OXFORD
UNIVERSITY PRESS

Systematic Biology

Craniodental and postcranial characters of non–avian Dinosauria often imply different trees

SCHOLARONE™
Manuscripts

1
2
Craniodental and postcranial characters of non-avian Dinosauria often imply different trees

3

4 Yimeng Li[1], Marcello Ruta[2], & Matthew A. Wills[1]

5

6 **Author Affiliations:** [1]*The Milner Centre for Evolution, The University of Bath, The Avenue,*
7 *Claverton Down, Bath BA2 7AY, UK*

8 [2]*School of Life Sciences, University of Lincoln, Joseph Banks*
9 *Laboratories, Green Lane, Lincoln LN6 7DL, UK*

10

11
12 **Correspondence:** Matthew A. Wills

13 The Milner Centre for Evolution

14 The University of Bath

15 The Avenue

16 Claverton Down

17 Bath BA2 7AY

18 United Kingdom

19

20 Telephone: 01225 383504

21 Fax: 01225 386779

22 E-mail: m.a.wills@bath.ac.uk

23

24

25 **Submitted:** 25th August, 2018

26

27

28 **Craniodental and postcranial characters of non-avian Dinosauria often imply**
29 **different trees**

30

31 Yimeng Li, Marcello Ruta, Matthew A Wills*

32 *Corresponding author

33

34 **Abstract**

35 Despite the increasing importance of molecular sequence data, morphology still makes an important

36 contribution to resolving the phylogeny of many groups, and is the only source of data for most

37 fossils. Most systematists sample morphological characters as broadly as possible on the principle of

38 total evidence. However, it is not uncommon for sampling to be focussed on particular aspects of

39 anatomy, either because characters therein are believed to be more informative, or because

40 preservation biases restrict what is available. Empirically, the optimal trees from partitions of

41 morphological data sets often represent significantly different hypotheses of relationships. Previous

42 work on hard-part versus soft-part characters across animal phyla revealed significant differences in

43 about a half of sampled studies. Similarly, studies of the craniodental versus postcranial characters

44 of vertebrates revealed significantly different trees in about one third of cases, with the highest

45 rates observed in non-avian dinosaurs. We test whether this is a generality here with a much larger

46 sample of 81 published data matrices across all major dinosaur groups. Using the incongruence

47 length difference (ILD) test and two variants of the incongruence relationship difference (IRD) test,

48 we found significant incongruence in about 50% of cases. Incongruence is not uniformly distributed

49 across major dinosaur clades, being highest (63%) in Theropoda and lowest (25%) in Thyreophora.

50 Our partition tests are not sensitive to the size of the tree, to differences in partition size, to the

51 amount of missing data in the matrix, difference in the amount of missing data between partitions or

52 the difference in levels of homoplasy between partitions. Incongruence must therefore reflect

53 differences in *patterns* of homoplasy between partitions, which may itself be a function of

54   modularity and mosaic evolution. Finally, we implement new tests to determine which partition

55   yields trees most similar to those from the entire matrix. Despite no bias across dinosaurs overall,

56   there are striking differences between major groups, with the craniodental characters of

57   Ornithischia and the postcranial characters of Saurischia yielding trees most similar to the 'total

58   evidence' trees derived from the entire matrix.

59   **Introduction**

60   The fossil record is notoriously incomplete, not only in terms of diversity and species richness

61   (Verriere et al., 2016; Davies et al., 2017; Tutin and Butler, 2017), but also with respect to

62   stratigraphy (Maxwell and Benton, 1990; Dunhill et al., 2012; Brocklehurst and Froebisch, 2014;

63   O'Connor and Wills, 2016b; Verriere et al., 2016) palaeobiogeography (Lieberman, 2002; Ksepka and

64   Boyd, 2012; Davies et al., 2017), palaeoecology (Stanley et al., 1989; Jablonski, 2005) and behaviour

65   (Hsiang et al., 2015; Daley and Drage, 2016; Fan et al., 2017). However, it is organismal

66   incompleteness – the selective preservation of tissues and body regions – that impinges most

67   directly on attempts to infer phylogeny (Kearney and Clark, 2003; Cobbett et al., 2007; Sansom,

68   2015). The fossil record of non-avian dinosaurs mostly comprises bones and other hard parts (Wills

69   et al., 2008a; Mannion and Upchurch, 2010), but there are further biases towards the preservation

70   of more heavily mineralised and massive elements (e.g., limb bones) at the expense of more

71   frangible and delicate structures (e.g., skulls). Inferred relationships may differ substantially

72   depending upon which subsets of characters are used, but palaeontologists may nonetheless wish to

73   infer the relationships of dinosaurs described from partial skeletal material.

74   We therefore address four related questions using a sample of 81 cladistic taxon-character matrices

75   published between 2011 and 2017, each comprising both craniodental and postcranial characters.

76   Firstly we determine whether levels of homoplasy differ between characters of the skull and

77   dentition on one hand and characters pertaining to the body on the other. Any such difference

78   might be used to argue for the superiority of one body region over the other for phylogenetic

79 inference (Pettigrew, 1991; Sanchez-Villagra and Williams, 1998; Williams, 2007; Song and Bucheli,

80 2010; Mounce et al., 2016; Parker, 2016). Secondly we use established (incongruence length

81 difference, ILD; Farris et al., 1994) and more recent (Mounce et al., 2016; Sansom et al., 2017) tests

82 to determine whether the most parsimonious trees inferred from craniodental and postcranial

83 character partitions are significantly different (Mounce et al., 2016; Sansom et al., 2017). Thirdly, we

84 ask whether there are differences in the incidence of significant craniodental/postcranial

85 incongruence across major taxonomic groups. Fourthly, we present a novel test to determine

86 whether the tree(s) inferred from craniodental characters or the tree(s) derived from

87 postcraniodental characters are more similar to those derived from the entire matrix, with the latter

88 being used as a proxy for the 'true' phylogeny. We do this by resampling from the partitions and the

89 entire matrix in order to control for differences in the number of characters in each partition.

90

91 **What is the rationale for comparing the signals within partitions of characters?**

92 Suites of morphological characters are often functionally and developmentally integrated into

93 modules (Clarke and Middleton, 2008; Klingenberg, 2008; Lue et al., 2010) that can be subject to

94 different selection pressures and consequently evolve at different speeds (J, 1993; Lue et al., 2010;

95 Parker, 2016). This has consequences for the rate at which new character states are utilised and the

96 subsequent exhaustion of character space (Wagner, 1995; Wagner, 1997; Oyston et al., 2015;

97 Oyston et al., 2016), resulting in different levels of homoplasy. For example, it has been shown that

98 the dental characters of mammals are particularly labile and prone to convergence/reversal

99 (Sanchez-Villagra and Williams, 1998; Sansom et al., 2017), which is explicable in terms of the strong

100 functional and biomechanical constraints upon the form and arrangement of teeth. This particular

101 bias is unfortunate given the predominance of teeth in the mammal fossil record. More generally,

102 the craniodental and postcranial characters of vertebrates have been shown to contain significantly

103 incongruent signals about one time in three (Mounce et al., 2016). Moreover, it is possible that

104   incongruence is partly a function of the extent to which the skull and the rest of the body are

105   biomechanically decoupled (Ji, Q. et al. 1999). Fishes (lacking a functional neck) typically show

106   integration, while the most striking incongruence has been observed in some of the long-necked

107   dinosaur groups.

108   As a prerequisite for combining data in early, multi-gene molecular analyses, systematists commonly

109   checked for homogeneity of signal across loci using a variety of partition tests (Templeton, 1983;

110   Rodrigo et al., 1993; Farris et al., 1994). This practice is rarely implemented nowadays, and for

111   principally three reasons (Cunningham, 1997). Firstly, as analyses of increasingly large numbers of

112   genes graded into phylogenomic studies, the concept of the contingent inclusion of individual genes

113   became largely obsolete. Secondly, more advanced analytical methods were developed that allow

114   heterogeneous rates across sites and branches to be modelled rigorously (Damgaard, J. 2012).

115   Thirdly, a consensus emerged amongst systematists in favour of the simultaneous analysis of all

116   available character data, on the principle of 'total evidence' (Kluge, 1989), not least because of

117   'hidden support'. This is the phenomenon whereby signals that are weak within particular partitions

118   of the data may be common to many (or all) such partitions, such that they become the dominant

119   signal when all partitions are analysed together (Kluge, 1989; Gatesy et al., 1999; Gatesy and

120   Arctander, 2000; Wahlberg et al., 2005; O'Leary and Gatesy, 2008; Padial et al., 2010; Damgaard,

121   2012; Mounce et al., 2016).

122   Despite the ascendance of molecular phylogenetics, morphological character data can still

123   contribute to our understanding of the relationships of many groups (Houde, 1994; Wiens, 2004;

124   O'Leary and Gatesy, 2008; Nicolalde-Morejon et al., 2009; Gainett et al., 2014; Lopardo and

125   Hormiga, 2015). Moreover, for extinct and particularly for fossil groups, morphology is usually the

126   only direct source of phylogenetic data, notwithstanding exceptional cases utilising sub-fossil DNA

127   (Dabney et al., 2013; Shapiro and Hofreiter, 2014; Orlando et al., 2015).

128

Materials & Methods

Datasets

The character matrices utilised here were obtained from peer-reviewed papers published between

2000 and 2017. We utilised Graeme Lloyd's online compilation of matrices (Lloyd, 2018) and

searches of Web of Science using the keywords phylog* and evolution* in combination with any

higher taxonomic name of non-avian dinosaurs. We sought to sample all major dinosaur groups,

including matrices of varying dimensions sampled at different taxonomic levels. We initially included

104 matrices, but these were further checked for overlap since systematists often repurpose data or

otherwise add modest numbers of novel taxa and sometimes characters to existing studies. In order

to remove any such pseudoreplication from our sample, each data set was compared with every

other, and for each pair the number of matrix cells in common (replicated taxa and characters) was

expressed as a percentage of the total number of cells in the smaller of the two matrices. For pairs

with 20% or more overlap, the least inclusive (or otherwise the oldest) was removed from

consideration, reducing our sample to 81 data sets (See Supplementary Materials 1 for the

percentage of character and species overlap between all pairs). Character lists were then used to

define partitions. The "craniodental" partition included all characters pertaining to the skull and

dentition. The "postcranial" partition encompassed all characters of the vertebral column, girdles

and limbs.

Poorly known taxa, or those that are otherwise scored for only a small number of characters, can be

highly mobile within parsimony trees. This, in turn, can result in large numbers of MPTs, prohibitively

extending search times, and yielding poorly resolved consensus trees (Wilkinson, 1995; Mounce et

al., 2016). Where such complications were found in our analyses, the matrix was edited by removing

taxa with more than 40% of characters scored as missing ("?") or non-applicable ("-") in either

partition. Any characters rendered uninformative or invariant by this process were also deleted

(Wiens, 1998). On average, 20 taxa and 18 characters were removed from each dataset in this way,

154   equating to 33% of taxa and 7% of characters. For a list of characters and taxa discounted, see

155   Supplemenary Materials 2. Matrices were manipulated using Mesquite Version 3.04 (build 725) for

156   Macintosh. The resulting sample of 81 matrices contained an average of 26 taxa scored for a mean

157   of 115 craniodental and 133 postcranial characters.

158   **Measuring homoplasy**

159   The ensemble Consistency Index (CI)(Kluge and Farris, 1969) is a commonly used and well-

160   characterised index of homoplasy, and was calculated here to compare levels of homoplasy across

161   partitions. However, the CI suffers from well-documented drawbacks, notably its correlation with

162   the number of characters and taxa in the dataset (Archie, 1989; Mounce et al., 2016). We remove

163   these biases empirically here, using the residuals from regression analyses of CI on both matrix

164   dimensions. In addition, we report the ensemble Retention Index (RI)(Kluge and Farris, 1969) as a

165   measure of retained synapomorphy. All indices were calculated in PAUP* 4.0a.154 for Macintosh

166   (Swofford, 2017).

167   **Statistical tests for congruence**

168   The Incongruence Length Difference (ILD) test (Mickevich and Farris, 1981; Farris et al., 1995) is a

169   widely implemented partition homogeneity test based upon the difference in most parsimonious

170   tree length for a matrix when analysed as a whole, and the sum of MPT lengths for the partitions of

171   the matrix analysed in isolation (MPTs) (Norman et al.) More formally, the ILD for a bipartitioned

172   matrix is given by $L_{AB} - (L_A + L_B)/L_{AB}$, where L_{AB} is the optimal tree length (in steps) from the analysis of

173   the entire matrix (the total evidence analysis), and L_A and L_B are the optimal tree lengths for

174   partitions A and B analyzed independently (Figure 1). This ILD is compared with a distribution of ILD

175   values (here, 999) for random bipartitions of the matrix in the same proportions as the original, and

176   a p value is derived from the fraction of these as large or larger than the original. The ILD test has

177   been criticized on philosophical grounds, and because it has a high Type I error rate(Dolphin et al.,

178   2000; Barker and Lutzoni, 2002; Ramirez, 2006; Sansom et al., 2017). However, it remains very

179   widely applied (Mounce et al., 2016) and is used here as a measure of matrix partition incongruence

180   rather than as a criterion for combining those partitions.

181   In addition to the ILD test, we also implemented the incongruence relationship difference (IRD) test

182   of Ruta & Wills (2016) and Mounce et al. (2016). This is analogous to the ILD test in that a measure

183   of incongruence for the original data partition is compared with a distribution of incongruence

184   values for a large number of random partitions. However, whereas for the ILD incongruence is

185   measured in terms of additional tree length, a tree-to-tree distance metric is used for the IRD. Many

186   such metrics are available, but here we used the symmetrical-difference (RF) distance

187   (IRD_{RF})(Robinson and Foulds, 1981) and matching (MD) distance ($IRD_{Matching}$) (Lin et al, 2012). The RF

188   distance is well characterised and widely applied but prone to saturation, while the matching

189   distance has a particularly favourable distribution and other theoretically desirable properties. It is

190   unusual for a single most parsimonious tree (Norman et al.) to result from a parsimony search, and

191   we therefore followed Mounce *et al.* (2016) in calculating the mean nearest neighbour distance

192   (NND) between each tree resulting from one partition and the most similar tree in the other

193   partition. In addition, we calculated the distances between strict, semi-strict and 50% majority rule

194   (plus compatible groupings) trees for the two partitions, although we caution that these offer poor

195   or summaries of the differences between sets of trees (Mounce et al., 2016). IRD tests were initially

196   based upon 99 random partitions of the data (c.f. 999 for the computationally much faster ILD).

197   However, in those cases where p≤0.10, we re-ran the test for that matrix using 499 random

198   partitions).

