How does dinosaur diversity through time change through time?

Jonathan P Tennant [Corresp., 1] , Alfio A. Chiarenza [1]

[1] Earth Science and Engineering, Imperial College London, London, United Kingdom

Corresponding Author: Jonathan P Tennant
Email address: jon.tennant.2@gmail.com

Our knowledge of the number of dinosaur species is dependent on the publication record, and how this taxonomic archive varies through historical time. Recently, a number of high-profile studies have analysed patterns in dinosaur diversity that assess the raw numbers of species, also accounting for ecological changes in the shape of the species abundance curve, both of which are dependent on databases complied from the primary published literature. However, how the shape of these 'diversity curves' is influenced by the history of publication remains largely unknown. Here, we investigate the influence of publication history between 1991 and 2015 on our understanding of dinosaur evolution using raw diversity estimates and Shareholder Quorum Subsampling for the three major subgroups, Ornithischia, Sauropodomorpha and Theropoda. In all three groups, the shape of raw global diversity through publication time remains intriguingly consistent. While theropod and sauropod diversity both increase notably throughout publication time, most apparent in the Cretaceous, these changes are fairly evenly distributed throughout the Cretaceous and do not distort the overall shape of diversity. When subsampling is applied, the relative magnitude of these global changes is greatly emphasised, highlighting a major decline in ornithischian diversity through the Early-Late Cretaceous transition, a major decline in sauropod diversity in the Campanian-Maastrichtian, and a double-dip decline in theropod diversity through the 'middle' Cretaceous, each of which become increasingly more emphasised through publication history. The continental signal reflects this global pattern too, with changes in subsampled diversity through time consistently indicating greater changes than the raw data reveal. When comparing each of these temporal diversity curves to sea level, we find that the correlation strength varies through publication time, is consistently weakly negative, but overall has decreased. Our results suggest that historical changes in database compilation, particularly in terms of the publication of additional specimens of previously identified species, affects the shape of macroevolutionary patterns for dinosaurs and our interpretations of the processes that govern them.

1 **How does dinosaur diversity through time change through time?**

2 Jonathan P. Tennant[1*], Alfio Alessandro Chiarenza[1]

3 [1]Department of Earth Science and Engineering, Imperial College London, London, SW6 2AZ, United

4 Kingdom.

5 *Corresponding author: jon.tennant.2@gmail.com

7 **Abstract**

8 Our knowledge of the number of dinosaur species is dependent on the publication record, and how this

9 taxonomic archive varies through historical time. Recently, a number of high-profile studies have analysed

10 patterns in dinosaur diversity that assess the raw numbers of species, also accounting for ecological

11 changes in the shape of the species abundance curve, both of which are dependent on databases

12 complied from the primary published literature. However, how the shape of these 'diversity curves' is

13 influenced by the history of publication remains largely unknown. Here, we investigate the influence of

14 publication history between 1991 and 2015 on our understanding of dinosaur evolution using raw

15 diversity estimates and Shareholder Quorum Subsampling for the three major subgroups, Ornithischia,

16 Sauropodomorpha and Theropoda. In all three groups, the shape of raw global diversity through

17 publication time remains intriguingly consistent. While theropod and sauropod diversity both increase

18 notably throughout publication time, most apparent in the Cretaceous, these changes are fairly evenly

19 distributed throughout the Cretaceous and do not distort the overall shape of diversity. When

20 subsampling is applied, the relative magnitude of these global changes is greatly emphasised, highlighting

21 a major decline in ornithischian diversity through the Early-Late Cretaceous transition, a major decline in

22 sauropod diversity in the Campanian-Maastrichtian, and a double-dip decline in theropod diversity

23 through the 'middle' Cretaceous, each of which become increasingly more emphasised through

24 publication history. The continental signal reflects this global pattern too, with changes in subsampled

25 diversity through time consistently indicating greater changes than the raw data reveal. When comparing

26 each of these temporal diversity curves to sea level, we find that the correlation strength varies through

27 publication time, is consistently weakly negative, but overall has decreased. Our results suggest that

28 historical changes in database compilation, particularly in terms of the publication of additional specimens

29   of previously identified species, affects the shape of macroevolutionary patterns for dinosaurs and our

30   interpretations of the processes that govern them.

