ELECTRONIC SUPPLEMENTARY MATERIAL

Decoupling of morphological disparity and taxic diversity during the adaptive radiation of anomodont therapsids

Marcello Ruta[1], Kenneth D. Angielczyk[2], Jörg Fröbisch[3] and Michael J. Benton[4]

[1] *School of Life Sciences, University of Lincoln, Lincoln LN2 2LG, UK*

[2] *Department of Geology, Field Museum of Natural History, Chicago 60605-2496, USA*

[3] *Museum für Naturkunde, Leibniz Institute for Research on Evolution and Biodiversity, Berlin 10115, Germany*

[4] *School of Earth Sciences, University of Bristol, Bristol BS8 1RJ, UK*

SUPPLEMENTARY MATERIAL AND METHODS

(**a**) *Anomodont phylogeny*

As one of the most successful radiations of therapsid (stem-group mammal) amniotes, anomodonts provide an excellent case study for exploring macroevolutionary patterns in terrestrial vertebrates during a major extinction event. There has been considerable interest in the phylogeny of this group over the past decade [1-10]. The analysis of anomodont interrelationships from [11], rooted on the genus *Biseridens* [9], offers the most up-to-date account of the published characters, their codings and formulation, and the rationale behind character-state delimitations. Our work represents the first example of a disparity analysis applied to a cladistic data matrix that includes both continuous and discrete characters.

(**b**) *Anomodont stratigraphy*

The stratigraphic framework used in this study is based on J. F.'s [12] review of anomodont-bearing tetrapod faunas, which used the well-established Permian-Triassic assemblage zones (AZ) from the South African Karoo Basin [13], the international marine stages of the current Geologic Time Scale (Standard Global Chronostratigraphic Scale; SGCS) [14], as well as the Permian and Triassic land-vertebrate faunachrons (LVFs; [15-21]). Changes to absolute ages of stage boundaries, specifically in the Triassic but also with respect to the Permian stages and to the chronological delimitation of the Permo-Triassic boundary (PTB), adopted from [22-24]), were already implemented in the study in [12] and are also employed here. In general, correlations of the Permian faunas primarily follow compendia in [18,25], whereas correlations of the Triassic faunas are chiefly based on Battail [26] and Lucas [15,26-28]. The rationale and details of the correlations of these faunas are discussed in detail in [12] and will not be repeated here. However, the present study further considers the most recent advances in the correlation schemes for southern and eastern Africa, namely those for the Upper Permian Upper Madumabisa Mudstone (best known from the Luangwa Basin in Zambia)

and the Tanzanian Usili Formation (Ruhuhu Basin), which are now both considered to be best correlated with the *Cistecephalus* Assemblage Zone of the South African Karoo Basin [29].

(c) *Disparity analyses*

We adopted standard protocols [30-32] to calculate character-based disparity (i.e., the amount of morphological dissimilarity measured between or within groups of taxa) from the data matrix in [11]. These were discussed in detail, implemented, and refined methodologically in [32-37], and have been applied to numerous fossil vertebrate groups in recent years [38-47]. Although detailed discussions of these protocols are available and have been extensively dealt with in various papers [32,38,40], we consider it necessary to provide a very short summary and add some clarifications. This is because character-based disparity analyses are less well known than those based on either traditional or geometric morphometric methods. Some issues, such as inclusion of autapomorphies and their impact on profiles of disparity and patterns of morphospace occupation, were addressed briefly in [40,41,46] (but see also below). Others, such as the geometric interpretation of the results, were discussed in [32,40]. However, we review both issues concisely here.

We consider character-based analyses of disparity in the context of measuring phenotypic dissimilarity [48-50]. Although much research on disparity focuses on biological shape and shape differences among organisms and/or their parts [51,52], we emphasize that there is no unique way to measure the morphological distinctiveness of organisms. Several researchers advocate using techniques for analyzing shape variation that rely solely on geometric properties of the objects (taxa or their constituent parts) being measured [49,53-57]. These techniques are employed in traditional (e.g., based on linear and other continuous measurements) and geometric (e.g., based on landmarks) morphometrics [58-63]. These methods permit a relatively intuitive representation of specimens or taxa in a multivariate morphospace in which the distances between specimens or taxa correspond to the magnitude of their shape differences. Geometric morphometrics has the additional advantage of a graphical means of displaying the nature of the shape differences between specimens or taxa using

63 thin plate spline deformations. However, in those cases where physical variables and/or landmarks
64 cannot be measured or determined across the whole spectrum of organisms that are being compared
65 (e.g., due to great variability, highly divergent body plans, or ambiguous primary homologies of the
66 landmarks), or where the variation among the organisms of interests transcends differences in shape
67 and includes other types of anatomical features, character-based analyses disparity are appropriate
68 [32,34-38,40,46]. Of course, such analyses are also suitable in those cases in which organisms are
69 directly comparable and statements of homology are clear, as in the anomodont example that is the
70 basis of the present study.
71 Character-based analyses of disparity rely on multivariate treatments of differences among
72 discrete and/or continuous traits, and these are codified in the form of states in phylogenetic data
73 matrices. The matrices may additionally encompass phenetic, ecological, environmental, and/or
74 functional data. When applied to morphological characters in a cladistic or phenetic framework, this
75 category of disparity methods provides a less intuitive treatment of morphological variation than
76 those offered by traditional or geometric morphometric methods. This is because the dissimilarity
77 among taxa is not measured using descriptors of physical shape (e.g., lengths, widths, angles, ratios,
78 positions of landmarks), but is instead quantified in terms of the amount of character-state or trait
79 variation that exists [32,40]. Therefore, the differences among taxa cannot be easily mapped back
80 on to the organisms as can be done for variations in linear dimensions or changes in overall shape.
81 Instead, they represent a summary of differences and these could take a variety of forms, ranging
82 from the presence or absence of, say, certain bones or organs, details of the contacts among bones,
83 features of the construction of complex structures, or the shape differences of those structures.
84 The characters used to build a phylogeny or a dendrogram can also be employed to construct
85 a multidimensional morphospace in which the distances between taxa reflect the magnitude of their
86 morphological differences. Again, it is important to remember that these distances are based on the
87 number of differences in the codings of characters that can cover a range of anatomical structures.
88 A given taxon may be located one unit away from a second taxon on account of the absence of a

89 bone, whereas it may be located one unit away from a third taxon on account of the combination of
90 bones that form its secondary palate. Therefore, although the distances among taxa in a character-
91 based morphospace reflect the fact that the taxa differ in a given number of character and/or trait
92 conditions, their positions in morphospace do not provide any direct and explicit information about
93 which characters differ between them.

94 Character-based analyses of disparity are not intrinsically superior/inferior to, nor are they a
95 substitute for, traditional or geometric morphometric studies; they are merely different. By focusing
96 on the number of differences among taxa, they can accommodate a greater variety of morphological
97 descriptors - cladistic, phenetic, functional, ecological, or otherwise [32,40,46] - in a simultaneous
98 analysis. They can further accommodate taxa that encmpass widely divergent body plans and few
99 homologous landmarks. This flexibility comes at the cost of the difficulty of relating the distances
100 among taxa in morphospace to specific anatomical and/or shape differences, thus limiting to some
101 extent the insight they provide into topics such as convergent evolution. Character-based analyses
102 of disparity have been extensively used with cladistic data matrices of discrete traits (or, at least,
103 matrices that include a large proportion of such traits), but their utility is by no means restricted to
104 alphanumeric data sets. In fact, the data sets can accommodate a mixture of traits, e.g., discrete
105 (binary and/or multistate) and continuous, as well as different options for ordering and weighting.
106 Many such options are implemented in the versatile free software MATRIX v. 1.0 [32,38-45,47].
107
108 (**d**) ***Pair-wise taxon distances***
109 We built a tabulation of inter-taxon generalized Euclidean distances (Dataset S1) in MATRIX v. 1.0,
110 treating all characters as unordered and having equal unit weight, but without replacing instances of
111 polymorphic codings with known states (i.e., polymorphisms were treated as such). This practice
112 contrasts with those in other studies [40], where uncertain and polymorphic codings were replaced
113 with one of the alternative states that are bracketed under uncertainty or polymorphism (e.g., using
114 consistently the lowest or the highest of the bracketed states in each case). Our solution is thus more

115 conservative than that of previous analyses, though in many of the case studies published thus far,
116 the main results from disparity and morphospace analyses are not altered to a significant degree
117 when alternative codings are introduced as a replacement for uncertainties and polymorphisms. For
118 the 143 discrete characters, there are 22 instances of polymorphic codings (there are no uncertainties
119 in the data matrix), representing 0.177% of the total number of cell entries (12,441) for the discrete
120 characters.

121 When missing entries (i.e., unknown, inapplicable states, as well as question marks replacing
122 uncertainties or polymorphisms) occur in a pair-wise comparison between any two taxa, MATRIX
123 assigns a mean weighted value to the character comparisons for which one or both taxa exhibit the
124 missing entries. It first calculates a weighted mean fractional similarity (WMFS), namely the ratio
125 between the total distance for all the scorable comparisons between any two taxa (i.e., the number of
126 instances of observed mismatches in the known scores assigned to those taxa) and the total potential
127 distance for all the scorable comparisons (i.e., the maximum possible number of mismatches in the
128 actual compared scores, if these were to differ entirely in the two taxa in question). Second, each
129 instance of a character comparison for which one or both taxa show missing entries is replaced with
130 a value given by multiplying the WMFS by the potential distance. Third, the value thus obtained is
131 added to the distance obtained for the scorable comparisons, in order to give a corrected Manhattan
132 distance. Fourth, all the character distances are squared and summed. Fifth, the resulting values are
133 rooted to yield a generalized Euclidean distance between the two taxa in question [32].