199   All parsimony searches were implemented using 25 random additions of taxa, followed by tree

200   bisection and reconnection branch swapping, and retaining 10 trees at each step. We also

201   condensed the resulting most parsimonious trees by collapsing branches with a minimum length of

202   zero (Goloboff's 'amb-') and removing all but one of any consequently identical trees. To expedite

203   the searches, we limited the number of trees stored in memory to 100,000, and for the IRD tests we

204   calculated nearest neighbour tree-to-tree distances based upon no more than 1,000 most

205   parsimonious trees (MPTs) from each partition (1,999,000 tree-to-tree distances calculated for each

206   metric in order to find the minima). We further checked that fewer than 1,000 MPTs were found

207   from the analysis of the original partitions, and from no more than 5% of the randomised partitions.

208   Consensus trees were calculated from all MPTs, up to the 100,000 limit. All analyses were carried

209   out in PAUP* 4.0a.154 for Macintosh (Swofford, 2017), and with the use of scripts (see ESM) that

210   produced batch files for PAUP* and summarised it's logged output.

211

212   **Determining whether craniodental or postcraniodental characters yield trees more similar to**

213   **those from the entire data matrix**

214   In cases where the tree(s) inferred from craniodental and postcranial characters differ (and

215   especially where these differences are significant), it is reasonable to ask which tree is likely to be

216   most accurate. Unfortunately, there are no objective tests of phylogenetic accuracy, except in those

217   exceptional cases where phylogeny is known (e.g., laboratory cultures or simulated data sets). One

218   approach for extant taxa (Sansom et al., 2017) is to determine the congruence of suites of

219   morphological characters with a robustly supported molecular tree for the same taxa (an

220   independent data source). However, this assumes that the molecular tree is likely to offer the best

221   approximation of the truth: a standpoint defended in many quarters (Scotland et al., 2003; Olmstead

222   and Scotland, 2005; Wortley and Scotland, 2006; Zou and Zhang, 2016). For fossil taxa, stratigraphic

223   congruence may be used as an ancillary criterion for choosing between alternative trees (Wills,

224   1998; Wills et al., 2009; O'Connor and Zhou, 2013; O'Connor and Wills, 2016a), but this is only

225   defensible where the fossil record is relatively complete, or at least where the order of first

226   occurrences for lineages is preserved with reasonable fidelity (Wills et al., 2008b; Wills et al., 2009).

227   Here, we ask whether trees from the craniodental or postcranial data partition are most similar to

228   those derived from the entire data matrix, with the underlying assumption that the total evidence

229   tree is likely to be the most accurate (Kluge, 1989; Gatesy et al., 1999; Gatesy and Arctander, 2000).

230   A straightforward approach would be to calculate mean nearest neighbour tree-to-tree distance

231   metrics for the craniodental to entire tree sets versus the postcranial to entire tree sets. However,

232   *all other things being equal*, a larger partition contributes more characters to the entire matrix than

233   a smaller one. In cases where the optimal trees for the two partitions differed, the larger partition

234   might therefore be expected to yield trees more similar to those from the entire data set. The

235   difference in character numbers in the partitions could be overcome by differential weighting of

236   characters, but the tree-to-tree distance metrics utilised here are sensitive to the reductions in

237   resolution that are likely as the character:taxon ratio declines (and this is not overcome by

238   weighting). We therefore adopted a resampling approach, repeatedly jack-knifing characters at the

239   sample size of the smallest partition (n) from both the larger partition and the entire matrix. For the

240   entire matrix, we randomly jack-knifed the same number of characters ($n/2$) from both the

241   craniodental and postcranial partition, such that neither was favoured with a larger sample size.

242   Where n was an uneven number, we alternately sampled the 'additional' character from either

243   partition. For each of 100 resamplings, we then calculated the mean nearest neighbour distance

244   between craniodental and entire trees, and the mean nearest neighbour distance between

245   postcranial and entire trees. We report the median of these 100 comparisons (which partition is

246   closest to the entire), as well as Mann-Whitney test results to approximate a p value at which to

247   reject the null that the medians of these distances are the same.

248   **Results**

249   **1. Craniodental and postcranial characters contain similar levels of homoplasy and retain similar**

250   **amounts of synapomorphy**

251   Statistics and test results for all data sets are given in Table 1, and we distil these further in Table 2

252   and Figure 2. We found no significant difference in the level of craniodental/postcranial ensemble

253   consistency index (CI) across all 81 data sets (Wilcoxon test paired V = 3453, p = 0.5645). Despite

254 remarkably similar medians (100 and 97) and overall distributions, the median number of

255 craniodental and postcranial characters differed non-significantly (V = 3112.5, p = 0.5747). We

256 therefore compared the residual CI values from a linear regression of CI on the log of the number of

257 characters and the log of the number of taxa, plus their interaction. This model was significant

258 overall (p < 2.2e-16), but none of the individual slope terms was significant (p > 0.1 in all cases).

259 Residuals from this model likewise showed no significant difference between partitions (paired t =

260 0.38211, p = 0.7032). Likewise, we found no significant difference in the level of

261 craniodental/postcranial ensemble retention index (RI) across all 81 data sets (V = 3717.5, p-value =

262 0.1437). Similarly, the residuals from the regression of RI onto the number of taxa, number of

263 characters and their interaction (p = 3.668e-12 overall, but with no significant slopes for individual

264 terms) also showed no difference between partitions. Neither partition of the data can be deemed

265 superior on the basis of these ensemble indices of internal consistency.

266

267 **2.** **Half of craniodental and postcranial data partitions yield significantly different trees**

268 Previous work on a broad sample of tetrapod matrices revealed significant incongruence between

269 craniodental and postcranial character partitions about one time in three, as measured by both the

270 incongruence relationship difference (IRD) test of Ruta & Wills (2016) and the incongruence length

271 difference (ILD) test (Mickevich and Farris, 1981; Farris et al., 1995). Here, we report that 50% of

272 dinosaur matrices yielded significantly (p<0.05) incongruent trees according to the IRD test for

273 nearest neighbours using matching distances ($IRD_{NND+matching}$) and 54% for the IRD test using the RF

274 distances (Robinson and Foulds, 1981) (IRD_{NND+RF}). Moreover, the $IRD_{NND+matching}$ and IRD_{NND+RF} values

275 were closely correlated (r_s = 0.649, p = 8.999e-14). We note that analyses of partitions reached the

276 limit of 1,000 trees for nearest neighbour (NND) computations in 47% of runs (see Supplementary

277 Materials 3). We therefore also report the results of IRD tests using majority rule consensus trees

278 derived from up to 10,000 optimal source trees. Inevitably, consensus trees cannot reflect accurately

279   the diversity of relationships within a set of source trees (Mounce et al., 2016) but they do permit

280   tests that incorporate all source trees more readily. The consensus results were closely similar to

281   those for the nearest neighbour tests: 63% of matrices were significantly incongruent using

282    $IRD_{MR+matching}$ and 60% were incongruent using IRD_{MR+RF}). Moreover, the consensus results were

283   strongly and significantly correlated with the NND results for both the $IRD_{MR+matching}$ (r_s = 0.688, p =

284   7.31e-16) and IRD_{MR+RF} (r_s = 0.560, p = 6.412e-10). The rate of significance (33% at p < 0.05) for the

285   ILD test was lower than that for variants of the IRD, and similar to that observed for tetrapods

286   overall (Mounce et al., 2016). Nevertheless, we found significant correlation between the ILD p value

287   and the p values for both the $IRD_{matching}$ (r_s = 0.308, p = 0.001) and the IRD_{RF} (r_s = 0.297, p = 0.002)

288   using NNDs.

289   Mounce *et al*. (2016) tested empirically whether the rate of null rejection was influenced by several

290   data matrix parameters. Here we used logistic regression to determine the outcome of each of our

291   five partition homogeneity tests (significant or not with p < 0.05) as a function of the overall number

292   of taxa, overall number of characters (both partitions), the interaction between characters and taxa,

293   the difference in partition size (scaled relative to the total number of characters), total amount of

294   missing data, the difference in the percentage of missing data between partitions and the higher

295   taxonomic group. Results are summarised in the supplementary materials. For IRD_{NND+RF}, only the

296   total number of characters and total number of taxa were retained in a minimum adequate model

297   (McDonald et al., 2012) selected using the AIC. For IRD_{MR+RF}, these parameters plus their interaction

298   and the difference in partition size (number of characters) was retained in the MAM. Hence, as

299   reported by Mounce *et al.* (2016), tests based upon symmetrical differences (Robinson and Foulds,

300   1981) are influenced, at least in part, by data matrix dimensions. By contrast, for the $IRD_{NND+matching}$

301   (based upon nearest neighbour matching distances), no matrix parameters were retained in the

302   MAM, but higher taxonomic group was highly significant. Similarly, taxonomic group was highly

303   significant for the $IRD_{MR+matching}$ (based upon majority rule matching distances), and the difference in

304   partition size was also retained by the AIC (although p>0.05).

305

3. Levels of incongruence vary significantly across major dinosaur groups

307 Of our 46 saurischian matrices, 31 (67%) showed significant (p < 0.05) incongruence using

308 $IRD_{NND+matching}$, compared with 9 from 35 (26%) ornithischian matrices: a significant difference in the

309 rate of null rejection (G = 20.1841, p = 0.00046). A similarly significant difference was observed for

310 the $IRD_{MR+matching}$ (G = 19.3857, p = 0.00066). For the IRD using Robinson Foulds distances, by

311 contrast, differences across saurischian and ornithischian matrices were non-significant (IRD_{NND+RF} G

312 = 6.7962, p = 0.14706: IRD_{MR+RF} G = 5.4241, p = 0.24648). At a finer taxonomic level (specifically

313 assigning trees to Theropda (e.g., Fig. 2), Sauropodomorpha (e.g., Fig. 3), Cerapoda, Ornithopoda,

314 and Thyreophora) there were highly significant differences in the rate of null rejection using the

315 matching distance (p = 0.00046 for the $IRD_{NND+matching}$: p = 0.00066 for $IRD_{MR+matching}$) no difference

316 using the symmetrical difference (RF) variants of the tests (Table 2) (Fig. 4). This difference between

317 'RF' and 'Matching' variants of the IRD test is consistent with that observed for the modelling of

318 outcomes above (i.e., taxonomic group was significant in determining the binary outcome of

319 $IRD_{matching}$ variants, but not for the IRD_{RF}). Rates of partition incongruence are relatively high in the

320 Sauropoda (53% for the $IRD_{NND+matching}$), Theropoda (66%) and Ceropoda (43%) compared with the

321 Ornithopoda (20%) and Thyreophora (14%). A similar hierarchy of outcomes pertained for the other

322 tests.

323

4. Craniodental and postcranial characters produce trees equally similar to the entire data set, but

with strong biases across major groups

326 Our second set of tests sought to determine whether the most parsimonious trees from the entire

327 matrix were more similar to those from the craniodental or postcranial partitions. The results from

328 these were in strong agreement. Considering the NNDs for matching distances across all 81 matrices,

329   42 were closer (using the matching distances for nearest neighbours: NND+matching) to the

330   postcranial partition, while 39 were closer to the craniodental partition (a non-significant bias:

331   binomial test p = 0.8243). Within Saurischia and Ornithischia, however, the biases were highly

332   significant, but in opposite directions (G = 6.42421, p = 0.01126). For Saurischia, trees from the

333   entire data set were most often more similar to those from the postcranial partition 29 cases) than

334   to those from the cranium (17 cases) (binomial p = 0.1038). For Ornithischia, by contrast, trees from

335   the whole data set were more often most to those from the craniodental partition (22 cases) than to

336   those from the postcranium (13 cases) (p = 0.1755). When partitioned into five groups as above,

337   there was also a significant difference in which partition was most similar to the entire matrix across

338   groups (G = 8.7347, p = 0.006209), with the bias for Sauropoda (13 postcranial verus 4 craniodental)

339   being the most striking. Similar findings were made for the other three tests (Tables 1 and 2).

340   These biases were much less marked if comparisons were restricted to only those data sets

341   for which there was a significant difference (p < 0.05) between the distributions of distances (entire

342   to craniodental versus entire to postcraniodental) according to the Mann-Whitney U test.

343   Considering NND+matching distances, 31 data sets favoured the craniodental partition and 33 the

344   postcranial partition (binomial p = 0.9007). Moreover, there was no longer a significant bias in favour

345   of postcranial characters for the Saurischia (23 out of 33: p = 0.0308) or in favour of craniodental

346   characters for the Ornithischia (17 out of 31: p = 0.7201) (overall G = 3.98451, p = 0.04592).

347   However, for the partition into Theropda, Sauropodomorpha, Cerapoda, Ornithopoda and

348   Thyreophora, we still marginally rejected the null that groups behave identically (G = 9.29294, p =

349   0.05418027). Similar findings for the other three tests are summarised in Table 2.

350

351

352

353   **Discussion**

354   **Implications for dinosaur phylogeny**

355   Our analyses of 81 published matrices demonstrate empirically that the relationships of dinosaurs

356   inferred from craniodental or postcranial characters in isolation differ significantly ($p < 0.05$) from

357   each other about half of the time. This is much more often than similar partitions for tetrapods in

358   general (about 1 in 3:(Mounce et al., 2016). At the same time, we find similar levels of homoplasy (as

359   measured by the ensemble consistency index: CI) and retained synapomophy (RI) in craniodental

360   and postcranial character partitions across all dinosaurs. Similarly, when character sample sizes are

361   controlled, the relationships inferred from either partition are equally congruent with those from

362   the entire matrix. Hence, there is no reason to prefer characters sampled from one partition versus

363   another across dinosaurs as a whole, and we concur with general recommendations to sample

364   characters widely from all anatomical regions in accordance with the principle of total evidence

365   (Kluge, 1989; Gatesy et al., 1999; Gatesy and Arctander, 2000; Mounce et al., 2016). However, we

366   also observe marked differences in levels of incongruence across major dinosaur groups, being

367   significantly higher in Ornithischia than Saurischia.

368   Homoplasy is always a problem for phylogenetic inference, but is least troublesome when

369   homoplastic states approximate to a random distribution across taxa (in which case it largely

370   contributes noise). Homoplasy is more problematic when it is correlated across complexes of

371   characters, especially when this occurs at higher frequencies. The morphological phylogeny of

372   mammals appears to have been subject to such problems. Phylogenomic trees (Dolphin, K. et al.

373   2000,) overturned many of the groups (e.g., Ungulata and Insectivora) that had emerged from nearly

374   all previous analyses of morphological characters over the preceding decades. Most problematic of

375   all are cases where correlated homoplasy is concentrated within a particular region of the body, and

376   especially where available data are limited to such regions. The teeth of mammals appear to be

377   subject to such convergence (Goswami A et al. 2011), with many aspects of their form changing in

378   concert and being controlled by a relatively small number of genes (Castelin et al., 2017). This is

379   singularly unfortunate for those studying mammalian evolution, since the high preservation

380   potential of teeth means that they dominate the fossil record. The fossil record of dinosaurs is also

381   biased, with sauropodomorphs and theropods being known predominantly from their postcranial

382   remains (skulls are often fragmentary or not recovered), and ceratopsians being more often

383   described from their much more massive skulls (Evans, D.C. and Ryan, M.J. 2015,). Our sample of

384   matrices suggest that the number of coded characters strongly reflects these differences. However,

385   it remains unclear whether this is solely a function of the available material, or whether systematists

386   preferentially code or more finely atomise characters from these regions. Whatever the case, we do

387   not find a higher concentration of homoplasy in either partition, whether across all dinosaurs or

388   within major clades. Moreover, although craniodental and postcranial characters often yield

389   significantly conflicting trees, we find no evidence that one partition is more likely to be congruent

390   with the "total evidence" tree than the other.