32   **Introduction**

33   In the latter half of the 20th Century, palaeobiology underwent a renaissance by adopting a more

34   quantitative analytical approach to understanding changes in the fossil record through time (Raup 1972;

35   Raup 1975; Sepkoski Jr 1982; Sepkoski Jr 1992; Sepkoski Jr 1996; Sepkoski Jr et al. 1981; Sepkoski Jr et al.

36   2002). This seminal work was largely focussed around estimating patterns of animal diversity, extinction

37   and speciation through time, and what the external controls and processes governing these were. To this

38   day, reconstructing the diversity of life through geological time remains one of the most crucial aspects

39   of palaeobiology. These earlier analyses were largely based on an archive of range-through taxa of marine

40   animals, known as the 'Sepkoski Compendium'. More recently, analytical palaeobiology has had a second

41   wave of innovation, in part thanks to development of large fossil occurrences databases such as the

42   Paleobiology Database (www.paleobiodb.org) and Fossilworks (fossilworks.org), and also due to

43   development of increasingly sophisticated analytical subsampling (Alroy 2000a; Alroy 2003; Alroy 2010a;

44   Starrfelt & Liow 2016) and modelling (Lloyd 2012; Smith & McGowan 2007) techniques. There are 1,920

45   published records in Google Scholar with the term "Paleobiology Database" in as of 27 April 2017,

46   highlighting its importance in research.

47   All of these studies, both older and more recent, are under-pinned by a single principle, in that they rely

48   on counts of the number of taxa present through geological time. Despite meticulous work to ensure that

49   these databases and compendia represent the best possible records of historical trends, there have been

50   continuing questions as to the accuracy of the data and the existence of what is broadly termed as 'bias'.

51   This includes factors such as sampling intensity, different sampling availability, and variable depth of

52   taxonomic research (Raup 1972; Raup 1976). In 1993, Sepkoski added an additional dimension to these

53   diversity studies by assessing how database compilation history through changes in taxonomy,

54   stratigraphic resolution, and sampling influences the shape of macroevolutionary patterns (Sepkoski Jr

55   1993). This analysis was based on comparing two records built in 1982 and 1992, and found that in spite

56   of numerous taxonomic changes over ten years, the overall pattern structure for marine animals remained

57   relatively constant, with the main notable change being that overall diversity was consistently higher in

58   the 1992 compilation. Following this example, Alroy (2000b) further showed that database age does have

59   an influence on North American mammal diversity estimates, but this apparent phenomenon has been

60   largely ignored since then.

61   At the present, there are three main arguments regarding the historical reliability of diversity curves:

62   firstly, that because independent datasets yield similar diversity curves, this suggests that convergence on

63   a common signal indicates accuracy; secondly, that the addition of new data to existing compilations

64   should yield only minor changes to resulting diversity estimates; and thirdly, that the addition of new data

65   can potentially dramatically alter shape of diversity. Since Sepkoski's library work, surprisingly little

66   consideration has been given to how publication or database history can influence macroevolutionary

67   patterns (Alroy 2000b), despite an enormous reliance on their research utility. In particular, to our

68   knowledge, no one has yet tested the potential influence using an occurrence-based dataset, such as

69   those available from the Paleobiology Database or Fossilworks. This is important, given that a wealth of

70   recent studies, and in particular on tetrapod groups, have been focussed on estimating diversity patterns

71   through geological time, with many using occurrence-based subsampling protocols, and interpreting what

72   the potential drivers of these large-scale evolutionary patterns might be (Benson et al. 2016a; Benson &

73   Butler 2011; Benson et al. 2016b; Brocklehurst et al. 2017; Butler et al. 2009; Butler et al. 2011; Grossnickle

74   & Newham 2016; Mannion et al. 2015; Nicholson et al. 2015; Nicholson et al. 2016; Tennant et al. 2016a;

75   Tennant et al. 2016b). Many of these studies employ subsampling methods that are sensitive to changes

76   in the shape of the taxonomic abundance distribution, which we would expect to change in a non-random