134 In theory, different types of taxon-taxon distances could be used. We opted for the generalized
135 Euclidean distances because ordination methods (e.g., Principal Coordinates Analysis [32,38,40]; but
136 see also [46]) conducted on square dissimilarity matrices of Euclidean distances result in many fewer
137 'negative' multivariate axes (which are difficult to interpret) than in the case of square dissimilarity
138 matrices consisting of, say, Manhattan distances. More fundamentally, the Euclidean nature of the
139 distance matrix is a logical prerequisite for geometric interpretations of patterns of morphospace
140 occupation (triangle inequality criterion holding strictly in a multidimensional Euclidean space). In

addition, although correction for negative axes is strictly speaking unnecessary with Euclidean distances, we found that in practice, a variable number of such axes may still result from matrix ordination. We therefore followed the practice of several recent works [38-45] in performing the multivariate analysis of tabulated distances with negative eigenvalue correction.

Because the data matrix in [11] includes discrete binary, discrete multistate, and continuous characters, we selected appropriate 'codes' in MATRIX for each character. Thus, code 'b' (for binary) is appropriate when one wishes to record shared derived similarity (for instance, shared loss of a feature; with this code, matches of '1's in two taxa yield a distance of '0', but other state-to-state comparisons and matches of '0's yield a distance of '1'). Code 'm' (for multistate) is suitable for multistate unordered discrete characters (identical states give a distance of '0'; different states give a distance of '1'). Finally, we applied code 'r' (for ranged) to continuous characters (this code is also suitable for multistate ordered discrete data), whereby the distance between two taxa is a fraction of the range between the highest and lowest values in the set of continuous measurements [32,63].

(**e**) ***Multivariate analyses of pair-wise distances***

We applied Principal Coordinates (PCo) Analysis to the pair-wise generalized Euclidean distances using GINKGO v. 1.5.5 [64], a free multi-platform Java application for multivariate analyses and statistics (http://biodiver.bio.ub.es/ginkgo/Ginkgo.htm). We performed the PCoa analysis with the Cailliez method [65] of negative eigenvalue correction [66], placing the centroid for the 87 taxa on the origin of the multivariate axes. PCo scores (i.e., PCo coordinates) for all taxa along each PCo axis are tabulated in Dataset S2. We determined the number of ordination axes (PCo components) to be retained for further analyses and tests by inspecting a semi-logarithmic scree plot in GINKGO [67]. In this plot, the $\log_{10}$-transformed eigenvalues (representing the amounts of variance explained by the principal coordinates) were plotted against axis number. The occurrence of a 'break' in the distribution of log-transformed eigenvalues between PCo axes 9 and 10 (after which point no major discontinuities could be observed) marked the threshold point for retaining 10 axes. Although other

criteria for axis retention are available, they provided less consistently meaningful results. As an example, we could have used the number of axes that, together, summarize a given percentage of the total variance in the data (e.g., 90%). However, it became apparent that many of the higher axes differed little from one another and contributed little structure to the morphospace. The use of the broken-stick method in the 'ape' library for 'R' (http://www.r-project.org/; [68]) also proved to be unsatisfactory. Without negative eigenvalue correction, this method selected axis PCo1 only. With negative eigenvalue correction, it selected axes PCo1-3. However, as explained above, higher axes PCo4-10 still contribute to morphospace structure.

(**f**) ***Disparity metrics***

We used the PCo scores on the first 10 PCo axes to calculate four disparity metrics, namely the sum and root-product of both ranges and variances. To this end, we used the RARE v. 1.2 disparity and rarefaction software [32,38-45,47,63]. We show exclusively the results based upon the two sums, as the root-products reveal very similar patterns. In a geometrically intuitive way, the sum of ranges measures the amount of morphospace occupation, whereas the sum of variances measures the spread of taxa relative to their own centroid (i.e., their mean configuration) [63]. For each disparity metric, we built a complete rarefaction profile of mean and median values by extracting numerous times (1000 bootstrap replicates) all the subsamples of taxa between 1 and the maximum number of taxa in any specific set, such as a clade or a time interval (Datasets S3 and S4). We calculated 95% confidence intervals for each subsampling routine. For disparity plots, we show both the un-rarefied median value with the associated confidence interval for each disparity metric (using the maximum number of taxa occurring in any specific taxon set), and the values generated by rarefaction of all samples to the smallest taxon set (in order to standardize calculations by taking into account the unevenness of sample sizes, i.e. taxon numbers, in the different sets).

(**g**) ***Morphospace plots***

193 Two- and three-dimensional plots using combinations of the first three PCo axes were inspected in
194 PAST v. 2.14 (http://folk.uio.no/ohammer/past [50]). To evaluate differences among groups, we
195 used a one-way analysis of similarities (ANOSIM: null hypothesis: equal median and range for
196 ranked dissimilarities within groups of interest; [69]) and a one-way non-parametric multivariate
197 analysis of variance (npMANOVA: null hypothesis: similar variances for groups of interest; [70]).
198 As non-parametric tests, ANOSIM and npMANOVA are useful in the absence of any information
199 on the distribution of scores on the PCo axes [40,41,50]. The results of both tests are reported in
200 Dataset S5.
201
202 (**h**) ***Experiments with autapomorphies***
203 The effects of autapomorphies on analyses of disparity and morphospace occupation require full
204 scrutiny. This issue was discussed briefly in [32,40,41], and has been raised more recently [46].
205 Although we reserve a more exhaustive treatment for a separate publication, some observations are
206 necessary before we describe our simulation experiments. These experiments provide empirical
207 evidence, based on the anomodont case study, that autapomorphies may not affect relative measures
208 of disparity (i.e., the profiles of increasing or decreasing disparity between groups) and patterns of
209 morphospace occupation to any significant extent.
210 There are practical limits to the way in which unique traits can be codified for any organisms.
211 Even within a specific morphofunctional complex, e.g. the skeletal system, autapomorphy selection
212 may be problematic for several reasons. First, the status of each character changes depending on the
213 level of phylogenetic and taxonomic inclusiveness, the practice of lumping or splitting species, and
214 different coding and ordering régimes. Second, the distinctiveness of taxa may rest either on truly
215 unique traits or on exclusive combinations of features that may be present in other taxa, including
216 plesiomorphies. However, recognition of genuinely unique features is hampered by subjective
217 evaluations of the comparative data at hand, character-state delimitations, failure to account for
218 variability, incompleteness of material, limited character atomization, and newly discovered taxa,

among other factors. Third, unless one or more taxa represent true outliers (e.g., due to their extreme morphologies), the distribution of unique features among taxa under study might approach randomness, and thus have little impact on relative disparity values (of course, absolute values will change with autapomorphy inclusion) and on patterns of morphospace occupation (within- and between-group distances will vary isometrically, or nearly so). However, even with outliers, it may be possible to atomize characters in such a way that, for a seemingly unique feature X in taxon A, there might be a unique feature Y in taxon B, a unique feature Z in taxon C, …, etc. Furthermore, the unique features X, Y, Z need not relate to the same structure. As a hypothetical example, we might find unique and very subtle differences in ten femora belonging to ten anomodont species, such that each has a distinctive trait of its own. For any given unique trait in the femur of an anomodont species, there might be a unique trait of the cervical vertebrae in another species, and a unique trait of the palate in yet another species. Ultimately, this atomization of traits will lead to uniform distribution of autapomorphies across the data matrix. We seek to understand how realistic this scenario is, and we want to determine how a non-random distribution of autapomorphies might affect either disparity analyses or patterns of morphospace occupation.

To examine these issues, we performed a set of experiments designed to simulate occurrences of various proportions of hypothetical unique features distributed in different ways among the 87 taxa. Starting with the original matrix of 163 characters [11], we built seven simulated data sets consisting of hypothetical autapomorphies only, and added these to the existing character set. We assumed that the hypothetical autapomorphies were discrete traits that could be coded in all taxa (i.e., there were no instances of unknown conditions), such that in each case 86 taxa would be coded as ‘0’ and a single taxon would be coded as a ‘1’. Simulated sets were as follows:

a) Simulation 1: 250 characters (163 original, 87 hypothetical autapomorphies; 1 autapomorphy per taxon).

b) Simulation 2: 424 characters (163 original, 261 hypothetical autapomorphies; 3 autapomorphies per taxon).

c) Simulation 3: 598 characters (163 original, 435 hypothetical autapomorphies; 5 autapomorphies per taxon).

d) Simulation 4: 1033 characters (163 original, 870 hypothetical autapomorphies; 10 autapomorphies per taxon).

e) Simulation 5: 580 characters (163 original, 417 hypothetical autapomorphies; average of 4.79 autapomorphies per taxon). In this experiment, each taxon was randomly given a number of autapomorphies ranging between 1 and 10.

f) Simulation 6: 667 characters (163 original, 504 hypothetical autapomorphies; average of 5.79 autapomorphies per taxon). In this experiment, each Permian taxon was randomly given a number of autapomorphies ranging from 5 to 10, and each Triassic taxon a number of autapomorphies ranging from 1 to 5.

g) Simulation 7: 541 characters (163 original, 378 hypothetical autapomorphies; average of 4.34 autapomorphies per taxon). In this experiment, each Permian taxon was randomly given a number of autapomorphies ranging from 1 to 5, and each Triassic taxon a number of autapomorphies ranging from 5 to 10.