391   **Interpreting incongruence**

392   The inference of significantly different trees from craniodental and postcranial character partitions

393   can be understood in terms of divergent selective pressures operating on different regions of the

394   body (Gould and Gould, 1977; J, 1993; Kemp and Kemp, 2005; Lue et al., 2010). This results in

395   different rates and patterns of character evolution (Mitteroecker and Bookstein, 2007; Klingenberg,

396   2008), in addition to distinct patterns of homoplasy. Anatomical modules are commonly recognised

397   in studying the evolution of form (Mitteroecker and Bookstein, 2007; Cardini and Elton, 2008;

398   Klingenberg, 2008; Lue et al., 2010; Goswami et al., 2011; Hopkins and Lidgard, 2012; Cardini and

399   Polly, 2013; Goswami et al., 2015), and it is reasonable to suppose that such modules will contain

400   phylogenetic characters that are more congruent with one another than with characters from other

401   modules (Clarke and Middleton, 2008)

402   The tetrapod skull is variously decoupled from the skeleton of the body, both biomechanically and in

403   terms of the selective pressures operating upon it (Ji et al., 1999; Koski, 2007; Mitteroecker and

404   Bookstein, 2007). However, this decoupling is particularly marked in the non-avian dinosaurs

405   (Mounce et al., 2016). The long necks of sauropodomorphs effect the greatest biomechanical

406   decoupling between the skull and the body, and sauropods unsurprisingly have one of the highest

407   levels of significant incongruence. Ornithischia, by contrast, show much lower levels of incongruence

408   overall: especially low (14%) in the thyreophorans.

409   Anatomical modules are typically envisaged as comprising physically proximate sets of characters or

410   aspects of form. However, particular selective pressures might result in the coordinated evolution of

411   suites of characters widely distributed across the body (Gardiner et al., 2011; Abourachid and

412   Hoefling, 2012; Godefroit et al., 2013). For example, a mode of predation or scavenging favoured by

413   many theropods entailed bracing a carcass with a back leg whilst ripping with powerful jaws and a

414   strong neck (Rayfield, 2004). This manner of feeding evolved in at least three large theropod clades,

415   and entailed coordinated changes in the limbs and skull (Snively et al., 2006; Snively and Russell,

416   2007a; Snively and Russell, 2007b; Hone and Rauhut, 2010). Similarly, the massive skulls of many

417   ceratopsians were braced into the body and pectoral girdle consistent with their ability to face and

418   ward off predators, and this may have effected other coordinated changes in the vertebral column

419   and back limbs.

420   In the infancy of molecular phylogenetics, trees were often inferred from single genes (Gatesy, J. and

421   Arctander, P. 2000), and it was not uncommon for the trees derived from different genes to be

422   strikingly at odds (Gatesy, J. and Arctander, P. 2000). In addition, molecular trees often differed

423   markedly from those inferred from morphology. A more cautious, combinatorial approach initially

424   evolved, therefore, along with partition tests designed to ensure homogeneity of signal (Mounce,

425   R.C.P. et al. 2016). This agglomerative approach may have been a function of the manner in which

426   data originally became available, with systematists exploring incongruent signals at a fine level of

427 granularity. The ascendance of phylogenomic analyses has brought its own bioinformatic challenges,

428 but all approaches seek to derive trees from increasingly inclusive data sets. More philosophically, a

429 consensus has emerged in favour of the principle of total evidence (Kluge, 1989): the procedure by

430 which all available character data are combined into a single matrix and analysis. One reason for this

431 is the phenomenon of 'hidden support' (Gatesy et al., 1999; Gatesy and Arctander, 2000), whereby

432 signals that are weak and therefore hidden within individual character partitions become dominant

433 when all data are analysed together. Various tests for partition homogeneity (Farris et al., 1994,

434 1995; Dolphin et al., 2000) will tend to return significant results in precisely those circumstances in

435 which support is hidden, and such tests are therefore no longer commonly used to preclude the

436 combination of data sources in this manner (Kluge, 1989; Gatesy et al., 1999; Gatesy and Arctander,

437 2000; Wahlberg et al., 2005; O'Leary and Gatesy, 2008; Padial et al., 2010; Damgaard, 2012; Mounce

438 et al., 2016). However, while molecular systematics has retained and elaborated the notion that

439 different suites of characters within large molecular matrices might be most effectively modelled

440 with different rate parameters (e.g., different sets of genes, and different codon positions),

441 morphological data are rarely treated in this manner. Moreover, there is relatively little quantitative

442 empirical data on the sorts of morphological characters that might be most useful for resolving

443 relationships at different hierarchical levels within a phylogeny, or for radiations of different ages.

444 This is partly because morphology is less likely to be constrained to evolve in a clock-like manner

445 throughout a tree (or to change it's rate in a manner amenable to modelling), and partly because

446 there are no standard repositories for morphological characters that enable data to be archived,

447 retrieved, coded and compared in a uniform manner. Despite this, trees are often inferred from

448 relatively restricted character sets (Sanchez-Villagra and Williams, 1998; Arratia, 2009; Song and

449 Bucheli, 2010; Mounce et al., 2016). In fossil taxa, this may reflect preservational biases, particularly

450 those favouring hard part preservation (Sansom et al., 2010; Sansom and Wills, 2013; Pattinson et

451 al., 2015; Sansom et al., 2017).

452

Conclusions

1. Across our sample of 81 data sets, practitioners have abstracted slightly fewer characters from the skull (mean 115, median 100) than the rest of the skeleton (mean 132 and median 97) overall. However, this masks particular biases in major dinosaur groups: markedly and significantly ($p < 0.02$ in all cases) in favor of postcranial characters in Sauropodomorpha and Theropoda and in favor of postcranial characters in Ornithopoda, Thyreophora and Cerapoda.

2. The overall frequency of significant ($p < 0.05$) incongruence between dinosaur trees inferred from craniodental and postcranial characters was about 50% for variants of the Incongruence Relationship Difference (IRD) test (Ruta and Wills, 2016); (Mounce et al., 2016). This was substantially higher than that previously reported (30%) for tetrapod clades in general. The Incongruence Length Difference Test (ILD) reported significant incongruence in 33% of cases: comparable to the level seen across tetrapods hitherto (Mounce et al., 2016).

3. Despite the high frequency of incongruence overall, rates of significance were heterogeneously distributed across major dinosaur groups, being highest (a mean of 48% for variants of the IRD) in Theropoda and lowest (25%) in Thyreophora. A number of factors have been purported to influence the outcome of the IRD test, notably the data matrix dimensions, relative partition sizes and the amount and distribution of missing entries (Mounce et al., 2016). We nevertheless found similar heterogeneity in rates of significant incongruence across groups when these variables were modelled out.

4. Tests to determine which partition (craniodental or postcranial) were most congruent with trees inferred from the entire character set were equivocal overall: equal numbers favoured the two partitions once differences in sample size were controlled for. However, there were significant asymmetries in many groups, with the bias for Sauropoda (13 postcranial verus 4 craniodental) being the most striking.

477   5. We show that our partition tests are not sensitive to the size of the tree (number of terminals), to

478   differences in partition size (differences in the number of characters), to the overall amount of

479   missing data in the matrix, the difference in the amount of missing data between partitions or the

480   difference in levels of homoplasy or retained synapomorphy between partitions. We also

481   demonstrate that there are similar levels of homoplasy and retained synapomorphy between

482   partitions overall. Incongruence therefore reflects differences in *patterns* of homoplasy between

483   partitions, which may itself be a function of modularity and mosaic evolution. We note that

484   incongruence is greatest in those groups (Sauropodomorpha and Theropoda) in which the

485   biomechanical decoupling between head and body is greatest.

486   6. Our results demonstrate clearly that phylogenies of dinosaurs inferred from craniodental and

487   postcranial characters differ significantly much more often than expected. We therefore make the

488   straightforward recommendation that characters should be sampled as broadly as possible from

489   across all body regions. This accords with the theoretical principle of total evidence (Kluge, 1989;

490   Gatesy et al., 1999; Gatesy and Arctander, 2000), as well as our empirical findings for tetrapods in

491   general (Mounce et al., 2016) and mammals in particular (Sansom and Wills, 2017).

492

References

494   Abourachid A, Hoefling E. 2012. The legs: a key to bird evolutionary success. Journal of Ornithology,

495   153:S193-S198.

496   Allain R, Xaisanavong T, Richir P, Khentavong B. 2012. The first definitive Asian spinosaurid

497   (Dinosauria: Theropoda) from the early cretaceous of Laos. Naturwissenschaften, 99:369-377.

498   Araujo R, Castanhinha R, Martins RMS, Mateus O, Hendrickx C, Beckmann F, Schell N, Alves LC. 2013.

499   Filling the gaps of dinosaur eggshell phylogeny: Late Jurassic Theropod clutch with embryos

500   from Portugal. Scientific Reports, 3.

501   Arbour VM, Burns ME, Sullivan RM, Lucas SG, Cantrell AK, Fry J, Suazo TL. 2014. A new ankylosaurid

502   dinosaur from the Upper Cretaceous (Kirtlandian) of New Mexico with implications for

503   ankylosaurid diversity in the Upper Cretaceous of western North America. PLoS One,

504   9:e108804.

505   Arbour VM, Currie PJ. 2013. Euoplocephalus tutus and the Diversity of Ankylosaurid Dinosaurs in the

506   Late Cretaceous of Alberta, Canada, and Montana, USA. PLoS ONE, 8:1-39.

507   Archie JW. 1989. Homoplasy Excess Ratios - New Indexes for measuring levels of Homoplasy in

508   Phylogenetic Systematics and a Critique of the Consistency Index. Systematic Zoology, 38:253-

509   269.

510   Arratia G. 2009. Identifying patterns of diversity of the actinopterygian fulcra. Acta Zoologica,

511   90:220-235.

512   Barker FK, Lutzoni FM. 2002. The utility of the incongruence length difference test. Systematic

513   Biology, 51:625-637.

514   Barrett PM, Butler RJ, Mundil R, Scheyer TM, Irmis RB, Sanchez-Villagra MR. 2014. A

515   palaeoequatorial ornithischian and new constraints on early dinosaur diversification.

516   Proceedings of the Royal Society B-Biological Sciences, 281.

517   Boyd CA. 2015. The systematic relationships and biogeographic history of ornithischian dinosaurs.

518   Peerj, 3.

519   Boyd CA, Pagnac DC. 2015. Insight on the anatomy, systematic relationships, and age of the Early

520   Cretaceous ankylopollexian dinosaur Dakotadon lakotaensis. Peerj, 3.

521   Brocklehurst N, Froebisch J. 2014. Current and historical perspectives on the completeness of the

522   fossil record of pelycosaurian-grade synapsids. Palaeogeography Palaeoclimatology

523   Palaeoecology, 399:114-126.

524   Brown CM, Boyd CA, Russell AP. 2011. A new basal ornithopod dinosaur (Frenchman Formation,

525   Saskatchewan, Canada), and implications for late Maastrichtian ornithischian diversity in

526   North America. Zoological Journal of the Linnean Society, 163:1157-1198.

527   Brusatte SL, Benson RBJ. 2013. The systematics of Late Jurassic tyrannosauroid theropods from

528   Europe and North America. Acta Palaeontologica Polonica, 58:47-54.

529   Brusatte SL, Lloyd GT, Wang SC, Norell MA. 2014. Gradual Assembly of Avian Body Plan Culminated

530   in Rapid Rates of Evolution across the Dinosaur-Bird Transition. Current Biology, 24:2386-

531   2392.

532   Burns ME, Currie PJ. 2014. External and internal structure of ankylosaur (dinosauria, ornithischia)

533   osteoderms and their systematic relevance. Journal of Vertebrate Paleontology, 34:835-851.

534   Burns ME, Currie PJ, Sissons RL, Arbour VM. 2011. Juvenile specimens of Pinacosaurus grangeri

535   Gilmore, 1933 (Ornithischia: Ankylosauria) from the Late Cretaceous of China, with comments

536   on the specific taxonomy of Pinacosaurus. Cretaceous Research, 32:174-186.

537   Butler RJ, Jin L, Chen J, Godefroit P. 2011. The postcranial osteology and phylogenetic position of the

538   small ornithischian dinosaur Changchunsaurus parvus from the Quantou Formation

539   (Cretaceous: Aptian-Cenomanian) of Jilin Province, north-eastern China. Palaeontology,

540   54:667-683.

541   Carballido JL, Pol D, Parra Ruge ML, Padilla Bernal S, Paramo-Fonseca ME, Etayo-Serna F. 2015. A

542   new early Cretaceous brachiosaurid (Dinosauria, Neosauropoda) from northwestern

543   Gondwana (Villa de Leiva, Colombia). Journal of Vertebrate Paleontology, 35:c980505-

544   c980505.

545   Carballido JL, Rauhut OWM, Pol D, Salgado L. 2011. Osteology and phylogenetic relationships of

546   Tehuelchesaurus benitezii (Dinosauria, Sauropoda) from the Upper Jurassic of Patagonia.

547   Zoological Journal of the Linnean Society, 163:605-662.

548   Carballido JL, Sander PM. 2014. Postcranial axial skeleton of Europasaurus holgeri (Dinosauria,

549   Sauropoda) from the Upper Jurassic of Germany: implications for sauropod ontogeny and

550   phylogenetic relationships of basal Macronaria. Journal of Systematic Palaeontology, 12:335-

551   387.

552   Cardini A, Elton S. 2008. Does the skull carry a phylogenetic signal? Evolution and modularity in the

553   guenons. Biological Journal of the Linnean Society, 93:813-834.

554   Cardini A, Polly PD. 2013. Larger mammals have longer faces because of size-related constraints on

555   skull form. Nature Communications, 4.

