs 2 and 3.

Purvis, 2002) and perhaps prehistoric horses (McFadden, 1992); bivalves and gastropods are counterexamples (Jablonski, 1996, 1996, 1997; see p. 263 for a full list). Of these studies, only Alroy (1998) and Jablonski (1996, 1997) applied a phylogeny to their analyses and both used within-genus comparisons and so did not necessarily require a higher-level phylogeny.

What was the mechanism behind the size increase in dinosuars? The significant difference between the rates estimated from phylogenetic and nonphylogenetic comparisons suggests that both within-lineage (natural selection or sexual selection) and among-lineage (Williams, 1992) processes may have been operating (Alroy, 1998, 2000). Our results appear to indicate that within-lineage size increases are driving the overall trend, and that among-lineage selection is opposing it. Kingsolver & Pfennig (2004) also proposed a conflict between within-lineage and among-lineage selection as one explanation for the strong selective advantage that they found larger size to confer. The strength of directional selection they recorded, if extrapolated, would lead to very much more rapid size increases than have been observed; they suggested that size-selective mass extinction could explain the discrepancy.
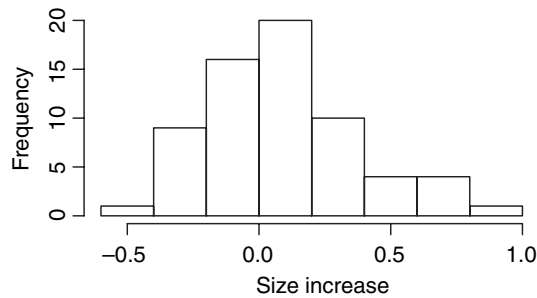
**Fig. 2** Histogram of size change in the 65 phylogenetically independent comparisons. Positive size change indicates a comparison supporting Cope's rule.
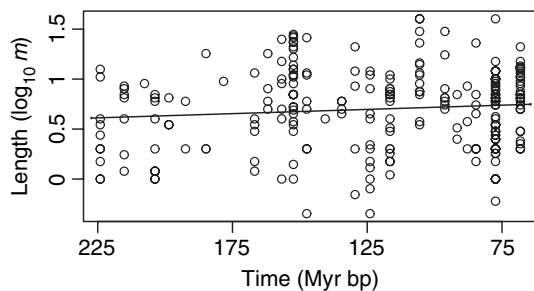


**Fig. 3** Sizes of dinosaur genera in the phylogeny, plotted against the midpoint of the earliest stratigraphic age from which they are known. The line is the least-squares linear regression line.

Limitations of our study mean that other explanations of our results are also possible, however. Ideally, we would be able to make direct ancestor–descendant comparisons based on a sufficiently detailed fossil record to permit exclusion of any among-lineage effects. Such comparisons would reflect only within-lineage

evolution, presumably due to natural or sexual selection (Alroy, 2000). However, such a design would require a sufficiently complete fossil record to trace population history through geological time; otherwise size-selective competition between newly formed lineages could in fact be responsible for 'within-lineage' trends. Our phylogenetically matched-pairs comparisons inevitably include some among-lineage component, possibly a considerable one, although within-lineage evolution is reflected more in the phylogenetic than the nonphylogenetic comparisons. Consequently, an alternative explanation for the difference between phylogenetic and nonphylogenetic rates (which we owe to an anonymous referee), involving only clade selection, cannot be discounted. If clade selection always favoured large size, but was strongest among closely related lineages (which might be expected to have competed most intensely: Darwin, 1859), the rate of size increase would appear higher in phylogenetic than nonphylogenetic analyses. We emphasize that neither of these scenarios requires that body size *per se* was a target of clade selection: speciation and extinction rates could instead have been influenced by any of the many traits correlated with body size among dinosaur lineages.

