Clades reach highest morphological disparity early in their evolution

Martin Hughes, Sylvain Gerber, and Matthew Albion Wills[1]

Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, United Kingdom

There are few putative macroevolutionary trends or rules that withstand scrutiny. Here, we test and verify the purported tendency for animal clades to reach their maximum morphological variety relatively early in their evolutionary histories (early high disparity). We present a meta-analysis of 98 metazoan clades radiating throughout the Phanerozoic. The disparity profiles of groups through time are summarized in terms of their center of gravity (CG), with values above and below 0.50 indicating top- and bottom-heaviness, respectively. Clades that terminate at one of the "big five" mass extinction events tend to have truncated trajectories, with a significantly top-heavy CG distribution overall. The remaining 63 clades show the opposite tendency, with a significantly bottom-heavy mean CG (relatively early high disparity). Resampling tests are used to identify groups with a CG significantly above or below 0.50; clades not terminating at a mass extinction are three times more likely to be significantly bottom-heavy than top-heavy. Overall, there is no clear temporal trend in disparity profile shapes from the Cambrian to the Recent, and early high disparity is the predominant pattern throughout the Phanerozoic. Our results do not allow us to distinguish between ecological and developmental explanations for this phenomenon. To the extent that ecology has a role, however, the paucity of bottom-heavy clades radiating in the immediate wake of mass extinctions suggests that early high disparity more probably results from the evolution of key apomorphies at the base of clades rather than from physical drivers or catastrophic ecospace clearing.

macroevolution | morphological disparity | morphospace | clade shape | clade center of gravity

Evolution is usually characterized as an essentially contingent and unpredictable process (1). This makes it very difficult to identify general rules comparable to those that typify the other natural sciences. Nonetheless, the prospect of formulating and testing macroevolutionary generalities is extremely seductive, because they seem to offer fundamental insights into the manner in which evolutionary processes operate throughout Earth's history (2). Patterns of increasing diversity (measured via proxies of species richness) (3, 4) and increasing maximal organismal size within clades (Cope's rule) (5) have been perennial foci, whereas more recent attention has turned to supposed trends in increasing organismal complexity (6, 7) and the mechanisms that might generate them (8). This paper tests another putative generality, namely, the tendency for taxa to reach maximal morphological diversity (disparity) relatively early in the lifespan of their parent clade (9–17) (early high disparity).

Disparity is conceptually and empirically distinct from diversity. For example, a relatively small sample of species that differ greatly from one another morphologically (e.g., one species from each order of insects) is likely to be more disparate than a much larger sample of species that are morphologically more homogeneous (e.g., a thousand beetles). Among the first questions to be addressed using disparity indices was the perceived magnitude of the Cambrian "explosion." From Charles Darwin (18) onward, evolutionary biologists have been perplexed by the apparently instantaneous first appearances of numerous phyla (a highly disparate sample of species) in the Cambrian fossil record (19). The

subsequent discovery of hitherto unknown fossil groups from the Cambrian Burgess Shale and similar localities added to the enigma, prompting the radical hypothesis that the disparity of metazoans peaked in the Cambrian (14, 20) and subsequent extinctions winnowed this down to much more modest levels soon thereafter. Surprisingly, a relatively small number of studies have tested this hypothesis directly in focal clades (10, 11, 21–23). These predominantly conclude that Cambrian animal groups had a disparity comparable to that of their modern counterparts (24–27). This nonetheless suggests that metazoans reached high levels of disparity relatively early in their history, the phenomenon of early high disparity. Unfortunately, such analyses are limited for two reasons. First, they discount the intervening trajectory of clade evolution. Second, the clade history is truncated both by the present and by a Precambrian fossil record that is enigmatic at best (17, 28). As a result, the focus of disparity studies has increasingly turned to clades that both originate and go extinct within the Phanerozoic (20). Once again, there is a purported tendency for clades to evolve their most disparate forms relatively early in their histories (11–14, 19, 29–31). However, the validity of this early high disparity model has never been tested systematically. If true, it represents a general macroevolutionary "rule" (19) on the broadest possible scale and is comparable to those proposed for increasing morphological complexity (6, 7) and increasing maximal organismal size within clades (2, 32).