The logic behind these experiments is that in simulations 1-4, autapomorphies are distributed uniformly among taxa, and their numbers per taxon increase regularly. Obviously one could create infinite data sets with increasing numbers of unique traits added equally to all taxa. However, this would result in little more than an isometric expansion of morphospace occupancy. Simulation 5 is probably the most realistic, because autapomorphies are distributed at random among taxa. Finally, the last two data sets (simulations 6 and 7) are designed to address the issue of uneven distributions of unique traits in taxon partitions. Note that in the last two simulated data sets, the difference in

autapomorphies is consistent for all Triassic taxa (i.e., the two Triassic emydopoids, the species of *Lystrosaurus* that have Triassic first appearances, and the kannemeyeriiforms).

Another way in which uneven distributions of autapomorphies could occur is by taxonomic groups (clades or otherwise). However, devising simulations that accommodate this scenario would entail a large number of possibilities. Furthermore, these simulations may be unwarranted for two related reasons. First, we would have to postulate that the taxa in a given clade display an unusually large number of traits not present in any other group. These traits would be autapomorphies for the clade as a whole, and thus synapomorphies for its constituent taxa. In this respect, they would be coded precisely as such - hypotheses of primary homology that have passed a congruence test [71,72]. Second, even if these traits were distributed in some of the taxa only, or could not pass a congruence test (thus appearing as homoplasies), they would still be coded as primary homologies and incorporated into the data matrix in that way.

For each of the simulated data sets described above, we created a tabulation of generalized Euclidean distances in MATRIX and subjected them to PCoA using GINKGO (Dataset S1). Subsequently, we carried out a global Mantel test [73] using the free software CADM [74], also implemented in the 'ape' library for 'R' (http://www.bio.umontreal.ca/casgrain/en/labo/cadm.html). CADM first outputs results for the global test of overall incongruence among distance matrices (null hypothesis: incongruence of all matrices; statistics: Kendall's W coefficient of concordance and Friedman's X^2 for permutational tests of W). The matrices are subsequently permuted at random and independently, and a second test is provided for the null hypothesis that each matrix is incongruent with each of the other matrices in turn (statistics: rank-based [Spearman] Mantel correlations, and associated permutational probability). In addition, we conducted a posteriori permutation tests of the contributions of each distance matrix from a particular simulation to the overall concordance among the matrices. Tests were performed with 999 random permutations of matrix structure, and each matrix was given equal weight. In addition, for a posteriori permutation

294 tests we applied the Holm-Bonferroni correction of *p* values for multiple pair-wise comparisons
295 (Dataset S7).
296
297 (**i**) ***Taxon partitions by groups***
298 We performed a series of experiments using different types of taxon and character partitions. Taxon
299 partitions for disparity analyses through time (Dataset S3) and by major groups (Dataset S4) were
300 detailed in the main text.
301 As a frame of reference for the study of disparity changes across the evolutionary history of
302 anomodonts, we used the tree topology in figure S1 in order to delimit nine groups. Five of these
303 groups are paraphyletic assemblages (grades), whereas the remaining four are monophyla (clades).
304 There are various ways in which we could delimit groups. Our preferred solution gives a trade-off
305 between reflecting the sequence of key branching events in anomodont evolution and including the
306 several newly recognized levels of morphological organization in the clade as a whole [11]. The
307 nine major groups (and associate colour-codings) are: g1 = basal anomodonts, dark gray; g2 =
308 endothiodonts, dark magenta; g3 = emydopoids, cyan; g4 = cryptodonts, green; g5 = '*Dicynodon*'-
309 grade taxa, brown; g6 = lystrosaurids, red; g7 = dinodontosaurids plus shansiodontids, lime; g8,
310 kannemeyeriids, light blue; g9 = stahleckeriids, dark blue.
311
312 (**j**) ***Taxon partitions by time intervals***
313 We considered two sets of chronological taxon partitions. The first partition consisted of Permian
314 (P) and Triassic (T) intervals. The second partition consisted of stage-level intervals, as detailed in
315 the main text. The merging of some stages into longer time intervals was dictated by the paucity
316 (Roadian; Norian) or absence (Olenekian) of taxa in some stages. Stage-level time intervals are as
317 follows: t1 = Roadian-Wordian; t2 = Capitanian; t3 = Wuchiapingian; t4 = Changhsingian; t5 =
318 Induan-Olenekian; t6 = Anisian; t7 = Ladinian; t8 = Carnian-Norian.
319

(**k**) ***Character partitions: cranial vs. postcranial and discrete vs. continuous***

In order to evaluate the extent to which different skeletal regions and different types of characters may affect patterns of morphospace occupancy, we partitioned the data matrix in [11] into cranial vs. postcranial characters, as well as into discrete vs. continuous characters. For each character partition, we performed Mantel tests to assess the correlation between the two sets of pair-wise taxon distances derived from the partitions of interest (e.g., cranial vs. postcranial; discrete vs. continuous) (Dataset S7). For the cranial vs. postcranial data partition, it was necessary to prune the original data matrix, so as to produce the largest possible set of taxa for which the two character subsets could be coded (i.e., we deleted all taxa that were entirely scored as unknown in either the cranial or the postcranial data). This resulted in 54 taxa being retained. No taxon deletion was necessary for the partition into discrete vs. continuous characters.

SUPPLEMENTARY RESULTS

(**a**) ***Results of Principal Coordinates Analysis***

PCoA of the pair-wise taxon distances built from the data matrix in [29] results in the first 10 PCo axes explaining 24.512% of the total variance. PCo axes 1-3, used to show patterns of distribution of taxa in morphospace, summarize 6.643%, 3.0989%, and 2,4111% of total variance, respectively.

(**b**) ***Distribution of anomodonts in morphospace***

The two-dimensional morphospace plots delimited by combinations of the first three PCo axes (figure 1*a*, *b*) show the following main features. First, basal anomodonts (g1) occupy a much larger region of morphospace than all other groups (g2-g9) and despite their low taxic diversity, they are highly dispersed around their own centroids. Second, endothiodonts and emydopoids (g2 and g3) overlap considerably in morphospace, but tend to remain separate from more apical groups (g4-g9). Among the latter groups, consistent and conspicuous overlap characterizes *Dicynodon*-grade taxa (g5) and lystrosaurids (g6), on the one hand, and kannemeyeriids (g8) and stahleckeriids (g9), on the

346 other. Third, increasingly apical groups - especially g8 and g9 - form progressively more compact
347 clusters in morphospace (though g4 - the cryptodonts - is an exception).
348 The ANOSIM and npMANOVA tests corroborate in part the patterns expounded above. Both
349 ANOSIM (R = 0.6531; p = 0.0001) and npMANOVA (F = 13.5; p = 0.0001) results reject the null
350 hypothesis of no differences in morphospace distribution among the nine major groups. In both tests,
351 two post-hoc pair-wise comparisons are not significant, specifically those between g2 and g3 and
352 between g8 and g9. Furthermore, in the case of ANOSIM an additional non-significant comparison
353 is between g1 and g2.
354 When we introduce autapomorphies according to the simulations 1-7 described above, a global
355 Mantel test indicates a strong and significant overall positive correlation for the ensemble of the
356 generalized Euclidean distances based on the original data matrix and the distances obtained from
357 each of the simulations (W = 0.90671; X^2 = 27128.76226; p = 0.001). Correlations in each pair-wise
358 comparison of distance sets generated from simulations tend to be higher than the correlations
359 between the original distances and the distances from each simulation. This result suggests that,
360 regardless of number and distribution of autapomorphies, they have little impact on the relationships
361 among intertaxon distances. However, correlation values are moderate to high even in the case of
362 original and simulated distances. The lowest value is for distances from simulation 6, where we
363 assigned at random a number of autapomorphies ranging from 5 to 10 to Permian taxa, and from 1 to
364 5 to Triassic taxa. The highest value is for distances from simulation 1 - perhaps unsurprisingly, as
365 this simulation implies addition of just one unique trait to each taxon - followed by the values for
366 simulations 4 (second highest) and 2 (third highest). The four lower values, for simulations 3, 5-7,
367 are nevertheless still substantial and of those, the value for simulation 5 is the highest (this is the
368 most realistic experiment, implying addition of autapomorphies at random across the matrix).
369 For each simulation, two-dimensional scatter biplots delimited by the first two PCo axes
370 (easily derived using Dataset S2) show that (1) placements of groups g1-g9 relative to each other
371 are only slightly altered by the inclusion of autapomorphies, (2) most of the individual taxa occur in

similar relative positions, and (3) the relative sizes of the groups in morphospace are affected to a negligible degree. The plot from simulation 4 may appear at odds with all the others; however, this is due to the fact that it is actually merely reflected along PCo axis 2, so that the basic relationships and arrangement of taxa are still essentially the same. The most conspicuous change relative to the plot from the original matrix is that adding autapomorphies tends to "smear" groups out over larger areas of morphospace (most convex hulls delimiting groups appear to be "stretched"). This change appears to be quite consistent, even when the number of autapomorphies per taxon is introduced at random (e.g., as in simulation 5). These observations also hold true for simulations 6 and 7, where autapomorphies have different frequencies among Permian and Triassic taxa.