Alternatively, the significant difference between the rates estimated from phylogenetic and nonphylogenetic comparisons could be artefactual. An important bias in the fossil record, which reduces the power to detect Cope's rule, could cause the observed rate difference, as follows. Ages of first appearance in the fossil record are generally underestimates of the true date of origination, and cannot be overestimates. This is particularly true for older taxa (whose fossils are more likely to have been destroyed over time) and smaller taxa (whose fossils are harder to find). Therefore the data are likely to be skewed towards large, late genera, making detection of Cope's rule less likely. However, the bias in ages of first



**Fig. 4** Scatter plot of the theropods (grey crosses), sauropods (black dashes) and eumaniraptorans (pale grey circles). The extremes of body size within periods are linked to form 'Jablonski polygons' demonstrating large-scale trends in size change. See text for explanation. In this case, we are lacking in taxa at the lowest horizons, and so the next lowest horizon has been used to provide a lower boundary. Pictured genera are: 1 *Masiakasaurus*, 2 *Tyrannosaurus*, 3 *Alamosaurus*, 4 *Spinosaurus*, 5 *Argentinosaurus*, 6 *Ornitholestes*, 7 *Amargasaurus*, 8 *Mamenchisaurus*, 9 *Vulcanodon*, 10 *Herrerasaurus*.
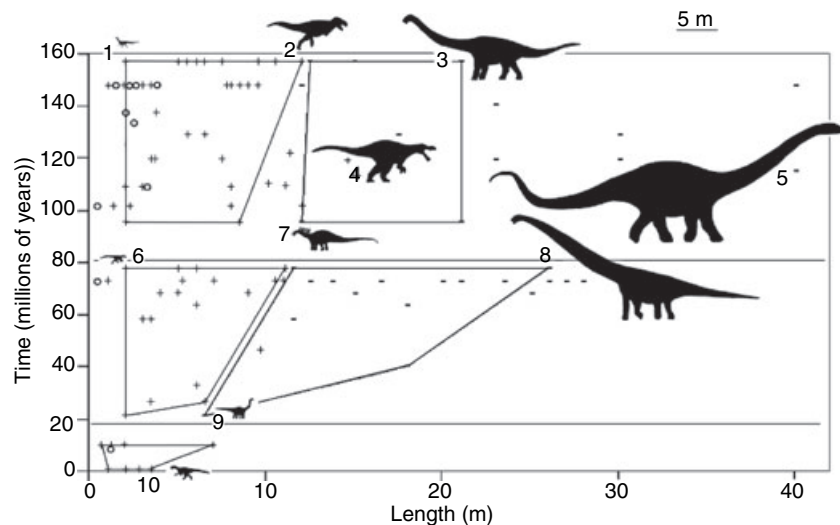
appearance also affects estimates of the rate of size increase. This bias will make little difference to the overall duration inferred for dinosaurs, and so will have negligible impact on the nonphylogenetic rate estimate. By contrast, the bias will tend to truncate the temporal separation between early and late genera within the phylogenetically matched pairs (typically much shorter than the overall duration of dinosaurs), resulting in an overestimate of the rate of size increase. An assessment of the strength of this effect would require more detailed information about sampling than is available, a situation that is common in paleontological analyses.

The Jablonski polygons further support the inference of Cope's rule. While the smallest forms do not increase in size (as would be expected from the strict interpretation of Cope's rule), the largest forms do show a marked increase in four of the five shapes. Only the Cretaceous sauropods do not follow the trend, and these are characterized by their extreme variance of sizes (the largest sauropods were also present in the late Cretaceous, but not late enough to form the borders of the polygon). In the case of the theropods, there were always small genera present, perhaps because there are always ecological niches for small predators. As the theropods were the only carnivorous lineage of dinosaurs, all of the small predatory genera had to arise from them. Notably, as a lineage, the tyrannosaurids began at a large size and got larger, despite being from a narrow stratigraphic range.