77   fashion based on new discoveries through time as they are published (Benton 2015; Benton et al. 2011;

78   Benton et al. 2013).

79   While the data used in these analyses is typically based on a 'mature' dataset, that has undergone rigorous

80   taxonomic scrutiny and data addition or refinement, they usually neglect explicit consideration of the

81   potential influence of temporal variations in the publication record (which these databases are explicitly

82   based on). This has important implications for several reasons. Firstly, we expect the shape of raw diversity

83   curves to change (either uniformly, randomly, or structurally) through time in concert with new taxonomic

84   discoveries and as sampling increases (Alroy 2000b; Sepkoski Jr 1993). Secondly, we therefore expect that

85   our interpretations of the relative magnitude, tempo and mode of radiations and extinctions to change as

86   these discoveries are made. As our data are refined, capturing this influence of sampling variation

87   becomes more important through longer periods of time. Thirdly, as the shape of estimated diversity

88   curves change (either based on raw or 'corrected' data), we expect that results from analyses based on

89   model-fitting   or   pairwise   correlation   analyses   with   extrinsic   factors   such   as   sea-level   or

90   palaeotemperature (Butler et al. 2011; Mannion et al. 2015; Martin et al. 2014; Nicholson et al. 2015;

91   Tennant et al. 2016a; Tennant et al. 2016b) will change. Recently, this was highlighted by Jouve et al.

92   (2017), who tested the conclusions of Martin et al. (2014) and their assertion that sea-surface

93   temperature was the primary factor driving marine crocodylomorph evolution, contra Mannion et al.

94   (2015) and Tennant et al. (2016a). They found that the correlations reported by the first study, also non-

95   replicable by (Mannion et al. 2015) and (Tennant et al. 2016a), were fairly unstable even based on very

96   recent changes in taxonomy, highlighting how small changes in publication history can lead to mixed

97   interpretations of macroevolutionary patterns.

98   In this study, we investigate the influence of how publication history can affect our reading and

99   understanding of diversity analyses through time. This study does not discuss the following: (1) what time-

100   binning methods are appropriate for the fossil record; (2) what analytical methods are optimal for

101   accounting for the incompleteness of the fossil record, and (3) what the impact of secular variation in the

102   rock record through space and time have on our understanding of diversity, and these factors are

103   appropriately discussed in more detail elsewhere (Benson & Mannion 2011; Benson & Upchurch 2013;

104   Benson & Butler 2011; Benson et al. 2016b; Benton 2015; Benton et al. 2013; Dunhill et al. 2014; Heim &

105   Peters 2011; Peters & Heim 2010; Peters & Heim 2011; Tennant et al. 2016a; Tennant et al. 2016b). We

106   want to test a simple hypothesis: How does our understanding of diversity patterns change through time

107   as we discover more new taxa and additional fossil occurrences of those taxa in time and space. For this,

108   we elected to use the clade Dinosauria (excluding Aves) as a study group, as more than almost any other

109   vertebrate group they have received an intensely sampled fossil record, and have a rich history of

110   taxonomy and macroevolutionary research.

111

112   **Material and Methods**

113   *Dinosaur occurrences dataset*

114   We used a pre-existing dataset of dinosaur body fossil occurrences drawn from the Paleobiology Database

115   that spans the entirety of the Jurassic to Cretaceous (201–66 Ma) (Carrano et al. 2015; Tennant et al.

116   2016b). This comprised only body fossil remains, and excluded ootaxa and ichnotaxa. This dataset was

117   divided into the three major clades, Sauropodomorpha (SI 1), Ornithischia (SI 2), and Theropoda

118   (excluding Aves) (SI 3), and resolution is provided at the genus level. We excluded Aves as, due the

119   differences in the skeletons of volant taxa, they have a fossil record dominated by exceptional modes of

120   preservation (Brocklehurst et al. 2012; Dean et al. 2016). We elected to use genera, as this means that we

121   can integrate occurrences that are resolved only to the genus level (e.g., *Allosaurus* sp.), and therefore

122   include a substantial volume of data that would be lost at any finer resolution (Robeck et al. 2000). A

123   potential issue with this approach is that many dinosaur genera are multispecific, but this is randomly

124   distributed throughout our dataset and therefore should not have any substantial impact on resulting

125   curves. It has also been repeatedly demonstrated that the shape of species and genus curves are strongly

126   correlated in spite of differential taxonomic treatment (Alroy 2000b; Butler et al. 2011; Mannion et al.