(**c**) ***Effects of character partitions***

Despite the reduced number of taxa in the comparison between morphospace plots generated from cranial and postcranial data, we found that pair-wise distances derived from cranial data correlate significantly, albeit only moderately, with those derived from postcranial data (Dataset S7). Since discrete characters constitute the greatest proportion of all characters in the original matrix [11], we expect the mutual positions of taxa in the discrete trait-based morphospace to be similar to the plot generated using all characters. This is indeed the case: groups g1-g9 show nearly identical relative arrangements and patterns of overlap to those in figure 1. When only continuous data are employed, however, substantial changes characterize the distribution of taxa. Thus, the nine major groups tend to overlap more extensively. Furthermore, except for a small number of outliers in groups 1 and 9, most remaining taxa form a compact cluster. These results are hardly surprising, given the highly divergent cladogram obtained from running a phylogenetic analysis employing only the continuous characters [11]. Results from the Mantel test show that the correlation between the distances from discrete characters and those from continuous characters are invariably very low for all correlation coefficients (though such coefficients differ significantly from random).

(**d**) ***Patterns of morphological disparity***

In the following, it is useful to report modalities of overlap among confidence intervals as a way to assess significant (non-overlap) or non-significant (overlap) differences between adjacent disparity values [30]. Invariably, both disparity metrics (sums of ranges and variances) show Permian taxa to be significantly more disparate than Triassic taxa, regardless of whether rarefaction is used (figure S2*a*, *b*). Temporal disparity profiles are more patterned when analyzed by stage subdivisions (figure S2*c*, *d*).

Plots using the rarefied sums of ranges and variances and the un-rarefied sum of variances are similar to one another and indicate a general, steady decrease in median disparity values from Roadian-Wordian (t1) (or Capitanian, t2, in the case of the rarefied sum of ranges) through to t8 (Carnian-Norian). In those three plots, disparity values in time bins t1-t4 (or t2-t4) appear to be significantly different from those in time bins t7 and t8. Again in the same plots, disparity in the Induan-Olenekian (t5) is not significantly different from disparity in all (or most, in the case of the un-rarefied sum of variances) other stages. The plot for the un-rarefied sum of ranges is utterly discordant from the other three plots examined so far. In particular, median values of disparity increase steeply from t1 to t2, fluctuate to a negligible extent in time bins t2-t4 (characterized by a small decrease), then drop dramatically across the PTB. During the Triassic, anomodont disparity experienced a new slight increase during t6, reaching a level comparable to that in t1. Finally, we observe a new steady decrease from t6 to t8.

Finally, we consider the contributions of individual groups to overall disparity (figure S2*e*, *f*). There are striking similarities in the profiles of disparity values from g1 through to g9 when either the sum of ranges or the sum of variances is used as metric. Furthermore, the un-rarefied and rarefied profiles also differ little from one another. The subtle differences between un-rarefied and rarefied plots, and between the two disparity metrics, are best captured in visual inspections of the plots. We therefore focus mostly on common patterns gleaned from each disparity plot.

In all plots, basal anomodonts (g1) and cryptodonts (g4) exhibit, respectively, the highest and lowest median disparity values. In addition, the confidence interval for g1 shows very small or no overlap with the confidence intervals for other groups, indicating significant or nearly significant differences from those. Somewhat intermediate between g1 and g4 - though closer to g4 than to g1 - are the comparable disparity values for endothiodonts (g3) and emydopoids (g4). Across groups g5 through to g9, the general trend is one of slow and negligible decrease in disparity, particularly evident in g6 (lystrosaurids), g7 (dinodontosaurids plus shansiodontids), and g8 (kannemeyeriids). In addition, confidence intervals for these five groups tend to overlap extensively, indicating no significant differences in disparity values among them.

(**e**) ***Decoupled disparity and diversity***

The plots in figures 3, S3, and S4 illustrate trends in disparity and diversity. These use all possible combinations of disparity metric (sum of ranges; sum of variances), correction for sample size (un-rarefied; rarefied), time series co-variation (raw data in time bins; generalized differencing of time series), taxic diversity (number of taxa used in cladogram; total number of anomodont taxa recorded so far; total number of major anomodont lineages). The non-significant correlation between the two sets of values (diversity; disparity) is borne out by the vast majority of calculations. However, there are three instances of significant comparisons. All involve generalized differencing treatment for the sum of ranges and generalized differencing for one of the following variables: number of taxa per time interval actually present in the phylogeny; total number of known anomodont taxa recorded in each time interval; and number of major lineages per time interval.

Taxic diversity is often decoupled from morphological disparity [75-77], and both variables may be affected by the use of different taxonomic ranks and preservation biases. It is commonly assumed that survivors of major extinction events undergo relatively rapid morphological diversification into supposedly empty ecospace. A clear understanding of these patterns in terrestrial

448 vertebrates requires study of diversity, ecological shifts, and morphofunctional change through the
449 time intervals of recovery [78-87].

450 There are potential issues with the quality of data in considering any aspect of the fossil record
451 [88-90]. A number of authors [89] have argued that the terrestrial record is especially incomplete,
452 and several have employed uncorrected data sets in macroevolutionary studies. However, for some
453 vertebrate clades the available paleontological documentation has been shown to be adequate [90],
454 and we reserve a detailed analysis of the impact of geological biases on anomodont diversity and
455 disparity through time for a separate publication. For the purposes of the present study, we note that
456 the sample of taxa is an adequate representation of known anomodont diversity at genus and, to a
457 more limited extent, species levels. As we employ different disparity metrics (each of which is
458 sensitive to various parameters), we can correct for possible biases introduced either by taxon sample
459 size (e.g., occurrence of outliers) or by taxonomic practice (e.g., by considering genera rather than
460 species, therefore avoiding the issue of taxonomic splitting/lumping).

461 At a more general level, note that local-scale investigations of both South African and Russian
462 Permo-Triassic sequences [91-93] have demonstrated that intensity of sampling increases through
463 the Early Triassic, the very time when diversity and disparity are lowest for several groups of
464 organisms. In fact, it has recently been pointed out [91,92,94,95] that the earliest Triassic time
465 intervals represent distinct outliers with unusually low taxon counts despite increased sampling.
466 Hence, the paleontological record (at least in those two key cases) may contain more biological than
467 geological signal, as these low taxon counts certainly are the result of the end-Permian mass
468 extinction. Interestingly, different sampling proxies used in the same study area often reflect the
469 same pattern, suggesting that they are converging on approximately the same sampling signal.
470 Hence, it remains essential to explore the relationship between varying sampling proxies and
471 diversity in more detail across a wide range of clades. The presentation of both raw and corrected
472 diversity estimates [96,97] should be preferred to enable a comparison between both signals and to
473 identify similarities as well as potential differences and their consequences.

SUPPLEMENTARY DATASETS

Dataset S1. Tabulation of pair-wise generalized Euclidean distances for the original data matrix, for each of the seven simulations with autapomorphy inclusions, and for each of the character partitions (cranial; postcranial; discrete; continuous).

Dataset S2. Tabulation of Principal Coordinates scores derived from Principal Coordinates analyses of the pair-wise generalized Euclidean distances for the original data matrix, for each of the seven simulations with autapomorphy inclusions, and for each of the character partitions (cranial; postcranial; discrete; continuous).

Dataset S3. Complete rarefaction profiles for sums and root-products of ranges and variances, including mean, median, and upper and lower bounds of 90% confidence intervals, calculated for eight stage-level time bins (t1 = Roadian-Wordian; t2 = Capitanian; t3 = Wuchiapingian; t4 = Changhsingian; t5 = Induan-Olenekian; t6 = Anisian; t7 = Ladinian; t8 = Carnian-Norian) and two system-level time bins (Permian; Triassic).

Dataset S4. Complete rarefaction profiles for sums and root-products of ranges and variances, including mean, median, and upper and lower bounds of 90% confidence intervals, calculated for nine major groups (g1 = basal anomodonts; g2 = endothiodonts; g3 = emydopoids; g4 = cryptodonts; g5 = 'Dicynodon'-grade taxa; g6 = lystrosaurids; g7 = dinodontosaurids plus shansiodontids; g8, kannemeyeriids; g9 = stahleckeriids).

Dataset S5. Tests of the separation among the nine major anomodont groups in morphospace, using Principal Coordinates on the first 10 axes, and Bonferroni correction of p values.

Dataset S6. Tests of the correlation between diversity and disparity through time, using un-rarefied and rarefied sums of ranges and variances, total number of described anomodonts and number of taxa in the phylogeny, and number of lineages. For all calculations, we also report results from generalized differencing.

Dataset S7. Tests of correlations among distance matrices obtained from the original data matrix, from each of the seven simulations with autapomorphy inclusions, and from each of the character partitions (cranial; postcranial; discrete; continuous).