Our phylogenetic comparisons indicate that Cope's rule has been in operation, but also hint at a more complex dynamic. Size changes were predicted by the size of the earlier taxon within a comparison: when the earlier taxon was smaller than about 7.8 m long, size tended to increase, whereas the reverse was true for earlier taxa larger than this. Taken at face value, this result implies that dinosaurs longer than about 7.8 m tended to become smaller over time, indicating that this could have been an optimal body size. However, the width of the 95% inverse prediction interval indicates that this result cannot be used to reliably predict whether a lineage of a given size increased or decreased, based on knowledge of its size. Furthermore, this analysis includes a 'regression to the mean' artefact (Alroy, 1998, 2000): unless the data contain no random error, early taxa larger than the average size will tend to be larger than the later taxa with which they are compared, whereas early taxa smaller than average will tend to be smaller than the corresponding later taxa. Correction for this artefact is problematic unless the error variance can be estimated independently (Alroy, 2000). The mean length of the genera in this study is 5.0 m, within the 95% confidence interval (4.9–17.5 m) associated with the estimate of the equilibrium length; regression to the mean therefore cannot be ruled out as the explanation for this relationship. The other analyses of the matched-pairs comparisons are all consistent with a constant trend over time and imprecise size data.

Phylogenetic analyses have the potential to give insights that would not otherwise be available into the evolutionary dynamics underlying trends (Alroy, 2000). However, they also often demand more of the data than do nonphylogenetic approaches (Purvis & Webster, 1999). Imperfections in the data greatly complicate interpretation. Thus, while we can safely conclude that dinosaurs provide an example of Cope's rule, we cannot yet be sure whether it arose from clade selection in tension with individual selection, or from clade selection alone. We cannot yet tell whether an apparent stable equilibrium body size of 7.8 m represents a true attractor or a 'regression to the mean' artefact. The full potential of phylogenetic analysis to shed light on evolutionary dynamics (Alroy, 2000) can be unlocked only with additional information. Detailed information about the sampling of each lineage would permit more reliable estimation of dates of origination and quantification of the associated uncertainty (Marshall, 1998); the latter could be used in sensitivity analyses. The 'regression to the mean' problem will become less severe as the size data improve, making it easier to test the hypothesis of an 'attractor' body size. Only an extremely densely sampled fossil record can permit clean separation of within-lineage from among-lineage change but, even without such an ideal, our analyses show how phylogenetic comparisons give useful insight into evolutionary trends.

## References

Alexander, R.McN. 1985. Mechanics of posture and gait of some large dinosaurs. *Zool. J. Linn. Soc.* **83**: 1–25.

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.

Arnold, A.J., Kelly, D.C. & Parker, W.C. 1995. Causality and Cope's rule – evidence from the planktonic foraminifera. *J. Pal.* **69**: 203–210.

Benton, M.J. 1995. Testing the time axis of phylogenies. *Phil. Trans. R. Soc. Lond. B* **349**: 5–10.

Benton, M.J. 2002. Cope's rule. In: *Encyclopedia of Evolution* (M. Pagel, ed.), pp. 185–186. Oxford University Press, Oxford.

Bonner, J.T. 1968. Size change in development and evolution. In: *Paleobiological Aspects of Growth and Development* (D. B. Macurda, ed.), pp. 1–15. Paleontological Society Memoirs. Supplement to J. Pal.

Brown, D. & Rothery, P. 1993. *Models in Biology: Mathematics, Statistics and Computing*. Wiley, New York.

Burt, A. 1989. Comparative methods using phylogenetically independent contrasts. *Oxf. Surv. Evol. Biol.* **6**: 33–53.

Carroll, S.B. 2001. Chance and necessity: the evolution of morphological complexity and diversity. *Nature* **409**: 1102–1109.

Colbert, E.C. 1962. The weights of dinosaurs. *Am. Mus. Novit.* **2076**: 1–16.

Cope, E.D. 1887. *The Origin of the Fittest*. Appleton Press, New York.