556   Castelin M, Williams ST, Buge B, Maestrati P, Lambourdiere J, Ozawa T, Utge J, Couloux A, Alf A,

557   Samadi S. 2017. Untangling species identity in gastropods with polymorphic shells in the genus

558   Bolma Risso, 1826 (Mollusca, Vetigastropoda). European Journal of Taxonomy, 288:1-21.

559   Cau A, Dalla Vecchia FM, Fabbri M. 2013. A thick-skulled theropod (Dinosauria, Saurischia) from the

560   Upper Cretaceous of Morocco with implications for carcharodontosaurid cranial evolution.

561   Cretaceous Research, 40:251-260.

562   Choiniere JN, Clark JM, Forster CA, Norell MA, Eberth DA, Erickson GM, Chu H, Xu X. 2014. A juvenile

563   specimen of a new coelurosaur (Dinosauria: Theropoda) from the Middle Late Jurassic

564   Shishugou Formation of Xinjiang, People's Republic of China. Journal of Systematic

565   Palaeontology, 12:177-215.

566   Clarke JA, Middleton KM. 2008. Mosaicism, modules, and the evolution of birds: Results from a

567   Bayesian approach to the study of morphological evolution using discrete character data.

568   Systematic Biology, 57:185-201.

569   Cobbett A, Wilkinson M, Wills MA. 2007. Fossils impact as hard as living taxa in parsimony analyses

570   of morphology. Syst Biol, 56:753-766.

571   Coria RA, Moly JJ, Reguero M, Santillana S, Marenssi S. 2013. A new ornithopod (Dinosauria;

572   Ornithischia) from Antarctica. Cretaceous Research, 41:186-193.

573   Cunningham CW. 1997. Can three incongruence tests predict when data should be combined?

574   Molecular Biology and Evolution, 14:733-740.

575   D'Emic MD. 2013. Revision of the sauropod dinosaurs of the Lower Cretaceous Trinity Group,

576   southern USA, with the description of a new genus. Journal of Systematic Palaeontology,

577   11:707-726.

578 Dabney J, Knapp M, Glocke I, Gansauge M-T, Weihmann A, Nickel B, Valdiosera C, Garcia N, Paeaebo

579   S, Arsuaga J-L, *et al.* 2013. Complete mitochondrial genome sequence of a Middle Pleistocene

580   cave bear reconstructed from ultrashort DNA fragments. Proceedings of the National

581   Academy of Sciences of the United States of America, 110:15758-15763.

582 Daley AC, Drage HB. 2016. The fossil record of ecdysis, and trends in the moulting behaviour of

583   trilobites. Arthropod Structure & Development, 45:71-96.

584 Damgaard J. 2012. Phylogeny of the semiaquatic bugs (Hemiptera-Heteroptera, Gerromorpha).

585   TreeBASE.

586 Davies TW, Bell MA, Goswami A, Halliday TJD. 2017. Completeness of the eutherian mammal fossil

587   record and implications for reconstructing mammal evolution through the

588   Cretaceous/Paleogene mass extinction. Paleobiology, 43:521-536.

589 Dolphin K, Belshaw R, Orme CDL, Quicke DLJ. 2000. Noise and incongruence: Interpreting results of

590   the incongruence length difference test. Molecular Phylogenetics and Evolution, 17:401-406.

591 Dunhill AM, Benton MJ, Twitchett RJ, Newell AJ. 2012. Completeness of the fossil record and the

592   validity of sampling proxies at outcrop level. Palaeontology, 55:1155-1175.

593 Eddy DR, Clarke JA. 2011. New Information on the Cranial Anatomy of Acrocanthosaurus atokensis

594   and Its Implications for the Phylogeny of Allosauroidea (Dinosauria: Theropoda). Plos One, 6.

595 Evans DC, Ryan MJ. 2015. Cranial Anatomy of Wendiceratops pinhornensis gen. et sp nov., a

596   Centrosaurine Ceratopsid (Dinosauria: Ornithischia) from the Oldman Formation (Campanian),

597   Alberta, Canada, and the Evolution of Ceratopsid Nasal Ornamentation. Plos One, 10.

598 Evans DC, Schott RK, Larson DW, Brown CM, Ryan MJ. 2014. The oldest North American

599   pachycephalosaurid and the hidden diversity of small-bodied ornithischian dinosaurs (vol 4,

600   3289, 2013)). Nature Communications, 5.

601 Evers S, Rauhut OWM, Milner AC, McFeeters B, Allain R. 2015. A reappraisal of the morphology and

602   systematic position of the theropod dinosaur Sigilmassasaurus from the "middle" Cretaceous

603   of Morocco. PeerJ, 3.

Systematic Biology

604   Fan R-Y, Uchman A, Gong Y-M. 2017. From morphology to behaviour: Quantitative morphological

605   study of the trace fossil Helminthorhaphe. Palaeogeography Palaeoclimatology Palaeoecology,

606   485:946-955.

607   Fanti F, Cau A, Cantelli L, Hassine M, Auditore M. 2015. New Information on Tataouinea hannibalis

608   from the Early Cretaceous of Tunisia and Implications for the Tempo and Mode of

609   Rebbachisaurid Sauropod Evolution. Plos One, 10.

610   Fanti F, Currie PJ, Badamgarav D. 2012. New Specimens of Nemegtomaia from the Baruungoyot and

611   Nemegt Formations (Late Cretaceous) of Mongolia. Plos One, 7.

612   Farke AA, Maxwell WD, Cifelli RL, Wedel MJ. 2014. A Ceratopsian Dinosaur from the Lower

613   Cretaceous of Western North America, and the Biogeography of Neoceratopsia. Plos One, 9.

614   Farke AA, Ryan MJ, Barrett PM, Tanke DH, Braman DR, Loewen MA, Graham MR. 2011. A new

615   centrosaurine from the Late Cretaceous of Alberta, Canada, and the evolution of parietal

616   ornamentation in horned dinosaurs. Acta Palaeontologica Polonica, 56:691-702.

617   Farke AA, Sertich JJW. 2013. An Abelisauroid Theropod Dinosaur from the Turonian of Madagascar.

618   Plos One, 8.

619   Farris JS, Kallersjo M, Kluge AG, Bult C. 1994. Testing significance of incongruence. Cladistics, 10:315-

620   319.

621   Farris JS, Kallersjo M, Kluge AG, Bult C. 1995. Constructing a significance test for incongruence.

622   Systematic Biology, 44:570-572.

623   Foth C, Tischlinger H, Rauhut OWM. 2014. New specimen of Archaeopteryx provides insights into the

624   evolution of pennaceous feathers. Nature, 511:79-U421.

625   Gainett G, Sharma PP, Pinto-da-Rocha R, Giribet G, Willemart RH. 2014. Walk it off: predictive power

626   of appendicular characters toward inference of higher-level relationships in Laniatores (

627   Arachnida: Opiliones). Cladistics, 30:120-138.

628   Gardiner JD, Codd JR, Nudds RL. 2011. An association between ear and tail morphologies of bats and

629   their foraging style. Canadian Journal of Zoology-Revue Canadienne De Zoologie, 89:90-99.

630   Gatesy J, Arctander P. 2000. Hidden morphological support for the phylogenetic placement of

631   Pseudoryx nghetinhensis with bovine bovids: A combined analysis of gross anatomical

632   evidence and DNA sequences from five genes. Systematic Biology, 49:515-538.

633   Gatesy J, O'Grady P, Baker RH. 1999. Corroboration among data sets in simultaneous analysis:

634   Hidden support for phylogenetic relationships among higher level artiodactyl taxa. Cladistics,

635   15:271-313.

636   Godefroit P, Cau A, Hu D-Y, Escuillie F, Wu W, Dyke G. 2013. A Jurassic avialan dinosaur from China

637   resolves the early phylogenetic history of birds. Nature, 498:359-362.

638   Gorscak E, Connor PMO, Stevens NJ, Roberts EM. 2014. The basal titanosaurian Rukwatitan

639   bisepultus (Dinosauria, Sauropoda) from the Middle Cretaceous Galula Formation, Rukwa Rift

640   Basin, southwestern Tanzania. Journal of Vertebrate Paleontology, 34:1133-1154.

641   Goswami A, Binder WJ, Meachen J, O'Keefe FR. 2015. The fossil record of phenotypic integration and

642   modularity: A deep-time perspective on developmental and evolutionary dynamics.

643   Proceedings of the National Academy of Sciences of the United States of America, 112:4891-

644   4896.

645   Goswami A, Milne N, Wroe S. 2011. Biting through constraints: cranial morphology, disparity and

646   convergence across living and fossil carnivorous mammals. Proceedings of the Royal Society B-

647   Biological Sciences, 278:1831-1839.

648   Gould SJ, Gould SJ. 1977. Ontogeny and phylogeny.

649   Han F-L, Barrett PM, Butler RJ, Xu X. 2012. Postcranial anatomy of Jeholosaurus shangyuanensis

650   (Dinosauria, Ornithischia) from the Lower Cretaceous Yixian formation of China. Journal of

651   Vertebrate Paleontology, 32:1370-1395.

652   He Y, Makovicky PJ, Wang K, Chen S, Sullivan C, Han F, Xu X. 2016. A New Leptoceratopsid

653   (Ornithischia, Ceratopsia) with a Unique Ischium from the Upper Cretaceous of Shandong

654   Province, China. Figshare.

655   Hone DWE, Rauhut OWM. 2010. Feeding behaviour and bone utilization by theropod dinosaurs.

656   Lethaia, 43:232-244.

657   Hopkins MJ, Lidgard S. 2012. Evolutionary mode routinely varies among morphological traits within

658   fossil species lineages. Proceedings of the National Academy of Sciences of the United States

659   of America, 109:20520-20525.

660   Houde P. 1994. Evolution of the Heliornithidae - reciprocal illumination by morphology,

661   biogeography and DNA hybridization (Aves, Gruiformes). . Cladistics, 10:1-19.

662   Hsiang AY, Field DJ, Webster TH, Behlke ADB, Davis MB, Racicot RA, Gauthier JA. 2015. The origin of

663   snakes: revealing the ecology, behavior, and evolutionary history of early snakes using

664   genomics, phenomics, and the fossil record. Bmc Evolutionary Biology, 15.

665   Hu H, O'Connor JK, Zhou Z. 2015. A New Species of Pengornithidae (Aves: Enantiornithes) from the

666   Lower Cretaceous of China Suggests a Specialized Scansorial Habitat Previously Unknown in

667   Early Birds. Plos One, 10.

668   Ignacio Canale J, Emilio Novas F, Pol D. 2015. Osteology and phylogenetic relationships of

669   Tyrannotitan chubutensis Novas, de Valais, Vickers-Rich and Rich, 2005 (Theropoda:

670   Carcharodontosauridae) from the Lower Cretaceous of Patagonia, Argentina. Historical

671   Biology, 27:1-32.

672   Ignacio Ruiz-Omenaca J, Ignacio Canudo J, Cuenca-Bescos G, Cruzado-Caballero P, Manuel Gasca J,

673   Moreno-Azanza M. 2012. A new basal ornithopod dinosaur from the Barremian of Galve,

674   Spain. Comptes Rendus Palevol, 11:435-444.

675   J MS. 1993. The theory of evolution, CHAPTER?

676   Jablonski D. 2005. Evolutionary innovations in the fossil record: the intersection of ecology,

677   development, and macroevolution. J. Exp. Zool., 304B:504-519.

678   Ji Q, Luo ZX, Ji SA. 1999. A Chinese triconodont mammal and mosaic evolution of the mammalian

679   skeleton. Nature, 398:326-330.

680   Kearney M, Clark JM. 2003. Problems due to missing data in phylogenetic analyses including fossils:

681   A critical review. Journal of Vertebrate Paleontology, 23:263-274.

682   Kemp TS, Kemp TS. 2005. The origin and evolution of mammals.

683   Klingenberg CP. 2008. Morphological Integration and Developmental Modularity. Annual Review of

684   Ecology Evolution and Systematics, p. 115-132.

685   Kluge AG. 1989. A concern for evidence and a phylogenetic hypothesis of relationships among

686   Epicrates (Boidae, Serpentes). Systematic Zoology, 38:7-25.

687   Kluge AG, Farris JS. 1969. Quantitative phyletics and evolution of anurans. Systematic Zoology, 18:1-

688   32.

689   Koski K. 2007. The mandibular complex. Eur. J. Orthod, 29:118-123.

690   Ksepka DT, Boyd CA. 2012. Quantifying historical trends in the completeness of the fossil record and

691   the contributing factors: an example using Aves. Paleobiology, 38:112-125.

692   Lacovara K. 2014. Supplementary Figure 18, Dreadnoughtus hind limb - Lacovara et al., 2014, A

693   Gigantic, Exceptionally Complete Titanosaurian Sauropod Dinosaur from Southern Patagonia,

694   Argentina, Scientific Reports. Figshare.

695   Lamanna MC, Sues H-D, Schachner ER, Lyson TR. 2014. A New Large-Bodied Oviraptorosaurian

696   Theropod Dinosaur from the Latest Cretaceous of Western North America. Plos One, 9.

697   Lee MSY, Cau A, Naish D, Dyke GJ. 2014. Data from: Morphological clocks in palaeontology, and a

698   mid-Cretaceous origin of crown Aves. Dryad.

699   Li L-G, Li D-Q, You H-L, Dodson P. 2014. A New Titanosaurian Sauropod from the Hekou Group

700   (Lower Cretaceous) of the Lanzhou-Minhe Basin, Gansu Province, China. Plos One, 9.

701   Lieberman BS. 2002. Phylogenetic biogeography with and without the fossil record: gauging the

702   effects of extinction and paleontological incompleteness. Palaeogeography Palaeoclimatology

703   Palaeoecology, 178:39-52.

704   Lloyd GT. 2018. Graeme T Lloyd morphological matrices.

705   Loewen MA, Irmis RB, Sertich JJW, Currie PJ, Sampson SD. 2013. Tyrant Dinosaur Evolution Tracks

706   the Rise and Fall of Late Cretaceous Oceans. Plos One, 8.

707   Longrich NR. 2011. Titanoceratops ouranos, a giant horned dinosaur from the late Campanian of

708   New Mexico. Cretaceous Research, 32:264-276.

709   Longrich NR. 2014. The horned dinosaurs Pentaceratops and Kosmoceratops from the upper

710   Campanian of Alberta and implications for dinosaur biogeography. Cretaceous Research,

711   51:292-308.

712   Lopardo L, Hormiga G. 2015. Out of the twilight zone: phylogeny and evolutionary morphology of

713   the orb-weaving spider family Mysmenidae, with a focus on spinneret spigot morphology in

714   symphytognathoids (Araneae, Araneoidea). Zoological Journal of the Linnean Society,

715   173:527-786.

716   Lu J, Yi L, Brusatte SL, Yang L, Li H, Chen L. 2014. A new clade of Asian Late Cretaceous long-snouted

717   tyrannosaurids. Nature Communications, 5.

718   Lue J, Unwin DM, Jin X, Liu Y, Ji Q. 2010. Evidence for modular evolution in a long-tailed pterosaur

719   with a pterodactyloid skull. Proceedings of the Royal Society B-Biological Sciences, 277:383-

720   389.

721   Mannion PD, Upchurch P. 2010. Completeness metrics and the quality of the sauropodomorph fossil

722   record through geological and historical time. Paleobiology, 36:283-302.

723   Mannion PD, Upchurch P, Barnes RN, Mateus O. 2013. Osteology of the Late Jurassic Portuguese

724   sauropod dinosaur Lusotitan atalaiensis (Macronaria) and the evolutionary history of basal

725   titanosauriforms. Zoological Journal of the Linnean Society, 168:98-206.