SUPPLEMENTARY REFERENCES

1. Angielczyk, K. D. & Kurkin, A. A. 2003 Phylogenetic analysis of Russian Permian dicynodonts (Therapsida, Anomodontia), implications for Permian biostratigraphy and Pangaean biogeography. *Zool. J. Linn. Soc.* **139**, 157-212.
2. Surkov, M. V. & Benton, M. J. 2004 The basicranium of dicynodonts (Synapsida) and its use in phylogenetic analysis. *Palaeontology* **47**, 619-638.
3. Vega-Dias, C., Maisch, M. W. & Schultz, C. L. 2004 A new phylogenetic analysis of Triassic dicynodonts (Therapsida) and the systematic position of *Jachaleria candelariensis* from the Upper Triassic of Brazil. *N. Jahr. Geol. Paläontol., Abhand.* **231**, 145-166.
4. Angielczyk, K. D. 2007 New specimens of the Tanzanian dicynodont "*Cryptocynodon*" *parringtoni* von Huene, 1942 (Therapsida, Anomodontia), with an expanded analysis of Permian dicynodont phylogeny. *J. Vert. Paleontol.* **27**, 116-131.
5. Damiani, R., Vasconcelos, C., Renaut, A., Hancox, J. & Yates, A. 2007 *Dolichuranus primaevus* (Therapsida, Anomodontia) from the Middle Triassic of Namibia and its phylogenetic relationships. *Palaeontology* **50**, 1531-1546.
6. Fröbisch, J. & Reisz, R. R. 2008 A new species of *Emydops* (Synapsida, Anomodontia) and a discussion of dental variability and pathology in dicynodonts. *J. Vert. Paleontol.* **28**, 770-787.
7. Surkov, M. V. & Benton, M. J. 2008 Head kinematics and feeding adaptations of the Permian and Triassic dicynodonts. *J. Vert. Paleontol.* **28**, 1120-1129.
8. Angielczyk, K. D. & Rubidge, B. S. 2010 A new pylaecephalid dicynodont from the *Tapinocephalus* Assemblage Zone, Karoo Basin, Middle Permian of South Africa. *J. Vert. Paleontol.* **30**, 1396-1409.

9. Liu, J., Rubidge, B. & Li, J. 2010 A new specimen of *Biseridens qilianicus* indicates its phylogenetic position as the most basal anomodont. *Proc. R. Soc. Lond. B* **277**, 285-292.

10 Cisneros, J. C., Abdala, F., Rubidge, B. S., Dentzien-Dias, D. & Bueno, A. O. 2011 Dental occlusion in a 260-million-year-old therapsid with saber canines from the Permian of Brazil. *Science* **331**, 1603-1605.

11. Kammerer, C. F., Angielczyk, K. D. & Fröbisch, J. 2011 A comprehensive taxonomic revision of *Dicynodon* (Therapsida, Anomodontia), and its implications for dicynodont phylogeny, biogeography, and biostratigraphy. *J. Vert. Paleontol. [sp. 1]* **31**, 1-158.

12. Fröbisch, J. 2009 Composition and similarity of global anomodont-bearing tetrapod faunas. *Earth Sci. Rev.* **95**, 119-157.

13. Rubidge, B. S. 1995 Biostratigraphy of the Beaufort Group (Karoo Supergroup). *South Afr. Comm. Stratig. Biostratig.* **Ser. 1**, 1.

14. Gradstein, F. M., Ogg, J. G. & Smith, A. G. 2004 *A geologic time scale*. Cambridge: Cambridge University Press.

15. Lucas, S. G. 1998 Global Triassic tetrapod biostratigraphy and biochronology. *Palaeogeog. Palaeocl. Palaeoecol.* **143**, 347-384.

16. Lucas, S. G. 2002 Tetrapods and the subdivision of Permian time. In *Carboniferous and Permian of the world* (eds L. V. Hills, C. Henderson & E. W. Bamber), pp. 479-491. *Can. Soc. Petr. Geol. Mem.* **19**.

17. Lucas, S. G. 2004 A global hiatus in the Middle Permian tetrapod fossil record. *Stratigraphy* **1**, 47-64.

18. Lucas, S. G. 2005 Permian tetrapod faunachrons. *New Mexico Mus. Nat. Hist. Sci. Bull.* **30**, 197-201.

19. Lucas, S. G. 2006 Global Permian tetrapod biostratigraphy and biochronology. In *Non-marine Permian biostratigraphy and biochronology* (eds S. G. Lucas, G. Cassinis & J. W. Schneider), pp. 65-93. *Geol. Soc. Spec. Publ.* **265**.

20. Lucas, S. G. & Huber, P. 2003 Vertebrate biostratigraphy and biochronology of the nonmarine Late Triassic. In *The Great Rift Valleys of Pangea in Eastern North America - volume II, Sedimentology, stratigraphy, and paleontology* (eds P. M. LeTourneau & P. E. Olsen), pp. 143-191. New York: Columbia University Press.

21. Lucas, S. G., Hunt, A. P., Heckert, A. B. & Spielmann, J. A. 2007 Global Triassic tetrapod biostratigraphy and biochronology, 2007 status. *New Mexico Mus. Nat. Hist. Sci. Bull.* **41**, 229-240.

22. Furin S. *et al*. 2006 High-precision U-Pb zircon age from the Triassic of Italy, implications for the Triassic time scale and the Carnian origin of calcareous nannoplankton and dinosaurs. *Geology* **34**, 1009-1012.

23. Mundil, R., Ludwig, K. R., Metcalfe, I. & Renne, P. R. 2004 Age and timing of the Permian mass extinctions, U/Pb dating of closed-system zircons. *Science* **305**, 1760-1763.

24. Ogg, J. G. 2008 Triassic Period. In *The concise geologic time scale* (eds J. G. Ogg, G. Ogg & F. M. Gradstein), pp. 95-106. Cambridge: Cambridge University Press.

25. Rubidge, B. S. 2005 Re-uniting lost continents - Fossil reptiles from the ancient Karoo and their wanderlust. *South Afr. J. Geol.* **108**, 135-172.

26. Battail, B. 1993 On the stratigraphy of the Triassic therapsid-bearing formations. *New Mexico Mus. Nat. Hist. Sci. Bull.* **3**, 31-35.

27. Lucas, S. G. 1995 Triassic dicynodont biochronology. *Albertiana* **16**, 33-40.

28. Lucas, S. G. 1999 Tetrapod-based correlation of the nonmarine Triassic. *Zentralblatt Geol. Paläontol. Teil I* **1998**, 497-521.

29. Angielczyk, K. D. *et al*. in press Permian and Triassic dicynodont (Therapsida: Anomodontia) faunas of the Luangwa Basin, Zambia: update and implications for dicynodont biogeography and biostratigraphy. In *Early evolutionary history of the Synapsida* (eds C. F. Kammerer, K. D. Angielczyk & J. Fröbisch). Dordrecht: Springer.

30. Foote, M. 1991 Analysis of morphological data. In *Analytical paleobiology. Shourt courses in paleontology, no. 4* (eds N. L. Gilinski & P. W. Signor), pp. 59-86. Knoxville: The Paleontological Society.

31. Foote, M. 1992 Paleozoic record of morphological diversity in blastozoan echinoderms. *Proc. Natl. Acad. Sci. USA* **89**, 7325-7329.

32. Wills, M. A., Briggs, D. E. G. & Fortey, R. A. 1994 Disparity as an evolutionary metric: a comparison of Cambrian and Recent arthropods. *Paleobiology* **20**, 93-130.

33. Roy, K. & Foote, M. 1997 Morphological approaches to measuring biodiversity. *Trends Ecol. Evol.* **12**, 277-281.

34. Wills, M. A. 1998 Cambrian and Recent disparity: the picture from priapulids. *Paleobiology* **24**, 177-199.

35 Wills, M. A. 1998 Crustacean disparity through the Phanerozoic: comparing morphological and stratigraphic data. *Biol. J. Linn. Soc.* **65**, 455-500.

36. Wills, M. A. 2001 Morphological disparity: a primer. In *Fossils, phylogeny and form. An analytical approach* (eds J. Adrain, B. Lieberman & G. Edgecombe), pp. 55-144. New York: Plenum.

37. Stockmeyer Lofgren, A., Plotnick, R. E. & Wagner, P. J. 2003 Morphological diversity of Carboniferous arthropods and insights on disparity patterns through the Phanerozoic. *Paleobiology* **29**, 349-368.

38. Brusatte, S. L., Benton, M. J., Ruta, M. & Lloyd, G. T. 2008 Superiority, competition, and opportunism in the evolutionary radiation of dinosaurs. *Science* **321**, 1485-1488.

39. Brusatte, S. L., Benton, M. J., Ruta, M. & Lloyd, G. T. 2008 The first 50 Myr of dinosaur evolution, macroevolutionary pattern and morphological disparity. *Biol. Lett*. **4**, 733-736.

40. Ruta, M. 2009 Patterns of morphological evolution in major groups of Palaeozoic Temnospondyli (Amphibia: Tetrapoda). In *Patterns and processes in early vertebrate evolution* (eds M. Ruta, J. A. Clack & A. C. Milner), pp. 91-120. *Spec. Papers Palaeontol.* **81**.