Cope, E.D. 1896. *The Primary Factors of Organic Evolution*. Open Court Publishing Co., New York.

Darwin, C. 1859. *On the Origin of Species by Means of Natural Selection*. John Murray, London.

Fastovsky, D.E. & Weishampel, D.B. 1996. *The Evolution and Extinction of the Dinosaurs*. Cambridge University Press, Cambridge.

Garland, T., Harvey, P.H. & Ives, I.R. 1992. Procedures for the analysis of comparative data using phylogenetically independent contrasts. *Syst. Biol.* **41**: 18–32.

Gould, S.J. 1997. Cope's rule as psychological artefact. *Nature* **385**: 199–200.

Henderson, D.M. 1999. Estimating the masses and centers of mass of extinct animals by 3-D mathematical slicing. *Paleobiology* **25**: 88–106.

Jablonski, D. 1996. Body size and macroevolution. In: *Evolutionary Paleobiology* (D. Jablonski, D. H. Erwin & J. H. Lipps, eds), pp. 256–289. University of Chicago Press, Chicago, IL.

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.

Lewin, R. 1990. Dinosaur count reveals surprisingly few species. *N. Sci.* **128**: 30.

Marshall, C.R. 1998. Determining stratigraphic ranges. In: *The Adequacy of the Fossil Record* (S. K. Donovan & C. R. C. Paul, eds), pp. 23–53. Wiley, London.

McFadden, B.J. 1992. *Fossil Horses: Systematics, Palaeobiology and Evolution of the Family Equidae*. Cambridge University Press, Cambridge, UK.

McKinney, M.L. 1997. Extinction vulnerability and selectivity: combining ecological and paleontological views. *Annu. Rev. Ecol. Syst.* **28**: 495–516.

McShea, D.W. 2000. Trends, tools, and terminology. *Paleobiology* **26**: 330–333.

Paul, G.S. 2000. Is there any evolutionary advantage to gigantism? Did the sauropod dinosaurs continue to grow throughout their lives, as some reptiles and fish do? *Sci. Am.* **284**: 116.

Pisani, D., Yates, A.M., Langer, M.C. & Benton, M.J. 2002. A genus-level supertree of the Dinosauria. *Proc. R. Soc. Lond. B* **269**: 915–921.

Poulin, R. 1995. Evolutionary influences on body-size in free-living and parasitic isopods. *Biol. J. Linn. Soc.* **54**: 231–244.

Purvis, A. & Rambaut, A. 1995. Comparative analysis by independent contrasts (CAIC): an Apple Macintosh application for analysing comparative data. *Comput. Appl. Biosci.* **11**: 247–251.

Purvis, A. & Webster, A.J. 1999. Phylogenetically independent comparisons and primate phylogeny. In: *Comparative Primate Socioecology* (P. C. Lee, ed.), pp. 44–70. Cambridge University Press, Cambridge.

Schmidt-Nielsen, K. 1984. *Scaling: Why is Animal Size So Important*. Cambridge University Press, Cambridge, UK.

Seebacher, F. 1999. A new method to calculate allometric length-mass relationships of dinosaurs. *J. Vert. Paleontol.* **21**: 51–60.

Sereno, P.C. 1997. The origin and evolution of dinosaurs. *Annu. Rev. Earth Planet Sci.* **25**: 435–489.

Stanley, S.M. 1973. An explanation of Cope's rule. *Evolution* **27**: 1–26.

Vrba, E.S. & Gould, S.J. 1986. The hierarchical expansion of sorting and selection: sorting and selection cannot be equated. *Paleobiology* **12**: 217–228.

Webster, A.J. & Purvis, A. 2002. Testing the accuracy of methods for reconstructing ancestral states of continuous characters. *Proc. R. Soc. Lond. B* **214**: 143–149.

Williams, G.C. 1992. *Natural Selection: Domains, Levels, Challenges*. Oxford University Press, Oxford, UK.