726   Maxwell WD, Benton MJ. 1990. Historical tests of the absolute completeness of the fossil record of

727   tetrapods. Paleobiology, 16:322-335.

728   McDonald AT. 2012. Phylogeny of Basal Iguanodonts (Dinosauria: Ornithischia): An Update. Plos

729   One, 7.

730   McDonald AT, Espilez E, Mampel L, Kirkland JI, Alcala L. 2012. An unusual new basal iguanodont

731   (Dinosauria: Ornithopoda) from the Lower Cretaceous of Teruel, Spain. Zootaxa:61-76.

732   McGarrity CT, Campione NE, Evans DC. 2013. Cranial anatomy and variation in Prosaurolophus

733   maximus (Dinosauria: Hadrosauridae). Zoological Journal of the Linnean Society, 167:531-568.

734   McPhee BW, Bonnan MF, Yates AM, Neveling J, Choiniere JN. 2015. A new basal sauropod from the

735   pre-Toarcian Jurassic of South Africa: evidence of niche-partitioning at the sauropodomorph-

736   sauropod boundary? Scientific Reports, 5.

737   Mickevich MF, Farris JS. 1981. The implications of congruence in Menidia. Systematic Zoology,

738   30:351-370.

739   Mitteroecker P, Bookstein F. 2007. The conceptual and statistical relationship between modularity

740   and morphological integration. Systematic Biology, 56:818-836.

741   Mounce RCP, Sansom R, Wills MA. 2016. Sampling diverse characters improves phylogenies:

742   Craniodental and postcranial characters of vertebrates often imply different trees. Evolution,

743   70:666-686.

744   Nicolalde-Morejon F, Vergara-Silva F, Gonzalez-Astorga J, Vovides AP, Espinosa de los Monteros A.

745   2009. Reciprocal illumination of morphological characters upon a molecular hypothesis

746   supports the proposal of a new species of cycad from Mexico. Systematics and Biodiversity,

747   7:73-79.

748   Norman DB. 2015. On the history, osteology, and systematic position of the Wealden (Hastings

749   group) dinosaur Hypselospinus fittoni (Iguanodontia: Styracosterna). Zoological Journal of the

750   Linnean Society, 173:92-189.

751   Norman DB, Crompton AW, Butler RJ, Porro LB, Charig AJ. 2011. The Lower Jurassic ornithischian

752   dinosaur Heterodontosaurus tucki Crompton & Charig, 1962: cranial anatomy, functional

753   morphology, taxonomy, and relationships. Zoological Journal of the Linnean Society, 163:182-

754   276.

755 Novas FE, Agnolín FL, Ezcurra MD, Porfiri J, Canale JI. 2013. Evolution of the carnivorous dinosaurs

756   during the Cretaceous: The evidence from Patagonia. Cretaceous Research, 45:174-215.

757 O'Connor A, Wills MA. 2016a. Measuring Stratigraphic Congruence Across Trees, Higher Taxa, and

758   Time. Systematic Biology, 65:792-811.

759 O'Connor A, Wills MA. 2016b. Measuring stratigraphic congruence across trees, higher taxa, and

760   time. Systematic Biology, 65:792-811.

761 O'Connor JK, Zhou Z. 2013. A redescription of Chaoyangia beishanensis (Aves) and a comprehensive

762   phylogeny of Mesozoic birds. Journal of Systematic Palaeontology, 11:889-906.

763 O'Leary MA, Gatesy J. 2008. Impact of increased character sampling on the phylogeny of

764   Cetartiodactyla (mammalia): combined analysis including fossils. Cladistics, 24:397-442.

765 Olmstead RG, Scotland RW. 2005. Molecular and morphological datasets. Taxon, 54:7-8.

766 Orlando L, Gilbert MTP, Willerslev E. 2015. Aplications of next-generation sequencing -

767   Reconstructing ancient genomes and epigenomes. Nature Reviews Genetics, 16:395-408.

768 Osi A, Prondvai E, Butler R, Weishampel DB. 2012. Phylogeny, Histology and Inferred Body Size

769   Evolution in a New Rhabdodontid Dinosaur from the Late Cretaceous of Hungary. Plos One, 7.

770 Oyston JW, Hughes M, Gerber S, Wills MA. 2016. Why should we investigate the morphological

771   disparity of plant clades? Annals of Botany, 117:859-879.

772 Oyston JW, Hughes M, Wagner PJ, Gerber S, Wills MA. 2015. What limits the morphological disparity

773   of clades? Interface Focus, 5.

774 Padial JM, Miralles A, De la Riva I, Vences M. 2010. The integrative future of taxonomy. Frontiers in

775   Zoology, 7.

776 Parker WG. 2016. Revised phylogenetic analysis of the Aetosauria (Archosauria: Pseudosuchia);

777   assessing the effects of incongruent morphological character sets. Peerj, 4.

778 Parsons WL, Parsons KM. 2015. Morphological Variations within the Ontogeny of Deinonychus

779   antirrhopus (Theropoda, Dromaeosauridae). Plos One, 10.

780   Pattinson DJ, Thompson RS, Piotrowski AK, Asher RJ. 2015. Phylogeny, paleontology, and primates:

781   do incomplete fossils bias the tree of life? Systematic Biology, 64:169-186.

782   Pettigrew JD. 1991. Wings or brain - convergent evoluton in the origins of bats. Systematic Zoology,

783   40:199-216.

784   Pol D, Garrido A, Cerda IA. 2011a. A New Sauropodomorph Dinosaur from the Early Jurassic of

785   Patagonia and the Origin and Evolution of the Sauropod-type Sacrum. Plos One, 6.

786   Pol D, Rauhut OWM, Becerra M. 2011b. A Middle Jurassic heterodontosaurid dinosaur from

787   Patagonia and the evolution of heterodontosaurids. Naturwissenschaften, 98:369-379.

788   Porfiri JD, Novas FE, Calvo JO, Agnolin FL, Ezcurra MD, Cerda IA. 2014. Juvenile specimen of

789   Megaraptor (Dinosauria, Theropoda) sheds light about tyrannosauroid radiation. Cretaceous

790   Research, 51:35-55.

791   Prieto-Marquez A, Dalla Vecchia FM, Gaete R, Galobart A. 2013. Diversity, Relationships, and

792   Biogeography of the Lambeosaurine Dinosaurs from the European Archipelago, with

793   Description of the New Aralosaurin Canardia garonnensis. Plos One, 8.

794   Prieto-Marquez A, Wagner JR. 2013. A new species of saurolophine hadrosaurid dinosaur from the

795   Late Cretaceous of the Pacific coast of North America. Acta Palaeontologica Polonica, 58:255-

796   268.

797   Ramirez MJ. 2006. Further problems with the incongruence length difference test:

798   "hypercongruence" effect and multiple comparisons. Cladistics, 22:289-295.

799   Rauhut OWM, Carballido JL, Pol D. 2015. A diplodocid sauropod dinosaur from the late Jurassic

800   Canadon Calcareo formation of Chubut, Argentina. Journal of Vertebrate Paleontology,

801   35:c982798-c982798.

802   Rayfield EJ. 2004. Cranial mechanics and feeding in Tyrannosaurus rex. Proceedings of the Royal

803   Society B-Biological Sciences, 271:1451-1459.

804   Robinson DF, Foulds LR. 1981. Comparison of phylogenetic trees. Mathematical Biosciences, 53:131-

805   147.

806   Rodrigo AG, Kellyborges M, Bergquist PR, Bergquist PL. 1993. A randomization test of the null

807   hypothesis that two cladograms are sample estimates of a parametric phylogenetic tree. New

808   Zealand Journal of Botany, 31:257-268.

809   Rubilar-Rogers D, Otero RA, Yury-Yanez RE, Vargas AO, Gutstein CS. 2012. An overview of the

810   dinosaur fossil record from Chile. Journal of South American Earth Sciences, 37:242-255.

811   Ruta M, Wills MA. 2016. Comparable disparity in the appendicular skeleton across the fish-tetrapod

812   transition, and the morphological gap between fish and tetrapod postcrania. Palaeontology,

813   59:249-267.

814   Saegusa H, Ikeda T. 2014. A new titanosauriform sauropod (Dinosauria: Saurischia) from the Lower

815   Cretaceous of Hyogo, Japan. Zootaxa, 3848:1-66.

816   Sanchez-Hernandez B, Benton MJ. 2014. Filling the ceratosaur gap: A new ceratosaurian theropod

817   from the Early Cretaceous of Spain. Acta Palaeontologica Polonica, 59:581-600.

818   Sanchez-Villagra MR, Williams BA. 1998. Levels of homoplasy in the evolution of the mammalian

819   skeleton. Journal of Mammalian Evolution, 5:113-126.

820   Sansom RS. 2015. Bias and sensitivity in the placement of fossil taxa resulting from interpretations of

821   missing data. Systematic Biology, 64:256-266.

822   Sansom RS, Gabbott SE, Purnell MA. 2010. Non-random decay of chordate characters causes bias in

823   fossil interpretation. Nature, 463:797-800.

824   Sansom RS, Wills MA. 2013. Fossilization causes organisms to appear erroneously primitive by

825   distorting evolutionary trees. Scientific Reports, 3.

826   Sansom RS, Wills MA. 2017. Differences between hard and soft phylogenetic data. Proceedings of

827   the Royal Society B-Biological Sciences, 284.

828   Sansom RS, Wills MA, Williams T. 2017. Dental Data Perform Relatively Poorly in Reconstructing

829   Mammal Phylogenies: Morphological Partitions Evaluated with Molecular Benchmarks.

830   Systematic Biology, 66:813-822.

831   Santucci RM, De Arruda-Campos AC. 2011. A new sauropod (Macronaria, Titanosauria) from the

832   Adamantina Formation, Bauru Group, Upper Cretaceous of Brazil and the phylogenetic

833   relationships of Aeolosaurini. Zootaxa:1-33.

834   Scotland RW, Olmstead RG, Bennett JR. 2003. Phylogeny reconstruction: The role of morphology.

835   Systematic Biology, 52:539-548.

836   Senter P, Kirkland JI, DeBlieux DD. 2012a. Martharaptor greenriverensis, a New Theropod Dinosaur

837   from the Lower Cretaceous of Utah. Plos One, 7.

838   Senter P, Kirkland JI, DeBlieux DD, Madsen S, Toth N. 2012b. New Dromaeosaurids (Dinosauria:

839   Theropoda) from the Lower Cretaceous of Utah, and the Evolution of the Dromaeosaurid Tail.

840   Plos One, 7.

841   Shapiro B, Hofreiter M. 2014. A Paleogenomic Perspective on Evolution and Gene Function: New

842   Insights from Ancient DNA. Science, 343.

843   Shibata M, Jintasakul P, Azuma Y, You H-L. 2015. A New Basal Hadrosauroid Dinosaur from the Lower

844   Cretaceous Khok Kruat Formation in Nakhon Ratchasima Province, Northeastern Thailand.

845   Plos One, 10.

846   Snively E, Henderson DM, Phillips DS. 2006. Fused and vaulted nasals of tyrannosaurid dinosaurs:

847   Implications for cranial strength and feeding mechanics. Acta Palaeontologica Polonica,

848   51:435-454.

849   Snively E, Russell AP. 2007a. Functional variation of neck muscles and their relation to feeding style

850   in tyrannosauridae and other large theropod dinosaurs. Anatomical Record-Advances in

851   Integrative Anatomy and Evolutionary Biology, 290:934-957.

852   Snively E, Russell AR. 2007b. Craniocervical feeding dynamics of Tyrannosaurus rex. Paleobiology,

853   33:610-638.

854   Song H, Bucheli SR. 2010. Comparison of phylogenetic signal between male genitalia and non-genital

855   characters in insect systematics. Cladistics, 26:23-35.

856　　Stanley SM, Roughgarden J, May RM, Levin SA. 1989. Fossils, macroevolution, and theoretical

857　　　　ecology.

858　　Swofford D. 2017. PAUP*: Phylogenetic Analysis Using Parsimony, Version 4.0a.154.

859　　Templeton AR. 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with

860　　　　particular reference to the evolution of humans and the apes. Evolution, 37:221-244.

861　　Thompson RS, Parish JC, Maidment SCR, Barrett PM. 2012. Phylogeny of the ankylosaurian dinosaurs

862　　　　(Ornithischia: Thyreophora). Journal of Systematic Palaeontology, 10:301-312.

863　　Tortosa T, Buffetaut E, Vialle N, Dutour Y, Turini E, Cheylan G. 2014. A new abelisaurid dinosaur from

864　　　　the Late Cretaceous of southern France: Palaeobiogeographical implications. Annales De

865　　　　Paleontologie, 100:63-86.

866　　Tschopp E, Mateus O. 2013. The skull and neck of a new flagellicaudatan sauropod from the

867　　　　Morrison Formation and its implication for the evolution and ontogeny of diplodocid

868　　　　dinosaurs. Journal of Systematic Palaeontology, 11:853-888.

869　　Tutin SL, Butler RJ. 2017. The completeness of the fossil record of plesiosaurs, marine reptiles from

870　　　　the Mesozoic. Acta Palaeontologica Polonica, 62:563-573.

871　　Verriere A, Brocklehurst N, Froebisch J. 2016. Assessing the completeness of the fossil record:

872　　　　comparison of different methods applied to parareptilian tetrapods (Vertebrata: Sauropsida).

873　　　　Paleobiology, 42:680-695.

874　　Wagner PJ. 1995. Testing evolutionary constraint hypotheses with early Palaeozoic gastropods.

875　　　　Paleobiology, 21:248-272.

876　　Wagner PJ. 1997. Patterns of morphologic diversification among the Rostroconchia. Paleobiology,

877　　　　23:115-150.

878　　Wahlberg N, Braby MF, Brower AVZ, de Jong R, Lee MM, Nylin S, Pierce NE, Sperling FAH, Vila R,

879　　　　Warren AD, *et al.* 2005. Synergistic effects of combining morphological and molecular data in

880　　　　resolving the phylogeny of butterflies and skippers. Proceedings of the Royal Society B-

881　　　　Biological Sciences, 272:1577-1586.

882   Wang M, Li D, O'Connor JK, Zhou Z, You H. 2015a. Second species of enantiornithine bird from the

883   Lower Cretaceous Changma Basin, northwestern China with implications for the taxonomic

884   diversity of the Changma avifauna. Cretaceous Research, 55:56-65.

885   Wang M, Zheng X, O'Connor JK, Lloyd GT, Wang X, Wang Y, Zhang X, Zhou Z. 2015b. The oldest

886   record of ornithuromorpha from the early cretaceous of China. Nature Communications, 6.

887   Wiens JJ. 1998. Does adding characters with missing data increase or decrease phylogenetic

888   accuracy? Systematic Biology, 47:625-640.

889   Wiens JJ. 2004. The role of morphological data in phylogeny reconstruction. Systematic Biology,

890   53:653-661.