602 41. Cisneros, J. C. & Ruta, M. 2010 Morphological diversity and biogeography of procolophonids
603 (Amniota, Parareptilia). *J. Syst. Palaeontol.* **8**, 607-625.
604 42. Young, M. T., Brusatte, S. L., Ruta, M. & de Andrade, M. B. 2010 The evolution of
605 Metriorhynchoidea [mesoeucrocodylia, thalattosuchia], an integrated approach using geometric
606 morphometrics, analysis of disparity, and biomechanics. *Zool. J. Linn. Soc.* **158**, 801-859.
607 43. Brusatte, S. L., Benton, M. J., Lloyd, G. T., Ruta, M. & Wang, S. C. 2011 Macroevolutionary
608 patterns in the evolutionary radiation of archosaurs (Tetrapoda, Diapsida). *Earth Environ. Sci.*
609 *Trans. R. Soc. Edin.* **101**, 367-382.
610 44. Prentice, K. C., Ruta, M. & Benton, M. J. 2011 Evolution of morphological disparity in
611 pterosaurs. *J. Syst. Palaeontol.* **9**, 337-353.
612 45. Thorne, P. M., Ruta, M. & Benton, M. J. 2011 Resetting the evolution of marine reptiles at the
613 Triassic-Jurassic boundary. *Proc. Natl. Acad. Sci. USA* **108**, 8339-8344.
614 46. Anderson, P. S. L., Friedman, M., Brazeau, M. D. & Rayfield, E. J. 2011 Initial radiation of
615 jaws demonstrated stability despite faunal and environmental change. *Nature* **476**, 206-209.
616 47. Brusatte, S. L., Montanari, S., Yi, H.-Y. & Norell, M. A. 2011 Phylogenetic corrections for
617 morphological disparity analysis, new methodology and case studies. *Paleobiology* **37**, 1-22.
618 48. McGhee, G. R., Jr. 1999 *Theoretical morphology. The concept and its applications*. New
619 York: Columbia University Press.
620 49. Zelditch, M. L., Swiderski, D. L., Sheets, H. D. & Fink, W. L. 2004 *Geometric morphometric*
621 *for biologists. A primer*. New York: Academic Press.
622 50. Hammer, Ø. & Harper, D. 2006 *Paleontological data analysis*. Oxford: Blackwell Publishing.
623 51. Erwin, D. H. 2007 Disparity, morphological patterns and developmental context.
624 *Palaeontology* **50**, 57-73.
625 52. Ciampaglio, C. N., Kemp, M. & McShea, D. V. 2001 Detecting changes in morphospace
626 occupation patterns in the fossil record, a characterization and analysis of measures of
627 disparity. *Paleobiology* **27**, 695-715.

628 53. Adams, D. C., Rohlf, F. J. & Slice, D. E. 2004 Geometric morphometrics: ten years of progress
629 following the 'revolution'. *Ital. J. Zool.* **71**, 5-16.
630 54. Slice, D. E. 2007 Geometric morphometrics. *Ann. Rev. Anthropol.* **36**, 261-281.
631 55. Mitteroecker, P. & Gunz, P. 2009 Advances in geometric morphometrics. *Evol. Biol.* **36**, 235-
632 247.
633 56. Mitteroecker, P. & Huttegger, S. M. 2009 The concept of morphospaces in evolutionary and
634 developmental biology, mathematics and metaphors. *Biol. Theory* **4**, 54-67.
635 57. MacLeod, N. & Forey, P. L. 2002 *Morphometrics, shape, and phylogeny*. London: Taylor and
636 Francis.
637 58. Marcus, L. F. 1990 Traditional morphometrics. In *Proceedings of the Michigan morphometric*
638 *workshop* (eds F. J. Rohlf & F. L. Bookstein), pp. 77-122. *Univ. Mich. Mus. Zool. Spec. Publ.*
639 **2**.
640 59. Bookstein, F. L. 1991 *Morphometric tools for landmark data: geometry and biology.*
641 Cambridge: Cambridge University Press.
642 60. Dryden, I. L. & Mardia, K. V. 1998 *Statistical shape analysis*. New York: Wiley.
643 61. MacLeod, N. 1999 Generalizing and extanding the eigenshape method of shape space
644 visualization and analysis. *Paleobiology* **25**, 107-138.
645 62. Elewa, A. M. T. 2004 *Morphometrics: applications in biology and paleontology*. Berlin:
646 Springer.
647 63. Villier, L. & Korn, D. 2004 Morphological disparity of ammonoids and the mark of Permian
648 mass extinctions. *Science* **306**, 264-266.
649 64. Cáceres, M., Font, X., Oliva, F. & Vives, S. 2007 GINKGO, a program for non-standard
650 multivariate fuzzy analysis. *Adv. Fuzzy Sets Syst.* **2**, 41-56.
651 65. Cailliez, F. 1983 The analytical solution to the additive constant problem. *Psychometrika* **48**,
652 305-308.
653 66. Legendre, P. & Legendre, L. 1998 *Numerical ecology*. Amsterdam: Elsevier Science.

67. Cangelosi, R. & Gorieli, A. 2007 Component retention in principal component analysis with application to cDNA microarray data. *Biol. Direct* **2**, 1-22.

68. Paradis, E., Claude, J. & Strimmer, K. 2004 APE, Analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289-290.

69. Clarke, K. R. 1993 Non-parametric multivariate analyses of changes in community structure. *Austral. J. Ecol.* **18**, 117-143.

70. Anderson, M. J. 2001 A new method for non-parametric multivariate analysis of variance. *Austral. Ecol.* **26**, 32-46.

71. Patterson, C. 1982 Morphological characters and homology. In *Problems of phylogenetic reconstruction* (eds K. A. Joysey & A. E. Friday, pp. 21-74. *Syst. Ass. Spec. Vol.* **21**.

72. de Pinna, M. G. G. 1991 Concepts and tests of homology in the cladistic paradigm. *Cladistics* **7**, 367-394.

73. Mantel, N. 1967 The detection of disease clustering and a generalized regression approach. *Cancer Res.* **27**, 209-220.

74. Legendre, P. & Lapointe, J.-F. 2004 Assessing congruence among distance matrices, single malt Scotch whiskies revisited. *Austral. New Zeal. J. Stat.* **46**, 615-629.

75. Foote, M. 1993 Discordance and concordance between morphological and taxonomic diversity. *Paleobiology* **19**, 185-204.

76. Ruta, M., Wagner, P. J. & Coates, M. I. 2006 Evolutionary patterns in early tetrapods. I. Rapid initial diversification followed by decrease in rates of character change. *Proc. R. Soc. Lond. B* **273**, 2107-2111.

77. Wagner, P. J., Ruta, M. & Coates, M. I. 2006 Evolutionary patterns in early tetrapods. II. Differing constraints on available character space among clades. *Proc. R. Soc. Lond. B* **273**, 2113-2118.

78. Sahney, S. & Benton, M. J. 2008 Recovery from the most profound mass extinction of all time. *Proc. R. Soc. Lond. B* **275**, 759-765.

79. Ward, P. D. *et al*. 2005 Abrupt and gradual extinction among Late Permian land vertebrates in the Karoo Basin, South Africa. *Science* **307**, 709-714.

80. Retallack, G. J. *et al.* 2006 Middle-Late Permian mass extinction on land. *Geol. Soc. Amer. Bull.* **118**, 1398-1411.

81. Botha, J. & Smith, R. M. H. 2006 Rapid vertebrate recuperation in the Karoo Basin of South Africa following the End-Permian extinction. *J. Afr. Earth Sci.* **45**, 502-514.

82. Roopnarine, P. D., Angielczyk, K. D., Wang, S. C. & Hertog, R. 2007 Trophic network models explain instability of Early Triassic terrestrial communities. *Proc. R. Soc. Lond. B* **274**, 2077-2086.

83. Ruta, M., Pisani, D., Lloyd, G. T. & Benton, M. J. 2007 A supertree of Temnospondyli, cladogenetic patterns in the most species-rich group of early tetrapods. *Proc. R. Soc. Lond. B* **274**, 3087-3095.

84. Angielczyk, K. D. & Walsh, M. L. 2008 Patterns in the evolution of nares size and secondary palate length in anomodont therapsids (Synapsida), implications for hypoxia as a cause for end-Permian terrestrial vertebrate extinctions. *J. Paleontol.* **82**, 528-542.

85. Stayton, C. T. & Ruta, M. 2006 Geometric morphometrics of the skull roof of stereospondyls (Amphibia, Temnospondyli). *Palaeontology* **49**, 307-337.

86. Ruta, M. & Benton, M. J. 2008 Calibrated diversity, tree topology and the mother of mass extinctions, the lesson of temnospondyls. *Palaeontology* **51**, 1261-1288.

87. Ruta, M., Cisneros, J. C., Liebrecht, T., Tsuji, L. A. & Müller, J. 2011 Amniotes through major biological crises, faunal turnover among parareptiles and the end-Permian mass extinction. *Palaeontology* **54**, 1117-1137.

88. Smith, A. B. 2007 Marine diversity through the Phanerozoic, problems and prospects. *J. Geol. Soc.* **164**, 731-745.

89. Smith, A. B. & Peterson, K. J. 2002 Dating the time of origin of major clades, molecular clocks and the fossil record. *Ann. Rev. Earth Planet. Sci.* **30**, 65-88.

90. 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.
91. Fröbisch, J. in press Synapsid diversity and the rock record in the Permian-Triassic Beaufort Group (Karoo Supergroup), South Africa. In *Early evolutionary history of the Synapsida* (eds C, F. Kammerer, K. D. Angielczyk & J. Fröbisch). Dordrecht: Springer.
92. Fröbisch, J. 2013 Vertebrate diversity across the end-Permian mass extinction - separating biological and geological signals. *Palaeogeog. Palaeocl. Palaeoecol.* **372**, 50-61.
93. Benton, M. J., Tverdokhlebov, V. P. & Surkov, M. V. 2004 Ecosystem remodelling among vertebrates at the Permian-Triassic boundary in Russia. *Nature* **432**, 97-100.
94. Smith, R. M. H., Rubidge, B. S. & van der Walt, M. 2011 T. Therapsid biodiversity patterns and palaeoenvironments of the Karoo Basin, South Africa. In *Forerunners of mammals* (ed. A. Chinsamy-Turan), pp. 31-64. Bloomington: Indiana University Press.
95. Nicolas, M. & Rubidge, B. S. 2010 Changes in Permo-Triassic terrestrial tetrapod ecological representation in the Beaufort Group (Karoo Supergroup) of South Africa. *Lethaia* **43**, 45-59.
96. Lloyd, G. T. 2012 A refined modelling approach to assess the influence of sampling on palaeobiodiversity curves: new support for declining Cretaceous dinosaur richness. *Biol. Lett.* **8**, 123-126.
97. McKinney, M. L. 1990 Classifying and analysing evolutionary trends. In *Evolutionary trends* (ed. K. J. McNamara), pp. 28-58. London: Belhaven.