127   2015), and therefore a genus level compilation should be sufficient for the scope of the present study.

128   These data were based on a comprehensive data compilation effort from multiple workers over a long

129   period of time, and represents the state of the art in dinosaur taxonomy and palaeontology at this time.

130   We elected to use a stage-level binning method based upon the Standard European Stages and absolute

131   dates provided by Gradstein et al. (2012), as opposed to a more equal-length time binning approach as

132   others have used recently (Benson et al. 2016a; Mannion et al. 2015). For our study, it is less important

133   what time binning scheme we use relative to the consistent treatment of it across different publication

134   intervals. Only body fossil occurrences that had their complete stratigraphic range contained within a

135   single stage bin were included in order to avoid the over-counting of taxa or occurrences that have poorly

136   constrained temporal durations or contained within multiple time bins. Each dinosaurian sub-group was

137   further sub-divided into approximately contiguous palaeocontinental regions: Africa, Asia, Europe, South

138   America, and North America (Mannion et al. 2015). Unfortunately, sampling is too poor to analyse

139   patterns in Antarctica, Australasia, or Indo-Madagascar, although these regions remain included in the

140   global analyses.

141

142   *Calculating diversity through time*

143   To test how diversity changes through time, we reduced this primary dataset by successively deleting data

144   from publications of occurrences recursively at 2 year intervals. Note that this is not the same as the date

145   that the entries were made into the database, but the explicit date of publication of that occurrence

146   record in the published version of record. We stopped at 1991, giving 12 sequential temporal datasets for

147   each dinosaurian clade. What this represents is the maturity of the dataset with respect to its present

148   state based on publication history. Two methods were used to assess diversity patterns. Firstly, empirical

149   diversity based on raw in-bin counts of taxa. This method has been repeatedly shown to be a 'biased' or

150   poor estimator of true diversity as it is influenced by heterogeneous sampling (Benson & Upchurch 2013;

151   Benson & Butler 2011; Benson et al. 2010; Butler et al. 2013; Mannion et al. 2015; Newham et al. 2014;

152   Smith & Benson 2013; Tennant et al. 2016a; Tennant et al. 2016b). Secondly, we employed the

153   shareholder quorum subsampling (SQS) method, which was designed to account for differences in the

154   shape of the taxon-abundance curve (Alroy 2010a; Alroy 2010c), and implemented in Perl (SI 4, 5). This

155   dual method is important, as not all publications name new taxa – some add to our knowledge of existing

156   taxa by publishing on new occurrences or collections, and therefore by applying a method that accounts

157   for changes in taxonomic abundance we can see how publication history influences diversity through

158   subsampling methods.

159

160   *Model-fitting of extrinsic parameters*

161   For our model-fitting protocol, we follow the standard procedure outlined in numerous recently analytical

162   studies, by employing simple pairwise correlation tests to the residuals of detrended time series at the

163   stage level (Benson & Butler 2011; Butler et al. 2013; Butler et al. 2011; Mannion et al. 2015; Tennant et

164   al. 2016a; Tennant et al. 2016b). Residuals for each of the two environmental parameters were calculated

165   using the arima() function, which uses maximum likelihood to fit a first-order autoregressive (AR(1)) model

166   to each time series (Gardner et al. 1980). This method eliminates the potential influence of any long-term

167   background trend (i.e., a directed change in the mean value of the complete time series through time)

168   within the time series, which has the potential to artificially inflate correlation coefficients in pairwise

169   tests (Box & Jenkins 1976), and also accounts for any potential serial autocorrelation (i.e., the correlation

170   of a variable with itself through successive data points). This protocol has become standard practice now

171   for palaeontological time series analysis following its recommendation by Alroy (2000b). For sea level, we

172   used the curve of Miller et al. (2005), which has been widely applied in recent analyses of tetrapod

173   diversification (Benson et al. 2010; Butler et al. 2011; Mannion et al. 2015; Martin et al. 2014; Tennant et

174   al. 2016a; Tennant et al. 2016b), and for palaeotemperature we used the data from Prokoph et al. (2008).