Unfortunately, it is impossible to interpret published case studies meta-analytically for several reasons. First, the type of data used is highly variable (outlines, landmarks, and discrete characters), as is the information that these data are intended to convey (shape, form, or homologous characters of the entire organism or of particular organ systems). Second, the manner in which these data have been analyzed is equally variable, although most studies implement some form of data reduction and ordination (10, 12). Species are typically plotted within an empirical, multidimensional space defined by morphological variables (a morphospace) (33). Third, there are many possible indices of morphological disparity, and these are known to describe different aspects of morphospace occupation (34). Fourth, the manner in which trajectories of disparity through time are quantified and classified is also variable. Several of the analyses that originally spurred the debate (10, 21–23, 35) used discrete character matrices to compare anatomically very disparate forms. Many studies have recently followed similar protocols (27, 36–38), and we adopted these methods here as a unifying approach. Where discrete and continuous character data have been compared for the same sets of taxa (39), relative estimates of disparity have been similar.

We collated morphological and stratigraphic data for 98 extinct and relict clades to answer three questions. (*i*) Is early high disparity the dominant pattern of clade evolution across the Metazoa and throughout the Phanerozoic? (*ii*) Is there a trend in clade disparity profile shape throughout the Phanerozoic? (*iii*) Do clades terminating at times of mass extinction have disparity profile shapes distinguishable from clades becoming extinct at other times? We addressed all three questions using the clade center of gravity (CG) index (31, 40, 41). This quantifies overall clade shape in a robust manner and has previously been applied to paleontological diversity and disparity data (*Materials and Methods* and Dataset S1) (Fig. 1). Values of CG <0.5 denote bottom-heavy clades, whereas CG >0.5 indicates top-heaviness. We considered extinct and some relict clades in our sample because clades with extant lineages are still evolving and may be at (or still approaching) their maximum disparity. Extant clades are more likely to have "flat-topped" disparity profiles, which will artifactually shift their CG upward (Fig. 1) relative to that which may have pertained for the (hypothetical) entire clade history (Dataset S1). However, clades terminating at a mass extinction event might be similarly truncated and are likely to have higher CGs for similar reasons. Mass extinctions have undoubtedly influenced the manner in which clades have explored morphospaces (42), but this phenomenon received little attention until recently (37, 43–47). Moreover, only one of these studies (44) focused on extinction selectivity per se; all others investigated the subsequent evolution of extinction survivors. Here, we determined whether the clades going extinct coincident with one of the "big five" mass extinction events [End Ordovician, Late Devonian (Fransian/Famenian), End Permian, End Triassic, and End Cretaceous] had disparity profiles distinguishable from those terminating at other times.

Results and Discussion

For diversity through time, random birth/death models with constant parameters predict that the average clade shape should be symmetrical (31, 48). However, for disparity, the predictions are less precise. New species can only arise from the fission of existing ones (clades initially diversify from a single species and therefore a single point in morphospace), whereas extinctions can be random with respect to this same tree (34). Therefore, if a clade follows a homogeneous birth–death model with characters evolving in a Brownian fashion, some top-heaviness would be expected (41). Our use of 0.5 as a null is slightly simplistic, therefore, but biased against our principal finding (namely that clades not terminating at a mass extinction event are bottom-heavy on average).

Across our sample of 98 clades (including those terminating coincident with a mass extinction), we found a mean disparity profile of CG of 0.495, with a median CG (0.501) indistinguishable from 0.5 (V = 2,429, P = 0.992). Time-averaged indices masked some apparent differences in clade disparity profiles within and between eras; most notably, there were more bottom-heavy (CG <0.5) clades in the Late Paleozoic than top-heavy (CG >0.5) clades, with the opposite pattern in the Mesozoic. However, comparison across four time bins [Early Paleozoic (Cambrian/Ordovician), Late Paleozoic, Mesozoic, and Cenozoic] revealed no significant differences in the frequencies (log likelihood ratio test; G = 2.298, P = 0.513). We then implemented a bootstrapping test (Dataset S1) for significant deviation from clade symmetry, allowing us to partition clades into three groups: significantly bottom-heavy, significantly top-heavy, and indistinguishable from symmetrical (which we discounted). Again, there were no significant differences in the relative frequencies of significantly top- and bottom-heavy clades across the four time bins (G = 3.558, P = 0.313). Finally, a plot of clade CG against the time of clade origin revealed no systemic trends throughout the Phanerozoic (Fig. 2).