SUPPLEMENTARY FIGURES

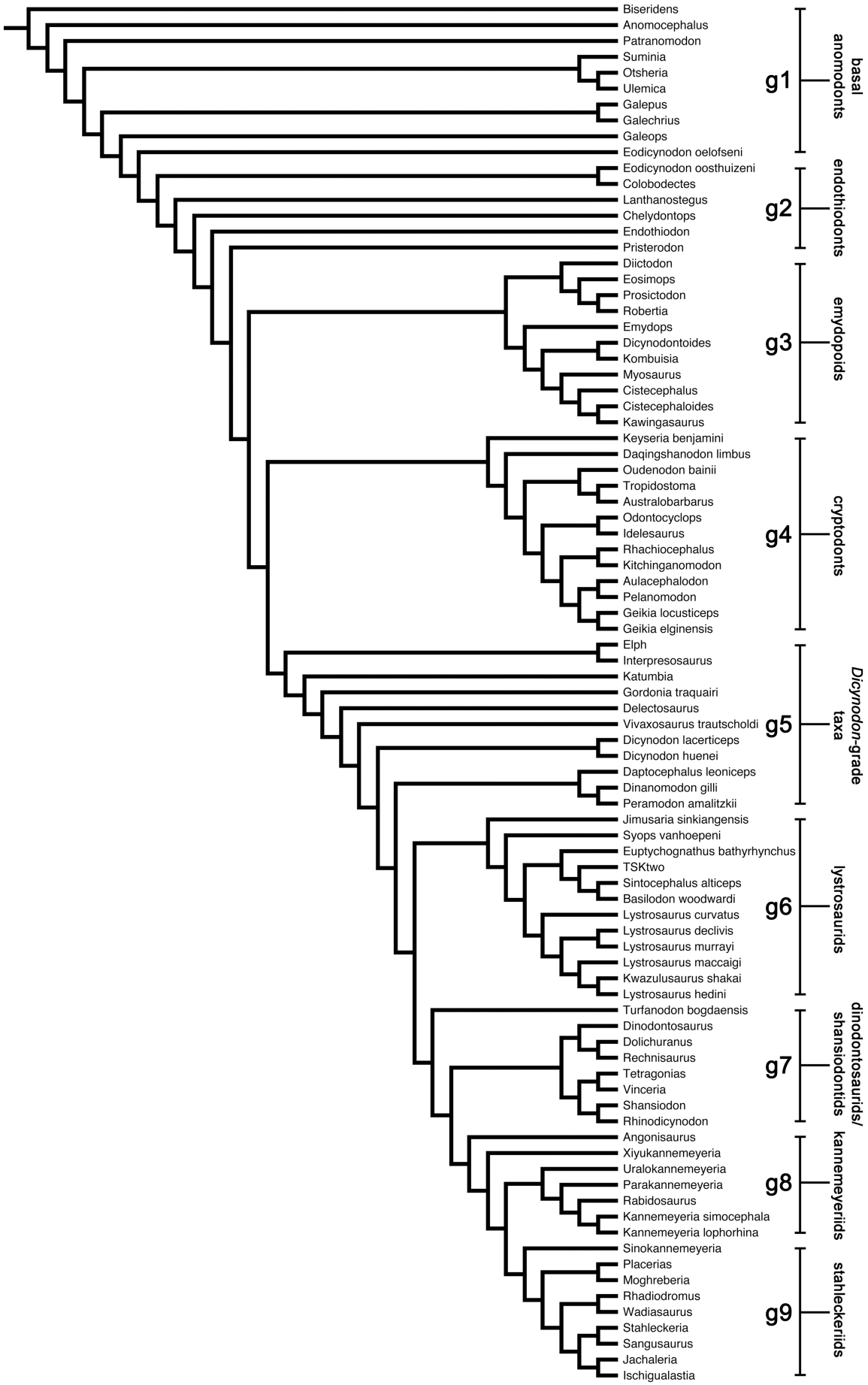
Figure S1. Anomodont phylogeny, based on analysis in [11].

Figure S2. Un-rarefied median disparity values and associated confidence intervals, based on the sum of ranges (left column plots) and the sum of variances (right column plots); vertical grey bars in (*a-d*) mark the Permian-Triassic boundary. (*a-b*) Disparity in Permian (P) and Triassic (T) taxa;

(*c*-*d*) disparity of taxa assigned to time intervals t1-t8 (see figure 2 for interval abbreviations); (*e*-*f*) disparity of taxa assigned to groups g1-g9 (see figure 1 for group notations and color codes).

Figure S3. Comparisons between anomodont disparity (un-rarefied median values, grey circles) and diversity (white squares) through time intervals t1-t8; disparity values are for the sum of ranges (left column plots) and the sum of variances (right column plots); symmetric error bars around the diversity values are calculated as $\pm\sqrt{N}$, where N is the number of taxa in a given time interval; vertical grey bars mark the Permian-Triassic boundary. (*a*-*b*) Comparisons based on the number of taxa actually present in the phylogeny; (*c*-*d*) comparisons based on the total number of known anomodont taxa.

Figure S4. Comparisons between anomodont disparity (median values, grey circles) and diversity (white squares) through time intervals t1-t8; disparity values are for the sum of ranges (left column plots) and the sum of variances (right column plots); diversity is expressed as the number of major lineages (groups g1-g9) present in a given time interval; vertical grey bars mark the Permian-Triassic boundary. (*a*-*b*) Comparisons based on rarefied disparity values; (*c*-*d*) comparisons based on un-rarefied disparity values.