175   We performed pairwise correlation tests between our diversity estimates and each environmental

176   parameter using parametric (Pearson's product moment correlation coefficient [r]) and non-parametric

177   (Spearman's rank [ρ]) tests. For each test, both the raw and adjusted p-values are reported, the latter

178   calculated using the p.adjust() function, and using the 'BH' model (Benjamini & Hochberg 1995). This

179   method accounts for the false-discovery test when performing multiple hypothesis tests with the same

180   data set, which can inflate type-2 error (i.e., in order to avoid falsely retaining a false null hypothesis; a

181   false negative). We avoided the more commonly used 'Bonferroni correction', due the undesirable

182   property it has of potentially increasing type 2 error to unacceptable levels (Nakagawa 2004). This

183   adjustment was performed on 'families' of analyses, rather than on all correlation tests together,

184   otherwise we potentially run the risk of setting the pass rate for statistical significance too low.

185   Differently to Tennant et al. (2016b), we excluded the first 5 Jurassic data points from our analyses instead

186   of treating them as missing data. We performed pairwise correlations for subsampled diversity estimates

187   at each two year iteration for each group to assess how the strength and direction of correlation changes

188   through publication history. We do not use a maximum likelihood model fitting approach as rather than

189   trying to distinguish between a set of candidate models, we are simply assessing how the strength of

190   correlations changes through publication history. All analyses were carried out in R version 3.0.2 (R

191   Development Core Team 2013).

192

193   **Results**

194   *Global patterns of raw diversity*

195   The overall shape of the raw theropod diversity curve remains consistent through publication history (Fig.

196   1A). Similar to ornithischians, we see steadily increasing Middle to Late Jurassic diversity, fluctuating

197   'middle' Cretaceous diversity, and then a major Campanian to Maastrichtian rise, where diversity remains

198   constant. The lowest diversity is the Coniacian, which is poorly sampled in terrestrial faunas. Notable

199   variations due to publication history are in the Barremian to Cenomanian, where the diversity peak

200   increases in magnitude through time, exceeding that of Late Jurassic diversity more recently.

201   Sauropodomorph empirical diversity is unique among other dinosaurian groups in that there are

202   numerous substantial changes in raw patterns through time, and particularly in the Cretaceous (Fig. 1B).

203   Late Jurassic patterns are consistent, documenting a steady increase from the Middle to the end of the

204   Jurassic, before a major decline across the J/K boundary. The Barremian to Cenomanian is approximately

205   constant in diversity, and consistently increasing through time. Diversity in the Turonian and Coniacian is

206   as low as the Berriasian, before steadily increasing until the end of the Cretaceous. The magnitude of this

207   Late Cretaceous increase has almost doubled in the last 20 years.

208   Raw ornithischian diversity (Fig. 1C) is remarkably constant and stable to changes in publication history,

209   similar to that for theropods. The magnitude of longer-term trends is obscured by the relative over-

210   sampling of the Campanian and Maastrichtian, which are almost an order of magnitude higher than any

211   other Jurassic or Cretaceous stage interval. Indeed, the Campanian shows no sign of slowing down in

212   increasing diversity, and is the highest and most rapidly increasing of any time interval. In spite of this, the

213   overall trends in raw diversity remain, with steadily increasing Middle to Late Jurassic diversity, a small

214   earliest Cretaceous decline followed by a 'middle' Cretaceous peak in the Aptian, a shallow decline into

215   the early Late Cretaceous, and a massive increase in the Campanian-Maastrichtian.