Although clade disparity profiles had a mean CG indistinguishable from 0.5, there was a marked and significant difference in CG between those clades terminating coincident with a mass extinction event and those becoming extinct at other times (Fig. 3). The latter group had significantly bottom-heavy disparity profiles on average (63 clades with a mean CG significantly less than 0.500; t = −2.420, P = 0.018). By contrast, the 35 clades ending at mass extinctions had a mean CG significantly greater than 0.500 (t = 3.901, P < 0.001). Likelihood ratio tests also confirmed that
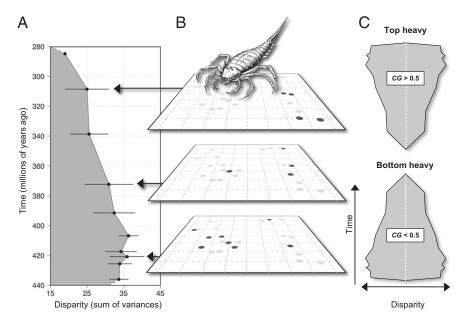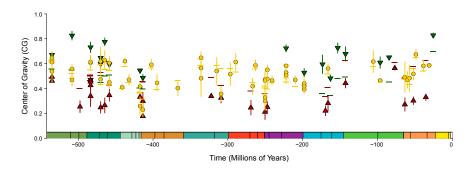

Fig. 1. Calculating the disparity profile of clades. (*A*) Disparity of Stylonurina (74) measured as the sum of variances on successive principal coordinate analyses at several time intervals. Mean of 1,000 bootstrap replicates ± SE. (*B*) Distribution of taxa on the first two principal coordinates of their empirical morphospace at three of the time intervals. Black symbols indicate taxa present in the interval; gray symbols indicate taxa present in other intervals. (*C*) Stylized representations of significantly top-heavy (*Upper*) and bottom-heavy (*Lower*) asymmetrical clade disparity profiles.

Fig. 2. Center of gravity (CG_{scaled}) values for all 98 datasets across the Phanerozoic. Case studies are sampled relatively evenly throughout this time, and there is no systemic temporal trend in disparity profile shape. Circles denote mean scaled CG (CG_{scaled}) from 1,000 bootstrap replicates of the variance-based disparity curves for each clade, plotted against the clade origination date. Vertical lines denote the SE around CG_{scaled}, derived from 1,000 bootstrap replicates. Green triangles, significantly top-heavy profiles ($CG_{scaled} > CG_i$ with $P < 0.05$); red triangles, significantly bottom-heavy profiles ($CG_{scaled} < CG_i$ with $P < 0.05$); yellow circles, profile indistinguishable from symmetrical; abscissa color scheme, International Stratigraphic Chart.
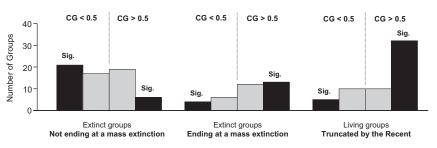
the relative frequencies of top- and bottom-heavy clades terminating at mass extinctions and at other times were different, whether including all clades ($G = 7.648$, $P = 0.006$) or only those with significant skew ($G = 13.022$, $P < 0.001$). For comparison, we also generated disparity profiles for 53 additional living clades with high diversity in the Recent (Dataset S1) (these are otherwise excluded from our sample unless stated). These extant clades (truncated by the present) had a median CG significantly greater than 0.500 ($V = 1,150$, $P < 0.001$) but indistinguishable from that for fossil clades terminating at a mass extinction ($W = 924$, $P = 0.980$).

Over half of our study clades had disparity profiles that were neither significantly top- nor bottom-heavy. However, these "symmetrical" clades may nonetheless have a variety of trajectories, with their own particular macroevolutionary implications. Most remarkable are groups [e.g., crinoids (35)] whose earliest exemplars have levels of disparity that are not significantly different from the maximum levels subsequently achieved by the clade; a simplistic null of early maximal disparity. For 29 of the 54 symmetrical groups, we were unable to reject this null. Such a pattern would be close to that often envisaged for explosive radiations (14, 49) and similar to that proposed as the trajectory for metazoans through the Phanerozoic (11). Early high patterns inevitably imply an unsampled period of cladogenesis (or the existence of ghost ranges) at the base of the clade, but this either occurs too fast for the available stratigraphic resolution or is not fossilized (11, 50). Late saturation is much less remarkable, because clades have already undergone radiation and diversification and had almost the entirety of their histories in which to colonize the extremities of their morphospaces. Although late saturation was observed in 32 symmetrical clades, 12 of these also ended at a mass extinction (and were