891   Wilkinson M. 1995. Coping with abundant missing entries in phylogenetic inference using parsimony.

892   Systematic Biology, 44:501-514.

893   Williams BA. 2007. Comparing levels of homoplasy in the primate skeleton. Journal of Human

894   Evolution, 52:480-489.

895   Wills MA. 1998. A phylogeny of recent and fossil Crustacea derived from morphological characters.

896   Systematics Association Special Volume Series, 55:189-209.

897   Wills MA, Barrett PM, Heathcote JF. 2008a. The modified gap excess ratio (GER*) and the

898   stratigraphic congruence of dinosaur phylogenies. Systematic Biology, 57:891-904.

899   Wills MA, Barrett PM, Heathcote JF. 2008b. The Modified Gap Excess Ratio (GER*) and the

900   Stratigraphic Congruence of Dinosaur Phylogenies. Systematic Biology, 57:891-904.

901   Wills MA, Jenner RA, Dhubhghaill CN. 2009. Eumalacostracan Evolution: Conflict between Three

902   Sources of Data. Arthropod Systematics & Phylogeny, 67:71-90.

903   Wilson JA, Allain R. 2015. Osteology of Rebbachisaurus garasbae Lavocat, 1954, a diplodocoid

904   (Dinosauria, Sauropoda) from the early Late Cretaceous-aged Kem Kem beds of southeastern

905   Morocco. Journal of Vertebrate Paleontology, 35:e1000701-e1000701.

906   Wortley AH, Scotland RW. 2006. Determining the potential utility of datasets for phylogeny

907   reconstruction. Taxon, 55:431-442.

908   Xing H, Wang D, Han F, Sullivan C, Ma Q, He Y, Hone DWE, Yan R, Du F, Xu X. 2014. A New Basal

909   Hadrosauroid Dinosaur (Dinosauria: Ornithopoda) with Transitional Features from the Late

910   Cretaceous of Henan Province, China. Plos One, 9.

911   Xing L, Miyashita T, Currie PJ, You H, Zhang J, Dong Z. 2015. A new basal eusauropod from the

912   Middle Jurassic of Yunnan, China, and faunal compositions and transitions of Asian

913   sauropodomorph dinosaurs. Acta Palaeontologica Polonica, 60:145-154.

914   Zanno LE, Makovicky PJ. 2013. Neovenatorid theropods are apex predators in the Late Cretaceous of

915   North America. Nature Communications, 4.

916   Zhou S, Zhou Z, O'Connor J. 2014. A new piscivorous ornithuromorph from the Jehol Biota. Historical

917   Biology, 26:608-618.

918   Zou Z, Zhang J. 2016. Morphological and molecular convergences in mammalian phylogenetics.

919   Nature Communications, 7.

920

921

922 **Table 1. Summary of all tests and statistics for all 81 partitioned data sets**

Reference	Craniodental Number	Postcranial Number	Taxa	Taxa removed	ILD test	IRD RF NN	IRD RF M	IRD Match NND	IRD Match MR	p-value	Craniodental missing %	Postcranial missing %	Craniodental CI	Craniodental RI	Postcranial CI	Postcranial RI
Allain_etal_2012a-Bc	74	94	20	9	0.140	0.11	0.04	0.07	0.04	4.19E-09	16.22	19.47	0.5172	0.6667	0.585	0.7615
Araujo_etal_2013a-Bc	159	187	22	42	0.011	0.01	0.01	0.03	0.03	3.12E-02	30.16	26.15	0.5064	0.6025	0.5899	0.729
Brusatte_et_Benson_2013a-Bc	200	107	18	16	0.164	0.01	0.01	0.02	0.05	2.20E-16	23.67	34.53	0.6809	0.8353	0.6648	0.846
Brusatte_etal_2014a-Bc	368	473	56	95	0.054	0.01	0.01	0.01	0.01	0.1693	49.53	39.33	0.4584	0.7616	0.4219	0.5781
Canale_etal_2015a-Bc	100	55	16	1	0.610	0.01	0.21	0.05	0.07	2.20E-16	35.8	39.27	0.6667	0.7617	0.6239	0.6583
Cau_etal_2012a-Bc	338	443	29	8	0.031	0.03	0.03	0.02	0.01	3.18E-08	38.32	45.39	0.3934	0.4944	0.4329	0.5189
Choiniere_etal_2014ba-Bc	216	290	58	56	0.001	0.01	0.01	0.01	0.01	0.0001219	48.83	46.83	0.2892	0.5817	0.3029	0.6013
Evers_etal_2015a-Bc	152	192	42	22	0.106	0.01	0.01	0.01	0.01	0.5229	46.2	43.42	0.434	0.6721	0.4636	0.7163
Eddy_et_Clarke_2011a-Bc	95	66	19	3	0.032	0.18	0.23	0.11	0.11	0.002423	26.93	35.35	0.6294	0.6702	0.7048	0.7304
Fanti_etal_2012a-Bc	109	69	17	2	0.001	0.01	0.02	0.03	0.03	2.20E-16	31.41	37	0.75	0.8442	0.5548	0.5988
Farke_et_Sertich_2013a-Bc	81	106	18	12	0.998	0.02	0.01	0.01	0.01	0.01551	46.52	43.64	0.7328	0.8558	0.5867	0.6773
Foth_etal_2014a-Bc	177	357	53	89	0.001	0.01	0.01	0.01	0.08	0.2493	39.2	35.9	0.4114	0.7684	0.392	0.7412
Godefroit_etal_2013b-Bc	373	614	27	28	0.001	0.01	0.01	0.01	0.01	8.94E-01	41.22	47.8	0.4647	0.5216	0.5207	0.5776
Hu_etal_2015a-Bc	41	171	27	35	0.231	0.63	0.76	0.02	0.02	5.02E-02	57.18	33.1	0.6761	0.783	0.4815	0.6337
Lamanna_etal_2014a-Bc	130	86	22	19	0.001	0.04	0.01	0.03	0.02	3.36E-16	34.44	45.4	0.6441	0.7711	0.5312	0.5909
Lee_etal_2014b-Bc	42	181	20	45	0.165	0.35	0.68	0.38	0.75	3.58E-11	37.5	33.5	0.6901	0.8211	0.6144	0.7315
LiZ_etal_2014a-Bc	45	162	16	20	0.372	0.84	0.85	0.67	0.7	0.003983	30.42	16.51	0.5446	0.6085	0.4641	0.5339
Loewen_etal_2013a-Bc	297	196	40	0	0.122	0.08	0.06	0.05	0.12	5.72E-07	29.69	29.96	0.4093	0.7603	0.4239	0.7713
Longrich_etal_2011a-Bc	46	163	26	20	0.002	0.01	0.01	0.01	0.03	2.58E-02	42.5	21.53	0.8154	0.9084	0.6353	0.7913
Lu_etal_2014b-Bc	203	107	20	5	0.233	0.01	0.01	0.01	0.01	2.20E-16	25.15	38.78	0.663	0.8345	0.6536	0.8438
Novas_etal_2013aa-Bc	91	188	21	24	0.475	0.03	0.05	0.01	0.09	0.03884	21.4	30.37	0.4823	0.6528	0.514	0.6661
Parsons_et_Parsons_2015a-Bc	254	188	28	90	0.016	0.01	0.01	0.03	0.02	3.49E-02	24.41	23.74	0.5523	0.7823	0.6348	0.8331
Porfiri_etal_2014a-Bc	122	159	31	14	0.374	0.04	0.04	0.38	0.31	0.1477	37.49	39.49	0.3989	0.6326	0.4371	0.6698
Sanchez-Hernandez_et_Benton_2014a-Bc	68	78	17	4	1.000	0.01	0.05	0.01	0.01	3.28E-08	45.16	44.34	0.7553	0.8655	0.6296	0.6875
Senter_etal_2012b-Bc	153	229	79	31	0.021	0.04	0.06	0.01	0.01	2.20E-16	20.73	17.83	0.6583	0.8601	0.6196	0.8179
Tortosa_etal_2014a-Bc	118	177	16	24	0.017	0.01	0.01	0.01	0.01	6.80E-11	44.6	34.03	0.7822	0.8327	0.6022	0.6613
Wang_etal_2015b-Bc	40	185	17	41	0.018	0.18	0.51	0.17	0.13	3.83E-04	43.38	30.84	0.7121	0.7841	0.6546	0.7103
Zanno_et_Makovicky_2013a-Bc	148	190	36	26	0.080	0.05	0.23	0.05	0.04	3.44E-15	42.59	33.81	0.4474	0.6555	0.4725	0.7113
Zhou_etal_2014a-Bc	46	94	15	18	0.001	0.02	0.02	0.03	0.12	5.59E-02	26.96	15.84	0.7746	0.8571	0.7462	0.8107
DEmic_2013a-Ba	15	96	13	12	0.018	0.05	0.06	0.03	0.08	6.40E-05	23.6	18.2	0.7273	0.8065	0.7769	0.8585
Gorscak_etal_2014a-Ba	100	145	22	13	0.074	0.02	0.02	0.02	0.01	6.10E-09	40.09	31.57	0.6687	0.7611	0.6681	0.7557
Santucci_et_Arruda-Campos_2011a-Ba	76	152	27	11	0.010	0.01	0.01	0.01	0.03	7.63E-15	32.2	21.36	0.7391	0.7778	0.7054	0.7692
Tschopp_et_Mateus_2013a-Ba	66	96	14	0	0.131	0.12	0.29	0.04	0.32	1.06E-07	47.29	30.73	0.6964	0.6634	0.5824	0.6181
Wilson_et_Allain_2015a-Ba	66	80	13	0	0.263	0.01	0.01	0.01	0.01	2.20E-16	31.47	15.38	0.8675	0.9106	0.6641	0.7676
Carballido_&_Sander-Bb	99	223	22	52	0.002	0.10	0.09	0.17	0.16	8.80E-13	33.47	29.03	0.6628	0.7852	0.5704	0.6525
Carballido_etal_2015a-Bb	95	231	18	55	0.041	0.05	0.03	0.04	0.05	6.60E-02	29.07	20.06	0.6467	0.7204	0.5859	0.5983
Fanti_etal_2015a-Bb	72	126	20	0	0.466	0.01	0.01	0.16	0.11	2.88E-10	41.32	27.18	0.7105	0.8187	0.5153	0.6333
Lacovara_etal_2014a-Bb	100	226	22	53	0.009	0.03	0.03	0.01	0.01	0.2951	34.4	27.82	0.6667	0.7765	0.5744	0.6679
LiL_etal_2014-c	99	220	21	50	0.020	0.34	0.24	0.26	0.28	2.05E-08	32.13	26.02	0.6686	0.7778	0.5752	0.6555
Mannion_etal_2013a-Bb	53	222	18	33	0.129	0.06	0.20	0.03	0.03	1.32E-01	39.41	22.87	0.5758	0.6693	0.4018	0.5108
McPhee_etal_2015a-Bb	117	233	32	22	0.496	0.04	0.20	0.16	0.15	2.20E-16	30.9	20.53	0.4227	0.6924	0.3894	0.6524
Rauhut_etal_2015a-Bb	102	216	29	43	0.677	0.01	0.01	0.01	0.01	2.20E-16	47.63	26.48	0.6031	0.7475	0.476	0.7051
Rubilar-Rogers_etal_2012a-Bb	68	150	22	8	0.080	0.01	0.01	0.01	0.01	2.20E-16	46.35	24.94	0.7064	0.7867	0.6614	0.8009
Xing_etal_2015b-Bb	101	237	28	17	0.096	0.01	0.01	0.01	0.01	2.20E-16	26.22	22.6	0.3269	0.5566	0.7135	0.6628
Pol_etal_2011a-Bb	102	160	38	12	0.120	0.52	0.58	0.37	0.29	7.14E-03	37.59	36.12	0.625	0.7931	0.6667	0.7978
Saegusa_et_Ikeda_2014a-Bb	15	99	13	16	0.025	0.06	0.04	0.01	0.03	0.001353	23.6	18.41	0.7273	0.8065	0.7803	0.8585
Evans-Aa	51	48	24	1	0.078	0.04	0.03	0.01	0.01	2.20E-16	29.17	43.66	0.6471	0.8266	0.7639	0.844
Farke_etal_2011-A	48	22	16	2	0.796	0.10	0.71	0.04	0.47	0.07588	16.5	31.1	0.7391	0.7857	0.8519	0.8519
Farke_etal_2014ba-Aa	103	43	14	19	0.082	0.04	0.22	0.03	0.01	5.04E-08	20.94	20.59	0.75	0.8413	0.746	0.8367
Han_etal_2015a-Aa	162	47	18	19	0.207	0.02	0.62	0.22	0.93	5.45E-03	17.18	20.8	0.5938	0.7784	0.68	0.831
Longrich_2011aa-Aa	90	23	19	3	1.000	0.47	0.45	0.43	0.41	4.79E-07	17	27	0.7983	0.8983	1	1
Longrich_2014a-Aa	159	32	16	19	0.727	0.05	0.09	0.02	0.31	8.64E-04	17.33	33.98	0.7449	0.8098	1	1
Boyd_2015a-Ab	128	99	34	65	0.814	0.85	0.74	0.59	0.6	4.24E-05	40.72	31.7	0.37	0.6056	0.4206	0.6477
Boyd_et_Pagnac_2015a-Ab	90	41	28	40	0.028	0.07	0.01	0.02	0.01	2.53E-12	25	27.44	0.5714	0.7894	0.5481	0.7991
Brown_etal_2011a-Ab	76	53	20	3	0.181	0.39	0.46	0.21	0.38	0.7277	27.43	23.11	0.5189	0.6877	0.5034	0.654
Evans_etal_2013a-Ab	28	10	12	6	0.993	0.01	0.07	0.07	0.37	0.2078	15.18	46.7	0.6792	0.7302	0.9167	0.9091
Godefroit_etal_2012b-Ab	100	46	19	2	0.108	0.02	0.01	0.02	0.02	2.20E-16	18.84	30.32	0.6875	0.7766	0.6782	0.7667
He_etal_2015a-Ab	191	144	34	27	0.007	0.68	0.49	0.83	0.71	0.1416	22.16	24.04	0.4275	0.6648	0.4595	0.6855
McDonald_etal_2012a-Ab	91	41	42	25	0.285	0.04	0.03	0.13	0.15	1.60E-09	21.89	24	0.5955	0.7977	0.596	0.8113
McGarrity_etal_2013a-Ab	102	46	20	2	0.202	0.01	0.01	0.01	0.01	2.20E-16	19.12	33.8	0.6759	0.7727	0.6782	0.7667
Norman_2015a-Ab	68	37	23	4	0.772	0.97	0.98	0.97	0.96	0.6486	10.29	17.63	0.594	0.781	0.6542	0.8405
Norman_etal_2011a-Ab	66	57	13	1	0.090	0.12	0.23	0.11	0.15	0.03156	12.82	18.89	0.5659	0.5591	0.6311	0.6481
Prieto-Marquez_2014b-Ab	155	67	19	15	0.088	0.34	0.83	0.12	0.37	2.20E-01	21.59	16.1	0.7233	0.7751	0.6013	0.7067
Prieto-Marquez_et_Wagner_2013a-Ab	194	91	34	17	0.303	0.69	0.68	0.39	0.32	0.02466	27	20.49	0.5737	0.8063	0.4843	0.7587
Prieto-Marquez_etal_2013a-Ab	143	66	26	8	0.475	0.01	0.03	0.14	0.43	0.0001373	27.65	29.37	0.5971	0.78	0.5577	0.6849
Shibata_etal_2015a-Ab	68	37	23	28	0.760	0.95	1.00	0.96	1	0.0393	10.29	17.62	0.5914	0.781	0.6542	0.8405
Xing_etal_2014a+Ab	233	111	32	30	0.645	0.16	0.60	0.06	0.21	9.42E-08	14.06	11.09	0.6359	0.8806	0.5991	0.8791
Arbour_et_Currie_2013c-Ac	70	56	17	1	0.436	0.11	0.73	0.58	0.83	0.04001	20.84	29.05	0.5692	0.72	0.7093	0.7642
Arbour_etal_2014b-Ac	57	66	19	1	0.460	0.14	0.31	0.11	0.08	7.00E-07	22.3	44.82	0.5455	0.7207	0.6729	0.7569
Barrett_etal_2014a-Ac	130	94	36	19	0.240	0.06	0.05	0.09	0.25	0.02561	22.54	29.28	0.4531	0.6841	0.5194	0.7324
Burns_et_Currie_2014-A	45	33	23	0	0.761	0.01	0.01	0.05	0.09	4.97E-05	16.14	45.85	0.4522	0.6957	0.6479	0.7312
Burns_et_Currie_2014-B	26	23	16	0	0.002	0.21	0.08	0.09	0.06	2.20E-16	13.94	32.07	0.6522	0.8261	0.6818	0.7407
Burns_etal_2011a-Ac	45	16	23	0	0.426	0.29	0.75	0.24	0.61	0.05312	12.8	36.76	0.5484	0.7103	0.7143	0.7838
Butler_etal_2011b-Ac	124	93	31	20	0.234	0.03	0.04	0.51	0.69	0.000465	22.11	23.93	0.4964	0.6745	0.5381	0.7217
Coria_etal_2013a-Ac	130	97	28	24	0.022	0.03	0.15	0.17	0.22	3.14E-02	19.86	20.95	0.5199	0.6787	0.5567	0.7463
Godefroit_etal_2014a-Ac	130	94	31	24	0.041	0.02	0.02	0.38	0.48	0.01808	19.35	21.17	0.4814	0.6585	0.512	0.6973
Han_etal_2012a-Ac	127	94	38	24	0.137	0.07	0.04	0.24	0.12	1.79E-01	18.06	18.56	0.5112	0.6881	0.5323	0.7134
Osi_etal_2012b-Ac	132	97	40	18	0.313	0.03	0.03	0.01	0.02	4.78E-02	18.56	18	0.4354	0.6708	0.5094	0.707
Pol_etal_2011b-Ac	129	97	42	9	0.020	0.01	0.03	0.01	0.11	0.01531	19.92	26.08	0.5238	0.6692	0.5567	0.7362
Ruiz-Omenaca_etal_2012a-Ac	130	94	35	18	0.641	0.04	0.01	0.23	0.05	6.85E-13	20.68	30.67	0.4545	0.6872	0.5347	0.7307
Thompson_etal_2012a-Ac	88	75	39	12	0.294	0.09	0.13	0.02	0.11	6.95E-09	24.45	35.8	0.445	0.729	0.4536	0.6753