747

748 **figure S1**

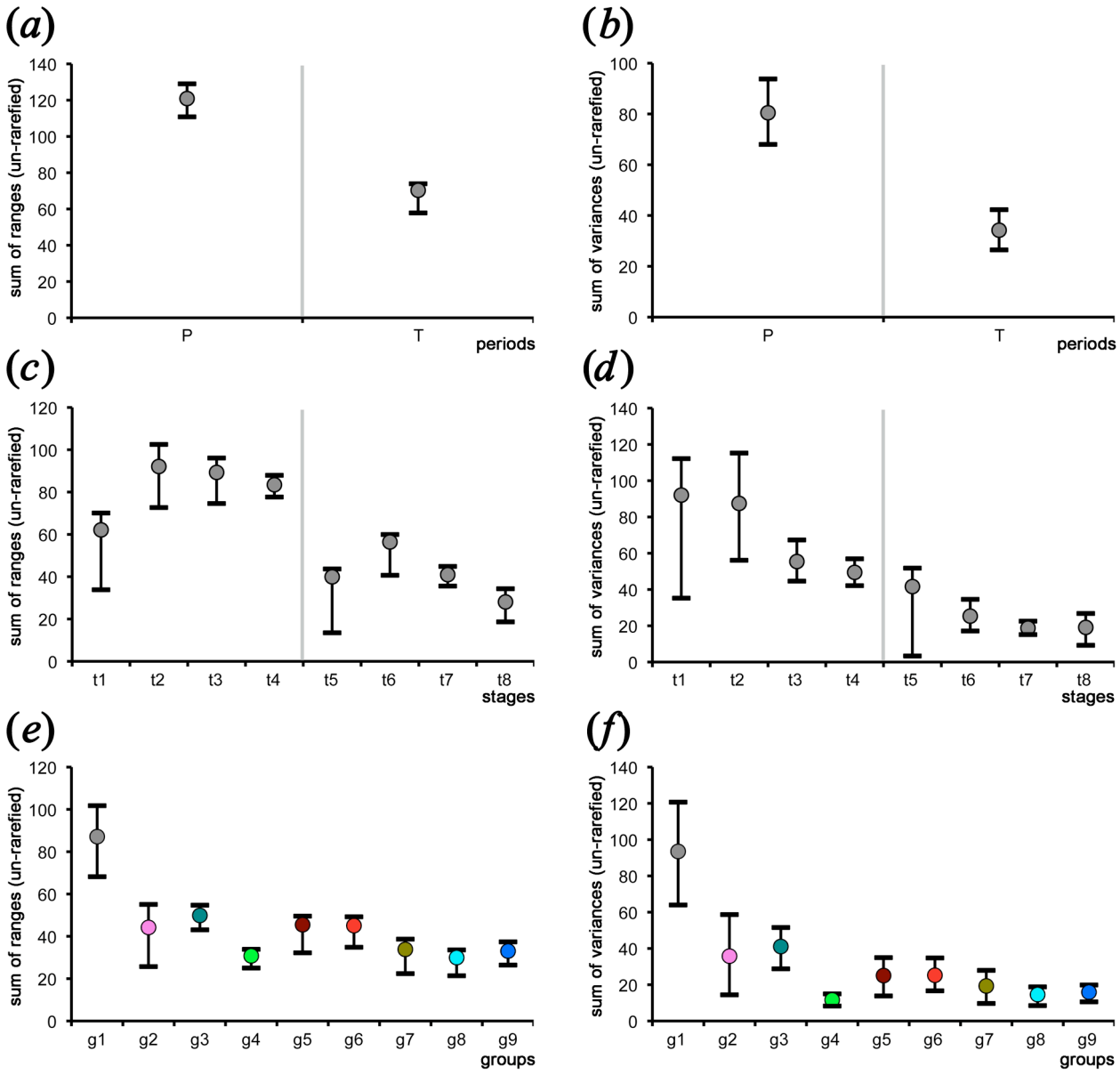

(a)
sum of ranges (un-rarefied)
140
120
100
80
60
40
20
0
P
T
periods
(b)
sum of variances (un-rarefied)
100
80
60
40
20
0
P
T
periods
(c)
sum of ranges (un-rarefied)
120
100
80
60
40
20
0
t1
t2
t3
t4
t5
t6
t7
t8
stages
(d)
sum of variances (un-rarefied)
140
120
100
80
60
40
20
0
t1
t2
t3
t4
t5
t6
t7
t8
stages
(e)
sum of ranges (un-rarefied)
120
100
80
60
40
20
0
g1
g2
g3
g4
g5
g6
g7
g8
g9
groups
(f)
sum of variances (un-rarefied)
140
120
100
80
60
40
20
0
g1
g2
g3
g4
g5
g6
g7
g8
g9
groups

figure S2

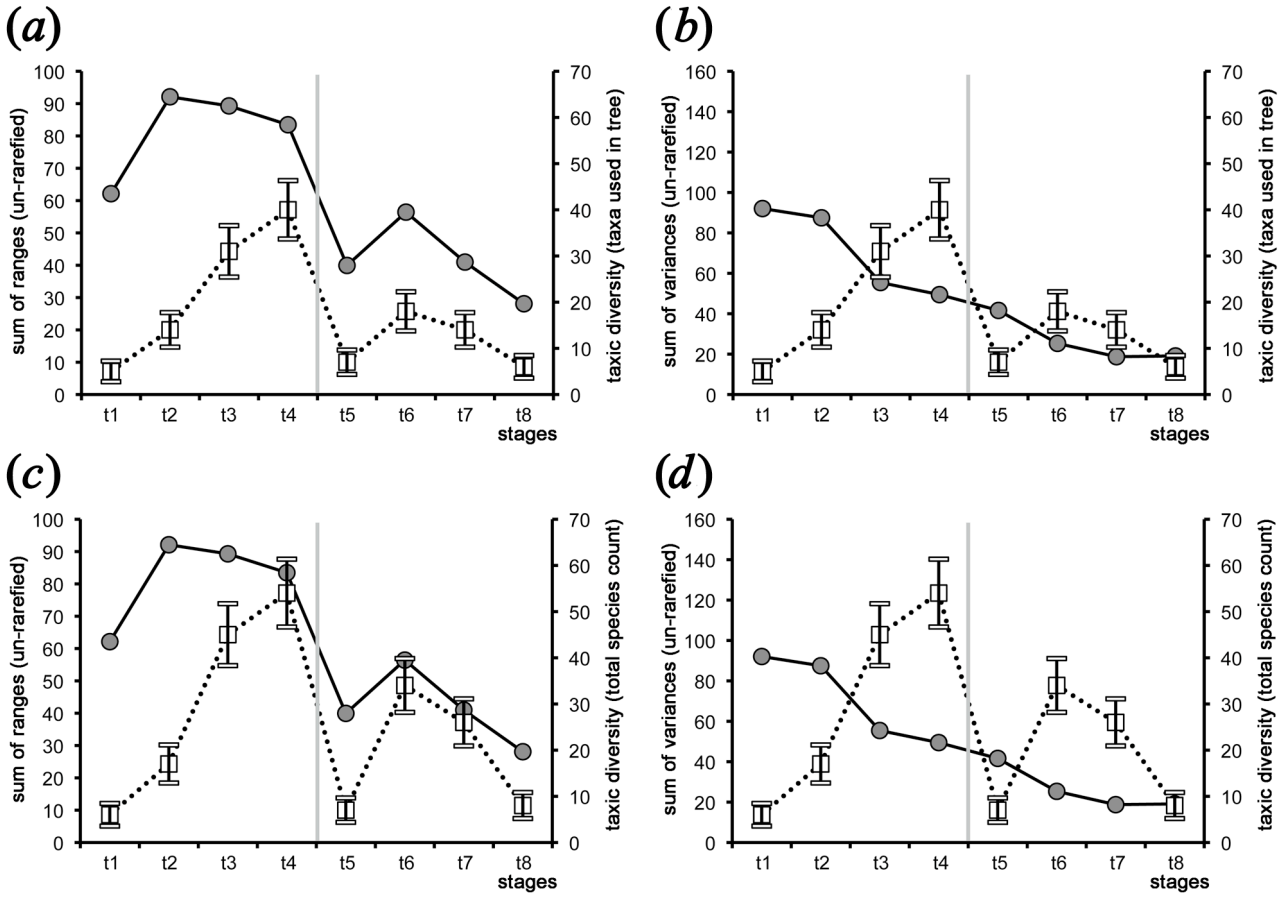
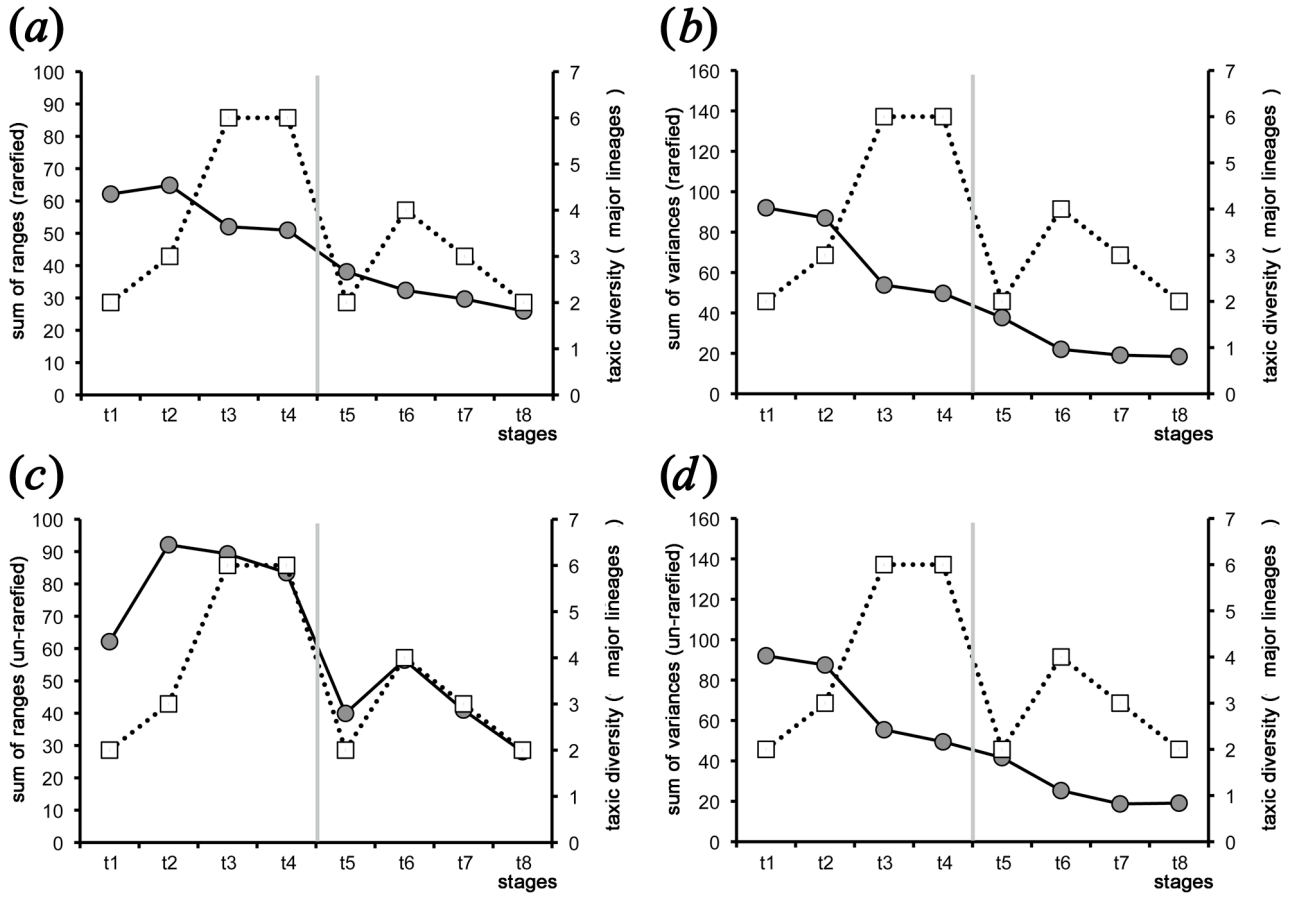

(a)
(b)
(c)
(d)
sum of ranges (un-rarefied)
sum of variances (un-rarefied)
taxic diversity (taxa used in tree)
taxic diversity (total species count)
stages
t1
t2
t3
t4
t5
t6
t7
t8

figure S3



774

775 **figure S4**