therefore likely to have been prematurely truncated). For this reason, we again focused on the 63 "free-evolving" clades that did not terminate at a mass extinction. Of these, the proportion (two-thirds) that were either significantly bottom-heavy or showed early saturation (two mutually compatible conceptions of early high disparity) was significantly greater than the proportion that were either significantly top-heavy or showed late saturation (late high disparity) (two-sample test for equality of proportions; $\chi^2 = 4.613$, $P = 0.016$). Therefore, clades that do not terminate at a mass extinction do indeed tend to reach their highest levels of disparity relatively early in their evolutionary histories (20). Moreover, this tendency occurs throughout the Phanerozoic.

Why Do Clades Have Early High Disparity?

What might explain the prevailing pattern of early high disparity in clade evolution (19, 51)? Both ecological and developmental explanations have been proposed, and our results remain consistent with both. The "empty ecospace" model predicts that clades will radiate and diversify more rapidly when colonizing a new environment. This colonization may occur because ecospace has been vacated by other occupants (e.g., in the wake of some other extinction, typically the result of external, physical factors) or because a hitherto inaccessible environment or other resource has been rendered viable by the acquisition of some novel, "key" adaptation (52–54) or series of characters (55) (an intrinsic, biological trigger). Morphological change under these circumstances may be rapid either because transitions are unusually large or because rates of cladogenesis are unusually high (even with "normal" step sizes at each splitting event) (29). In this context, we also note that major clades are often distinguished from their paraphyletic

Fig. 3. Groups terminating at one of the "big five" mass extinction events (and living groups that are still diversifying) are more top-heavy than those terminating at other times. (*Left*) Disparity profile frequencies for extinct clades that do not terminate at a mass extinction boundary. (*Center*) Disparity profile frequencies for extinct clades that terminate at a mass extinction boundary. (*Right*) Disparity profile frequencies for living clades (truncated by the Recent). Bars to the left and right of the dotted lines indicate the frequencies of bottom-heavy (CG <0.5) and top-heavy (CG >0.5) clades, respectively. Black bars indicate the frequencies of significantly bottom- or top-heavy clades ($P < 0.05$), while gray bars indicate the frequencies of clades for which $P \geq 0.05$. Mass extinctions: Late Ordovician, 443.7 Ma; Late Devonian, 374.5 Ma; Late Permian, 251 Ma; Late Triassic, 199.6 Ma; and Late Cretaceous, 65.5 Ma.

progenitors because they possess distinct and defining sets of derived characters, or because an extant crown is defined relative to an extinct stem. These divisions into a clade and its residual paraphylum would otherwise often be arbitrary. For example, rather than delimiting a clade of Aves from within the para-phyletic nonavian dinosaurs, it would be possible to define a clade of Aves plus some arbitrary "depth" of theropod dinosaurs. However, birds are defined in the manner they are because they acquired a distinctive suite of apomorphies pertaining to the evolution of flight; key innovations, in this case, that also enabled them to exploit a new environment. These shifts in anatomy, physiology, behavior, and ecology may themselves explain the differential survival of crowns and stems.

More generally it is likely that global shifts in climate, sea level, and ocean chemistry [coupled with the elevated rates of extinction and turnover that these phenomena engender (56–58)] affected the availability of ecospace throughout the Phanerozoic. The only temporal pattern in disparity profile shapes detected in our data was the significant tendency toward top-heaviness in those clades terminating coincident with a mass extinction [which predominantly result from physical drivers (59)]. However, mass extinctions need not increase the subsequent availability of ecospace but may actually cause its collapse (60). The absence of any systemic trends in clade disparity patterns through time, or of any increased propensity for early high disparity in clades radiating in the immediate wake of mass extinctions (Dataset S1) suggest that if ecological mechanisms have a role, then this is more likely to be mediated via key innovations (which can evolve at any time) and the opening up of new adaptive zones rather than from ecospace clearing.