923

924 IRD tests are based either upon the mean tree-to-tree distances between nearest neighbours (NND)
925 or the distance between majority rule (plus compatible grouping) trees (MR). The tree-to-tree
926 distance metric used is either the Robinson Foulds (symmetrical difference) distance (RF) or the
927 Matching distance (Match). Sources: (Brown et al., 2011; Burns et al., 2011; Butler et al., 2011;
928 Carballido et al., 2011; Eddy and Clarke, 2011; Farke et al., 2011; Longrich, 2011; Pol et al., 2011a;
929 Pol et al., 2011b; Santucci and De Arruda-Campos, 2011; Allain et al., 2012; Fanti et al., 2012; Han et
930 al., 2012; Ignacio Ruiz-Omenaca et al., 2012; McDonald, 2012; Osi et al., 2012; Rubilar-Rogers et al.,
931 2012; Senter et al., 2012a; Senter et al., 2012b; Thompson et al., 2012; Araujo et al., 2013; Arbour
932 and Currie, 2013; Brusatte and Benson, 2013; Cau et al., 2013; Coria et al., 2013; D'Emic, 2013; Farke
933 and Sertich, 2013; Godefroit et al., 2013; Loewen et al., 2013; Mannion et al., 2013; McGarrity et al.,
934 2013; Novas et al., 2013; Prieto-Marquez et al., 2013; Prieto-Marquez and Wagner, 2013; Tschopp
935 and Mateus, 2013; Zanno and Makovicky, 2013; Arbour et al., 2014; Barrett et al., 2014; Brusatte et
936 al., 2014; Burns and Currie, 2014; Carballido and Sander, 2014; Choiniere et al., 2014; Evans et al.,
937 2014; Farke et al., 2014; Foth et al., 2014; Gorscak et al., 2014; Lacovara, 2014; Lamanna et al., 2014;
938 Lee et al., 2014; Li et al., 2014; Longrich, 2014; Lu et al., 2014; Porfiri et al., 2014; Saegusa and Ikeda,
939 2014; Sanchez-Hernandez and Benton, 2014; Tortosa et al., 2014; Xing et al., 2014; Zhou et al., 2014;
940 Boyd, 2015; Boyd and Pagnac, 2015; Carballido et al., 2015; Evans and Ryan, 2015; Evers et al., 2015;

941   Fanti et al., 2015; Hu et al., 2015; Ignacio Canale et al., 2015; McPhee et al., 2015; Norman, 2015;

942   Parsons and Parsons, 2015; Rauhut et al., 2015; Shibata et al., 2015; Wang et al., 2015a; Wang et al.,

943   2015b; Wilson and Allain, 2015; Xing et al., 2015; He et al., 2016)

944

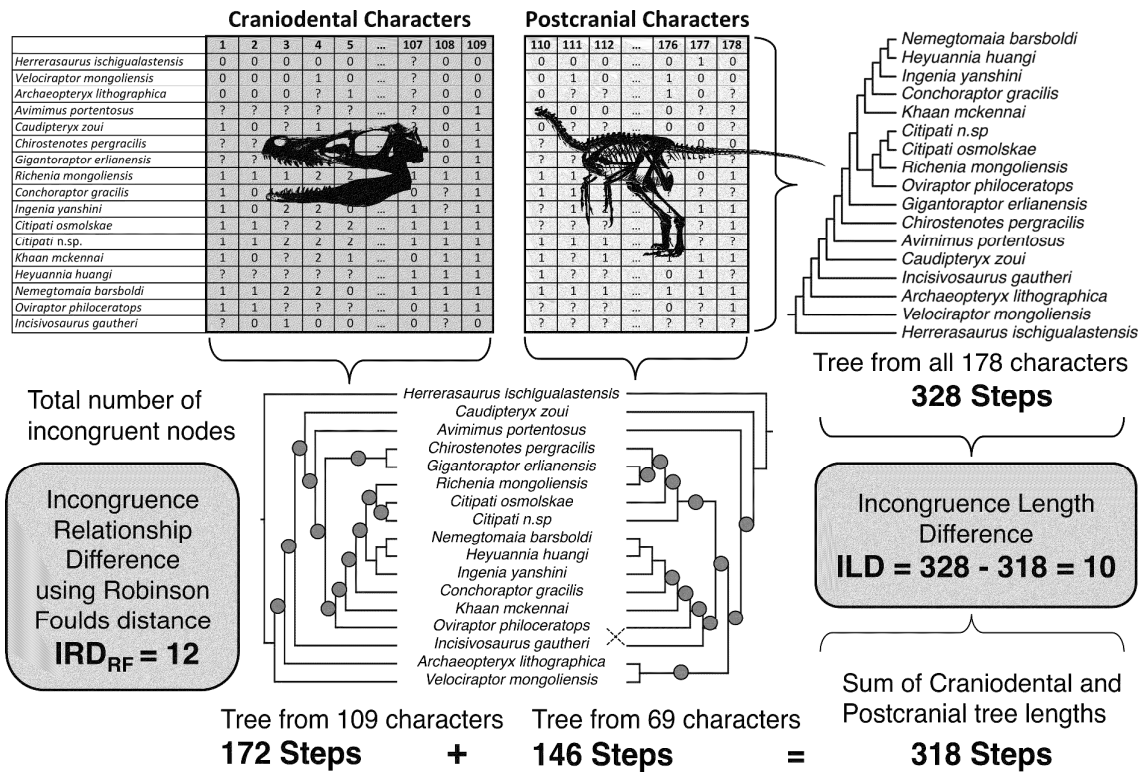945   **Table 2. Summary of matrix partitions and results of tests**

	All matrices	Theropoda	Sauropoda	Cerapoda	Ornithopoda	Thyreophora	W or G	p
No. Matrices	81	29	17	6	15	14	-	-
Median Number Craniodental/Postcranial Characters	100/97	122/177	95/160	96.5/37.5	100/53	125.5/93.5	-	-
Median Craniodental/Postcranial CI	0.5971/0.5867	0.6294/0.585	0.6686/0.5824	0.742/0.8079	0.5914/0.5991	0.5038/0.5474	-	-
Median Craniodental/Postcranial RI	0.7711/0.7307	0.7684/0.7103	0.7778/0.6679	0.8182/0.84795	0.7766/0.7667	0.68765/0.7318	-	-
*Craniodental/Postcranial most similar to entire (NND+Matching)	39/42	13/16	4/13	5/1	10/5	7/7		
f p<0.05 ILD	27	14	7	0	2	4	11.3810	0.0226
f p<0.05 IRD NND+RF	45	21	9	3	5	7	6.7962	0.1471
f p<0.05 IRD MR+RF	39	17	9	1	5	7	5.4241	0.2465
f p<0.05 IRD NND+Matching	40	19	12	4	3	2	20.1841	**0.0005**
f p<0.05 IRD MR+Matching	35	19	9	3	3	1	19.3857	**0.0007**

946

947   *Craniodental/Postcranial most similar to entire (NND+Matching). Both partitions and the entire

948   matrix were repeatedly (x100) resampled at the sample size (number of characters) of the smaller

949   partition, and most parsimonious trees were inferred from all three samples of characters. For each

950   resampling, the mean matching distance between nearest neighbours was used to determine which

951   partitioned tree(s) (craniodental or postcranial) were most similar to the tree(s) from the

952   subsampled entire matrix. The mean of these distances across all random re-samplings was then

953   used to determine which partition (craniodental or postcranial) yielded trees mot similar to that

954   from the entire matrix overall. The last five rows of the table indicate the frequency with which

955   partitions yield trees that are significantly different (with p < 0.05) for the ILD and variants of the IRD

956   test. IRD tests are based either upon the mean tree-to-tree distances between nearest neighbours

957   (NND) or the distance between majority rule (plus compatible grouping) trees (MR). The tree-to-tree

958   distance metric used is either the Robinson Foulds (symmetrical difference) distance (RF) or the

959   Matching distance.

960

961

Craniodental Characters

	1	2	3	4	5	...	107	108	109
Herrerasaurus ischigualastensis	0	0	0	0	0	...	?	0	0
Velociraptor mongoliensis	0	0	0	1	0	...	?	0	0
Archaeopteryx lithographica	0	0	0	?	1	...	?	0	0
Avimimus portentosus	?	?	?	?	?	...	?	0	1
Caudipteryx zoui	1	0	?	1		...		0	1
Chirostenotes pergracilis	?	?				...		0	1
Gigantoraptor erlianensis	?	?				...		0	1
Richenia mongoliensis	1	1	1	2	2	...	1	1	1
Conchoraptor gracilis	1	0				...	0	?	1
Ingenia yanshini	1	0	2	2	0	...	1	?	1
Citipati osmolskae	1	1	?	2	2	...	1	1	1
Citipati n.sp.	1	1	2	2	2	...	1	1	1
Khan mckennai	1	0	?	2	1	...	0	1	1
Heyuannia huangi	?	?	?	?	?	...	1	1	1
Nemegtomaia barsboldi	1	1	2	2	0	...	1	1	1
Oviraptor philoceratops	1	1	?	?	?	...	0	1	1
Incisivosaurus gautheri	?	0	1	0	0	...	0	?	0

Postcranial Characters

110	111	112	...	176	177	178
0	0	0	...	0	1	0
0	1	0	...	1	0	0
0	?	?	...	1	0	?
0	0	0	...	0	0	?
0	?	?	...	0	0	?
?			...	0	0	0
?	?	?	...	0	0	?
1	1		...	0	1	1
1	1		...	?	?	?
?	1		...	1	1	1
?	?	?	...	1	1	?
1	1	1	...	1	1	1
1	?	?	...	1	1	1
1	?	?	...	0	1	?
1	1	1	...	1	1	1
?	?	?	...	0	1	1
?	?	?	...	?	?	?