We stress that ecological and developmental explanations for early high disparity are not mutually exclusive; neither do our results allow us to distinguish between them. The hypothesis of increasing developmental constraint predicts that the increasing complexity and interdependence of ontogenetic processes with evolutionary time effectively lock down the potential for subsequent morphological innovation (14, 61–65). Such mechanisms purportedly explain why bodyplans become invariant and inflexible with time, although mechanisms by which these constraints may be lifted have been posited (66). Notable examples are the tetrapod pentadactyl limb [early tetrapods explored a range of higher digit numbers (67)], the seven cervical vertebrae of all mammals except sloths and manatees [otherwise invariant from mice to giraffes (68)] and the diagnostic head segmentation of arthropod subphyla [Cambrian genera explored numerous alternatives with relative freedom (14, 69)]. Such body patterning characters are usually controlled by *Hox* (homeobox) genes, which are also frequently exapted for other (often functionally and positionally unrelated) developmental roles (70). This increasing pleiotropy (more and more varied roles for the same regulatory genes) may account for the observed reduction of developmental lability. Testing this hypothesis would require detailed ontogenetic data far beyond the scope of this study.

The prevalence of early high disparity as the dominant pattern of clade evolution ranks alongside the well-known tendencies for increasing complexity (7, 8, 71, 72) and diversity (2, 8) underpinning putative macroevolutionary trends of the widest possible generality. Moreover, it seems to apply throughout the Phanerozoic, and not merely at times of global diversification (e.g., the early Paleozoic).

Materials and Methods

Collation of Data. We compiled published discrete morphological and stratigraphic data for 98 vertebrate and invertebrate clades radiating throughout the Phanerozoic (Dataset S2). For a subset of analyses (where expressly

stated), we also compiled morphological and stratigraphic data for an additional 53 extant clades (Dataset S2). We avoided taxonomically overlapping cases or datasets obviously derivative of others. Individual datasets were sampled at a variety of taxonomic levels, although most were familial and ordinal in their coverage. Within datasets, strict rules were applied to ensure that sampling was adequately uniform across known operational taxonomic units (OTUs) and through time, amalgamating taxa where necessary (Dataset S1).

Analyses. All analyses were conducted in R using our own scripts (Dataset S3). Empirical morphospaces were derived as multidimensional spaces in which the proximity of OTUs correlated with their morphological similarity (10, 21). Disparity was measured using the sum of variances on successive axes of the morphospace (10, 22, 73). To derive a trajectory of disparity through time, we divided the duration of the clade into time bins, defined so as to balance the competing requirements of stratigraphic resolution and sample size (73) (Fig. 1). To provide a single index of the shape of clade disparity profiles, we calculated the CG metric previously applied to paleontological diversity and disparity data (31, 40, 41). The CG in absolute time (CG_m) was given by

$$CG_m = \Sigma \, d_i t_i / \Sigma \, d_i,$$

where d_i is the disparity at the ith stratigraphic interval and t_i is the temporal midpoint in absolute time (millions of years) of the ith stratigraphic interval. We then scaled this value between the ages of the oldest (t_{oldest}) and youngest ($t_{youngest}$) representatives of the clade to yield a scaled index of observed CG (CG_{scaled}) between 0 and 1:

$$CG_{scaled} = \frac{t_{oldest} - CG_m}{t_{oldest} - t_{youngest}}.$$

If time bins were all of the same duration, then clades with uniform or symmetrical disparity profiles would have CG_{scaled} of 0.50 (midway). Clades with a relatively early disparity maximum (bottom-heavy) would have CG_{scaled} <0.50, whereas those with a late disparity maximum (top-heavy) would have CG_{scaled} >0.5. In practice the expected CG_{scaled} for a clade of constant disparity through time is not necessarily 0.50, but rather is determined by the durations of the time bins over which the profile was measured. This is because stratigraphic stages are of variable durations, and because taxa are not always dated to series and stages. Hence, we compared CG_{scaled} with the inherent CG_{scaled} (CG_i) for a hypothetical clade of uniform disparity spanning the same intervals. A bootstrapping test determined when this deviation was significant [clades for which >97.5% of 1,000 bootstrapped replicates lay either above or below the center of gravity inherent in the time scale (P value <0.05) (41). Finally, we adjusted the observed scaled CG_{scaled} relative to CG_i as a zero baseline, hereafter simply CG. Clades were then partitioned into one of three categories according to CG: significantly bottom-heavy, significantly top-heavy, or indistinguishable from symmetrical. Log likelihood ratio goodness-of-fit tests (G-tests) were used to compare frequencies of different profile shapes (e.g., in different time bins).

Clades that were not significantly top- or bottom-heavy could nonetheless have a variety of profile shapes. We therefore devised an ancillary test to determine whether the taxa observed at the beginning and end of the history of each clade (those in the first and last time bins) had a disparity that could be distinguished from the maximum observed in any time bin. The disparity profile of the clade was resampled using 1,000 bootstraps of all of the OTUs in the dataset. For each replicate curve, the difference in disparity between the first (or last) intervals and the disparity maximum elsewhere in the curve was calculated, yielding a distribution. If a difference of zero was within the 95% limits of this distribution, we were unable to reject the null hypothesis of no difference between the initial disparity and the maximum achieved by the clade.

1. Ghiselin MT (2005) The Darwinian revolution as viewed by a philosophical biologist. *J Hist Biol* 38(1):123–136.

2. McShea DW (1998) Possible largest-scale trends in organismal evolution: Eight "live hypotheses." *Annu Rev Ecol Syst* 29:293–318.

3. Benton MJ, Emerson BC (2007) How did life become so diverse? The dynamics of diversification according to the fossil record and molecular phylogenetics. *Palaeontology* 50(1):23–40.

4. Alroy J (2010) Geographical, environmental and intrinsic biotic controls on Phanerozoic marine diversification. *Palaeontology* 53(6):1211–1235.

5. Clauset A, Erwin DH (2008) The evolution and distribution of species body size. *Science* 321(5887):399–401.

6. Arthur W (2007) *Creatures of Accident: The Rise of the Animal Kingdom* (Hill & Wang, New York).

7. Adamowicz SJ, Purvis A, Wills MA (2008) Increasing morphological complexity in multiple parallel lineages of the Crustacea. *Proc Natl Acad Sci USA* 105(12):4786–4791.

8. McShea DW, Brandon RN (2010) *Biology's First Law: The Tendency for Diversity and Complexity to Increase in Evolutionary Systems* (Univ of Chicago Press, Chicago).

9. Ruta M, Wagner PJ, Coates MI (2006) Evolutionary patterns in early tetrapods. I. Rapid initial diversification followed by decrease in rates of character change. *Proc Biol Sci* 273(1598):2107–2111.

10. Wills MA, Briggs DEG, Fortey RA (1994) Disparity as an evolutionary index: A comparison of Cambrian and Recent arthropods. *Paleobiology* 20(2):90–130.

11. Wills MA, Fortey RA (2000) The shape of life: How much is written in stone? *Bioessays* 22(12):1142–1152.

12. Erwin DH (2007) Disparity: Morphological pattern and developmental context. *Palaeontology* 50(1):57–73.

13. Foote M (1997) The evolution of morphological diversity. *Annu Rev Ecol Syst* 28:129–152.

14. Gould SJ (1989) *Wonderful Life* (W. W. Norton, New York).

15. Wagner PJ (1997) Patterns of morphologic diversification among the Rostroconchia. *Paleobiology* 23(1):115–150.

16. Foote M (1993) Discordance and concordance between morphological and taxonomic diversity. *Paleobiology* 19(2):185–204.

17. Valentine JW, Erwin DH, Jablonski D (1996) Developmental evolution of metazoan bodyplans: the fossil evidence. *Dev Biol* 173(2):373–381.

18. Darwin C (1859) *On the Origin of Species by Natural Selection* (Murray, London).

19. Wagner PJ (2010) *Evolution Since Darwin: The First 150 Years*, eds Bell M, Futuyma D, Eanes W, Levinton J (Sinauer, Sunderland, MA), pp 451–478.

20. Gould SJ (1991) The disparity of the Burgess Shale arthropod fauna and the limits of cladistic analysis: Why we must strive to quantify morphospace. *Paleobiology* 17(4):411–423.

21. Briggs DEG, Fortey RA, Wills MA (1992) Morphological disparity in the Cambrian. *Science* 256(5064):1670–1673.

22. Foote M (1992) Paleozoic record of morphologica diversity in blastozoan echinoderms. *Proc Natl Acad Sci USA* 89(16):7325–7329.

23. Wills MA (1998) Cambrian and Recent disparity: The picture from priapulids. *Paleobiology* 24(2):177–199.

24. Valentine JW, Jablonski D, Erwin DH (1999) Fossils, molecules and embryos: New perspectives on the Cambrian explosion. *Development* 126(5):851–859.