Nemegtomaia barsboldi
Heyuannia huangi
Ingenia yanshini
Conchoraptor gracilis
Khaan mckennai
Citipati n.sp
Citipati osmolskae
Richenia mongoliensis
Oviraptor philoceratops
Gigantoraptor erlianensis
Chirostenotes pergracilis
Avimimus portentosus
Caudipteryx zoui
Incisivosaurus gautheri
Archaeopteryx lithographica
Velociraptor mongoliensis
Herrerasaurus ischigualastensis

Tree from all 178 characters
328 Steps

Total number of incongruent nodes

Incongruence Relationship Difference using Robinson Foulds distance
IRD$_{RF}$ = 12

Herrerasaurus Ischigualastensis
Caudipteryx zoui
Avimimus portentosus
Chirostenotes pergracilis
Gigantoraptor erlianonsis
Richenia mongoliensis
Citipati osmolskae
Citipati n.sp
Nemegtomaia barsboldi
Heyuannia huangi
Ingenia yanshini
Conchoraptor gracilis
Khaan mckennai
Oviraptor philoceratops
Incisivosaurus gautheri
Archaeopteryx lithographica
Velociraptor mongoliensis

Incongruence Length Difference
ILD = 328 - 318 = 10

Sum of Craniodental and Postcranial tree lengths
318 Steps

Tree from 109 characters   Tree from 69 characters
172 Steps   +   146 Steps   =

Figure 1. Calculation of the Incongruence Length Difference (ILD) and the Incongruence Relationship Difference (IRD) for the cranial and postcranial character partitions in the data of Fanti et al., 2012. Parsimony analysis of all 178 characters together yields a single most parsimonious tree (MPT) of 328 steps. Analysis of 109 craniodental characters alone yields an MPT length of 172 steps, while 69 postcranial characters alone yield an MPT length of 146 steps. The sum of these partitioned lengths (172+146 = 318) is less than the length of the global MPT by 10 steps (328-318 = 10). This discrepancy is the incongruence length difference (ILD = 10). The partitioned trees imply different relationships, and the magnitude of this difference can be measured using a diversity of tree-to-tree distance metrics. The Robinson Foulds (RF) or symmetrical difference distance (Robinson and Foulds, 1981) is among the most widely applied, and is calculated as the sum of the number of internal tree nodes that are present in one tree but not the other. The IRD$_{RF}$ is therefore the incongruence relationship difference measured using the RF distance. The RF distance has the disadvantage that it can saturate quickly. Numerous other tree-to-tree distance metrics are available, and we also implement the matching distance here (MD) (Lin et al, 2012), to yield the IRD$_{Matching}$. The ILD, IRD$_{RF}$ and IRD$_{Matching}$ each enables a test of incongruence, implemented by randomly partitioning the data set into character sets of the same size as the original (here, 109 and 69) and recalculating the metric. This is repeated a large number of times to yield a null distribution for randomised metric values, and the value for the original partition is compared with this in order to yield an empirical p value. Because the ILD, IRD$_{RF}$ and IRD$_{MD}$ all measure different things (tree length versus different aspects of tree shape and relationships), the results of these tests do not invariably coincide.
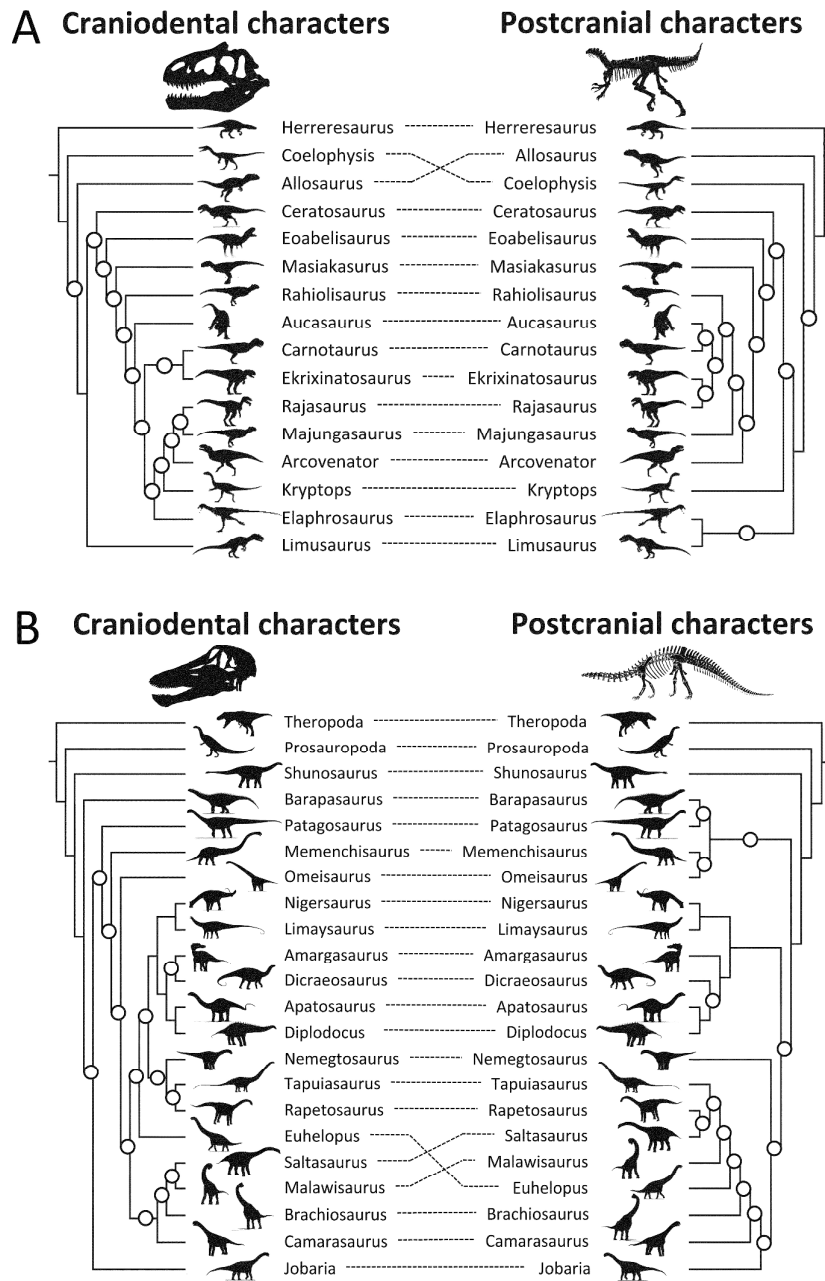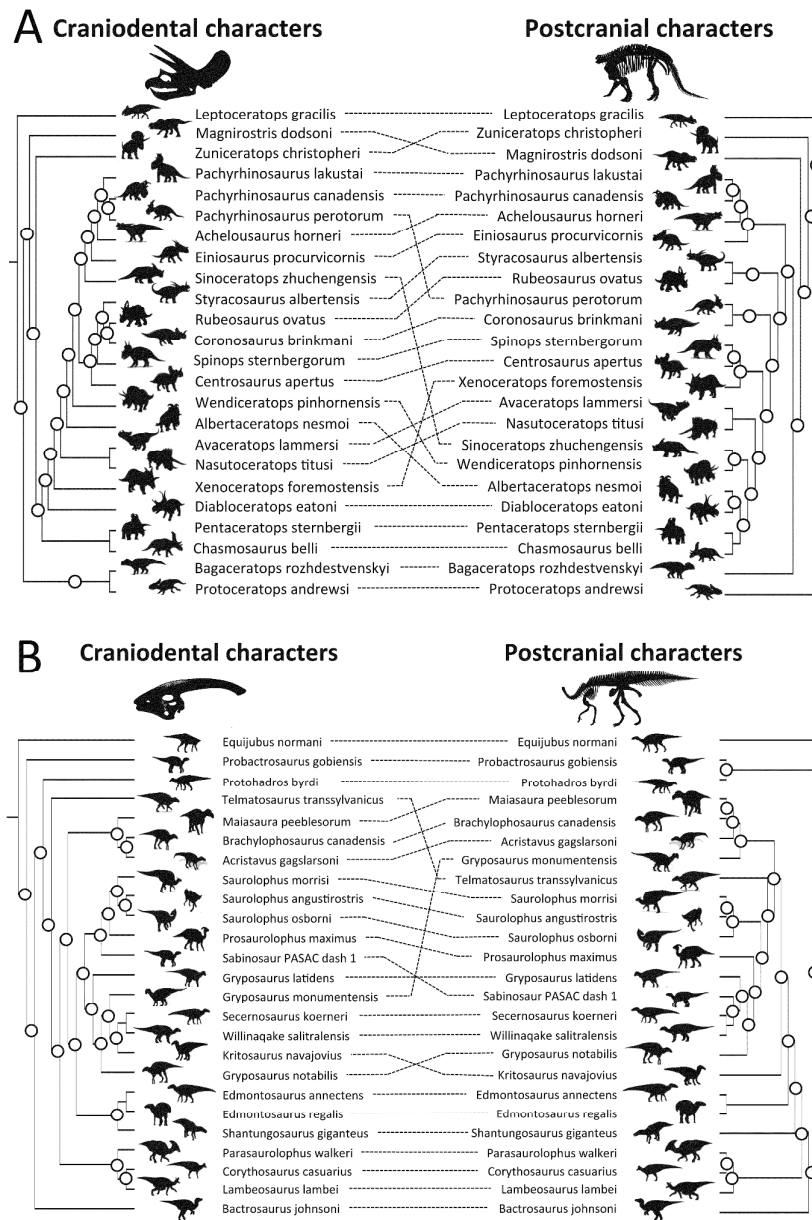
Figure 2. Exemplary tanglegrams for two groups of Saurischia. All trees are majority rule trees, plus compatible groupings. We do not necessarily recommend the use of majority rule trees in calculating IRD statistics (although we summarise these IRD_{MR} calculations in Tables 1 and 2), because they can be far from the centroid of tree space (Mounce et al., 2016). Rather, we advocate the use of mean distances between nearest neighbours in the two sets of trees for comparison. The left hand tree in each panel is derived from craniodental characters, while the right hand tree is derived from postcranial characters. Circled internal nodes are those present in one tree but not the other, and are tallied to give the Robinson Foulds (RF) or symmetrical difference distance (Robinson and Foulds, 1981). **A.** Tanglegram for Theropoda using data from Tortosa et al., (2014). (ILD p=0.017; IRD_{NND+RF} p=0.01; $IRD_{NND+Matching}$ p=0.01; IRD_{MR+RF} p=0.01; $IRD_{MR+Matching}$ p=0.01). **B.** Tanglegram for Sauropoda using data from Gorscak et al., 2014. (ILD p=0.074; IRD_{NND+RF} p=0.02; $IRD_{NND+Matching}$ p=0.02; IRD_{MR+RF} p=0.02; $IRD_{MR+Matching}$ p=0.01).

999

Figure 3. Exemplary tanglegrams for two groups of Ornithischia. All trees are majority rule trees, plus compatible groupings. We do not necessarily recommend the use of majority rule trees in calculating IRD statistics (although we summarise these IRD$_{MR}$ calculations in Tables 1 and 2), because they can be far from the centroid of tree space (Mounce et al., 2016). Rather, we advocate the use of mean distances between nearest neighbours in the two sets of trees for comparison. The left hand tree in each panel is derived from craniodental characters, while the right hand tree is derived from postcranial characters. Circled internal nodes are those present in one tree but not the other, and are tallied to give the Robinson Foulds (RF) or symmetrical difference distance (Robinson and Foulds, 1981). **A.** Tanglegram for Ceratopsia using data from Evans and Ryan,(2015). (ILD p=0.078; IRD$_{NND+RF}$ p=0.04; IRD$_{NND+Matching}$ p=0.01; IRD$_{MR+RF}$ p=0.03; IRD$_{MR+Matching}$ p=0.01). **B.** Tanglegram for Hadrosaurida using data from Prieto and Marquez, 2014. (ILD p=0.088; IRD$_{NND+RF}$ p=0.34; IRD$_{NND+Matching}$ p=0.12; IRD$_{MR+RF}$ p=0.83; IRD$_{MR+Matching}$ p=0.37).

1013
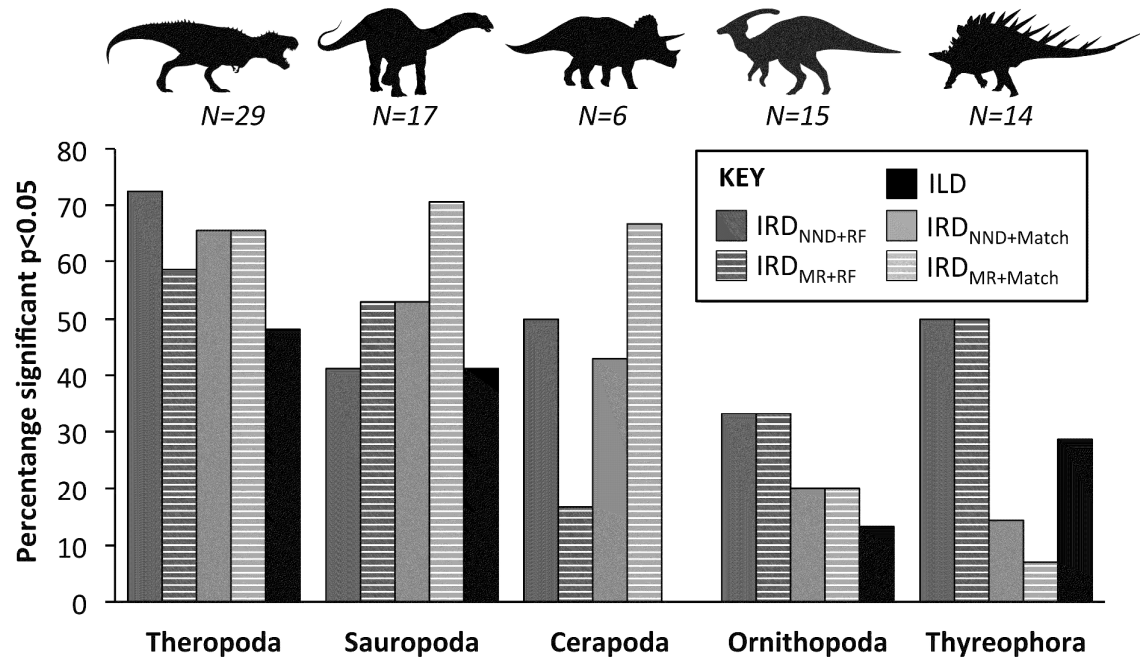
Figure 4. Summary of results of Incongruence Length Difference (ILD) and variants of the Incongruence Relationship Difference (IRD) tests, partitioned by major taxonomic group. Bars denote the percentage of data sets for which $p<0.05$. 'RF' in the subscript denotes IRD tests utilising the symmetrical-difference distance of Robinson and Foulds (1981), while 'Match' in the subscript denotes tests utilising the Matching distance of Linn et al. (2012). Comparisons are either made using majority rule consensus trees (MR), or the mean nearest neighbour distance between each tree in one set and its nearest neighbour in the other.

```
Data file(s):
Nexus files and partition data
Basic script for performing IRD tests in conjunction with
PAUP*
PARTRELAT is a Basic script for determining which of two
character matrix partitions (e.g., craniodental or post-
cranial) yields trees most similar to those derived from the
entire matrix.
Supplemenatry Materials 1 – Table of overlap between
matrices
```

Thank you for submitting your data package to Dryad for
journal review. Please read the following information
carefully and save it for future reference.

(1) Your data package is temporarily available to journal
editors and reviewers via a private review link (see below).

(2) It has been assigned a Digital Object Identifier (DOI),
but this DOI will not work until/unless your manuscript is
accepted for publication and your data package is approved
by our curators.

(3) If/when your manuscript is accepted, please make sure
that the DOI, and NOT the review link, is included in the
published article. More information about citing your data
is below.

(4) If you need to modify your submission, please see the
instructions below. Do not create multiple Dryad submissions
associated with the same manuscript.

(5) If you were asked to enter a credit card number, your
card will not be charged unless/until your manuscript is
accepted and your submission is approved. You will receive a
confirmation when your card is charged.

YOUR TEMPORARY REVIEW LINK:

https://datadryad.org/review?doi=doi:10.5061/dryad.05961bm

(for journal access prior to manuscript acceptance only, DO
NOT link in your manuscript)

YOUR DRYAD DOI:

doi:10.5061/dryad.05961bm

(for inclusion in the published article)