25. Thomas RDK, Shearman RM, Stewart GW (2000) Evolutionary exploitation of design options by the first animals with hard skeletons. *Science* 288(5469):1239–1242.

26. Knoll AH, Carroll SB (1999) Early animal evolution: Emerging views from comparative biology and geology. *Science* 284(5423):2129–2137.

27. Shen B, Dong L, Xiao S, Kowalewski M (2008) The Avalon explosion: Evolution of Ediacara morphospace. *Science* 319(5859):81–84.

28. Waggoner B, Collins AG (2004) Reductio ad absurdum: Testing the evolutionary relationships of Ediacaran and Paleozoic problematic fossils using molecular divergence dates. *J Paleontol* 78(1):51–61.

29. Foote M (1996) *Evolutionary Paleobiology*, eds Jablonski D, Erwin DH, Lipps JH (Univ of Chicago Press, Chicago), pp 62–86.

30. Erwin DH (2011) Evolutionary uniformitarianism. *Dev Biol* 357(1):27–34.

31. Gould SJ, Gilinsky NL, German RZ (1987) Asymmetry of lineages and the direction of evolutionary time. *Science* 236(4807):1437–1441.

32. Hone DWE, Benton MJ (2005) The evolution of large size: How does Cope's Rule work? *Trends Ecol Evol* 20(1):4–6.

33. Wills MA (2001) *Fossils, Phylogeny and Form: An Analytics Approach*, eds Adrain JM, Edgecombe GD, Lieberman BS (Kluwer, Dordrecht, The Netherlands), pp 402.

34. Ciampaglio CN, Kemp M, McShea DW (2001) Detecting changes in morphospace occupation patterns in the fossil record: Characterization and analysis of measures of disparity. *Paleobiology* 27(4):695–715.

35. Foote M (1994) Morphological disparity in Ordovician-Devonian crinoids and the early saturation of morphological space. *Paleobiology* 20(3):320–344.

36. Brusatte SL, Benton MJ, Ruta M, Lloyd GT (2008) The first 50Myr of dinosaur evolution: Macroevolutionary pattern and morphological disparity. *Biol Lett* 4(6):733–736.

37. Bapst DW, Bullock PC, Melchin MJ, Sheets HD, Mitchell CE (2012) Graptoloid diversity and disparity became decoupled during the Ordovician mass extinction. *Proc Natl Acad Sci USA* 109(9):3428–3433.

38. Cisneros JC, Ruta M (2010) Morphological diversity and biogeography of procolophonids (Amniota: Parareptilia). *J Syst Palaeontology* 8(4):607–625.

39. Foth C, Brusatte SL, Butler RJ (2012) Do different disparity proxies converge on a common signal? Insights from the cranial morphometrics and evolutionary history of Pterosauria (Diapsida: Archosauria). *J Evol Biol* 25(5):904–915.

40. Uhen MD (1996) An evaluation of clade-shape statistics using simulations and extinct families of mammals. *Paleobiology* 22(1):8–22.

41. Foote M (1991) Morphological and taxonomic diversity in clade's history: The blastoid record and stochastic simulations. *Contrib Mus Paleontol Univ Mich* 28(6):101–140.

42. Jablonski D (2005) Mass extinctions and macroevolution. *Paleobiology* 31(2):192–210.

43. Friedman M (2010) Explosive morphological diversification of spiny-finned teleost fishes in the aftermath of the end-Cretaceous extinction. *Proc Biol Sci* 277(1688):1675–1683.

44. Brusatte SL, Butler RJ, Prieto-Márquez A, Norell MA (2012) Dinosaur morphological diversity and the end-Cretaceous extinction. *Nat Commun* 3(804):1–8.

45. Ausich WI, Deline B (2012) Macroevolutionary transition in crinoids following the Late Ordovician extinction event (Ordovician to Early Silurian). *Palaeogeogr Palaeoclimatol Palaeoecol* 361(15):38–48.

46. Girard C, Renaud S (2012) Disparity changes in 370 Ma Devonian fossils: The signature of ecological dynamics? *PLoS ONE* 7(4):e36230.

47. Thorne PM, Ruta M, Benton MJ (2011) Resetting the evolution of marine reptiles at the Triassic-Jurassic boundary. *Proc Natl Acad Sci USA* 108(20):8339–8344.

48. Kitchell JA, MacLeod N (1989) Asymmetries of clade shape and the direction of evolutionary time: Response. *Science* 243(4898):1614–1615.

49. Gavrilets S, Losos JB (2009) Adaptive radiation: Contrasting theory with data. *Science* 323(5915):732–737.

50. Fortey RA, Briggs DEG, Wills MA (1996) The Cambrian evolutionary 'explosion': Decoupling cladogenesis from morphological disparity. *Biol J Linn Soc Lond* 57(1):13–33.

51. Erwin D (1994) Early introduction of major morphological innovations. *Acta Palaeontol Pol* 38(3):281–294.

52. Simpson GG (1944) *Tempo and Mode in Evolution* (Columbia Univ Press, New York).

53. Etienne RS, Haegeman B (2012) A conceptual and statistical framework for adaptive radiations with a key role for diversity dependence. *Am Nat* 180(4):E75–E89.

54. Dumont ER, et al. (2012) Morphological innovation, diversification and invasion of a new adaptive zone. *Proc Biol Sci* 279(1734):1797–1805.

55. Donoghue MJ (2005) Key innovations, convergence, and success: Macroevolutionary lessons from plant phylogeny. *Paleobiology* 31(2):77–93.

56. Hannisdal B, Peters SE (2011) Phanerozoic Earth system evolution and marine biodiversity. *Science* 334(6059):1121–1124.

57. Mayhew PJ, Bell MA, Benton TG, McGowan AJ (2012) Biodiversity tracks temperature over time. *Proc Natl Acad Sci USA* 109(38):15141–15145.

58. Benton MJ (2009) The Red Queen and the Court Jester: Species diversity and the role of biotic and abiotic factors through time. *Science* 323(5915):728–732.

59. Peters SE (2008) Environmental determinants of extinction selectivity in the fossil record. *Nature* 454(7204):626–629.

60. Erwin DH (2001) Lessons from the past: Biotic recoveries from mass extinctions. *Proc Natl Acad Sci USA* 98(10):5399–5403.

61. Arthur W (2004) The effect of development on the direction of evolution: Toward a twenty-first century consensus. *Evol Dev* 6(4):282–288.

62. Arthur W, Farrow M (1999) The pattern of variation in centipede segment number as an example of developmental constraint in evolution. *J Theor Biol* 200(2):183–191.

63. Smith M, et al. (1985) Developmental constraints and evolution: A perspective from the mountain lake conference on development and evolution. *Q Rev Biol* 60(3):265–287.

64. Kauffman SA (1985) New questions in genetics and evolution. *Cladistics* 1(3):247–265.

65. Wimsatt WA (1986) *Integrating Scientific Disciplines*, ed Bechtel W (Springer, New York), pp 185–208.

66. Wagner A (2011) Genotype networks shed light on evolutionary constraints. *Trends Ecol Evol* 26(11):577–584.

67. Clack JA (2006) The emergence of early tetrapods. *Palaeogeogr Palaeoclimatol Palaeoecol* 232(2-4):167–189.

68. Galis F (1999) Why do almost all mammals have seven cervical vertebrae? Developmental constraints, *Hox* genes, and cancer. *J Exp Zool* 285(1):19–26.

69. Wills MA, Briggs DEG, Fortey RA, Wilkinson M (1995) The significance of fossils in understanding arthropod evolution. *Verh Dtsch Zool Ges* 88(2):203–215.

70. Chipman AD (2001) Developmental exaptation and evolutionary change. *Evol Dev* 3(5):299–301.

71. McShea DW (1996) Metazoan complexity and evolution: Is there a trend? Perspective. *Evolution* 50(2):477–492.

72. Valentine JW, Collins AG, Meyer CP (1994) Morphological complexity increase in metazoans. *Paleobiology* 20(2):131–142.

73. Foote M (1992) Rarefaction analysis of morphological and taxonomic diversity. *Paleobiology* 18(1):1–16.

74. Lamsdell JC, Braddy SJ, Tetlie OE (2010) The systematics and phylogeny of the Stylonurina (Arthropoda: Chelicerata: Eurypterida). *J Syst Palaeontology* 8(1):49–61.

75. Clarkson ENK (1998) *Invertebrate Palaeontology & Evolution* (Wiley, New York).

EVOLUTION

EARTH, ATMOSPHERIC, AND PLANETARY SCIENCES