216

217   *Global patterns of subsampled diversity*

218   The overall pattern, and the magnitude of disparity through publication time, are emphasised for

219   theropod diversity (note that SQS reveals changes in relative diversity, not absolute diversity) (Fig. 2A). In

220   the Late Jurassic, we see a switch from steadily increasing subsampled diversity to a Callovian-Oxfordian

221   peak and subsequent diversity crash in the Kimmeridgian. The Tithonian peak and J/K boundary decline

222   are both consistently recovered through publication time. Diversity is highest during the Aptian than at

223   any other stage during the Jurassic/Cretaceous interval, and has doubled in the last 20 year, revealing an

224   overall decline in diversity throughout the 'middle' Cretaceous and early Late Cretaceous. Campanian

225   diversity is as high as the Cenomanian, and increases slightly towards the end of the Cretaceous, a pattern

226   that remains consistent through publication time.

227   When we look at subsampled sauropodomorph diversity, this pattern changes dramatically too (Fig. 2B).

228   The Late Jurassic shows a double dip in diversity in the Oxfordian and Tithonian that is stable through

229   publication time. Barremian, Aptian, and Cenomanian diversity has almost doubled in the last 20 years.

230   Campanian diversity has more than doubled, whereas Maastrichtian diversity has increased at a much

231   slower pace, revealing an increasing magnitude of diversity decline before the end-Cretaceous.

232   Subsampled ornithischian diversity shows a distinctly different pattern from the raw curve, both in terms

233   of overall trends, and in terms of the magnitude of the effect of publication history (Fig. 2C). The Jurassic

234   is generally too poorly sampled to reveal a constant signal, but there is evidence of a 'double-dip' decline

235   over the Jurassic/Cretaceous transition, which remains constant through publication time. This is followed

236   by a Barremian peak, an Aptian decline, and then a large Albian radiation, in which ornithischian diversity

237   is at its joint highest throughout the Cretaceous. The magnitude of this radiation has rapidly increased

238   over publication time, the result being that originally what appeared to be stable diversity over the

239   Early/Late Cretaceous transition now shows a major decline from the Albian to Coniacian. Santonian

240   subsampled diversity remains unknown, but when we see a signal emerge in the Campanian, diversity is

241   as high as the Albian, before declining by more than half into the Maastrichtian. This overall structure,

242   besides the Albian, remains consistent throughout publication time with no major perturbations to the

243   global curve.

244

245   *The geographical structure of dinosaur diversity*

246   ***Ornithischians***

247   In North America, diversity is flat and low throughout the Late Jurassic and most of the Cretaceous (Fig.

248   3A). There is a Campanian peak, and order of magnitude higher than any prior interval, which is rapidly

249   increasing through publication time. Diversity decreases from this into the Maastrichtian, in which

250   diversity has remained stable through publication time. Subsampled diversity greatly reduces the

251   magnitude of this Campanian radiation, but the major decline into the Maastrichtian remains, and is

252   constant through publication history (Fig. 3D). Raw diversity in Europe shows increasing diversity across

253   the J/K boundary, constant 'middle' Cretaceous diversity, and an increase from the Campanian to

254   Maastrichtian (Fig. 3B). Subsampled diversity reveals a shift from increasing diversity across the J/K

255   boundary to stable diversity through publication time, and gradually decreasing diversity throughout the

256   remainder of the Early Cretaceous, punctuated by a Hauterivian trough (Fig. 3E). The increasing diversity

257   from the Campanian to Maastrichtian remains, with Maastrichtian diversity increasing by almost a third

258   through publication time. Raw Asian diversity is fairly constant through the Cretaceous, until a Campanian

259   peak and Maastrichtian decline (Fig. 3C). Subsampled diversity reveals a major decline across the Early-

260   Late Cretaceous transition, and a moderate increase from the Campanian to Maastrichtian (Fig. 3F). Both

261   raw and subsampled African ornithischian diversity is too inconsistent to analyse any changes through

262   geological time or publication time. Sampling in South America is also relatively poor, with the only

263   notable change being the increasing magnitude of raw Campanian diversity and the subsequent

264   Maastrichtian diversity crash, a pattern partially reflected in the subsampled diversity patterns, with

265   Maastrichtian diversity becoming unknown.

266

267   ***Theropods***

268   In North America, raw pre-Campanian diversity is again dwarfed by the intensive sampling of latest

269   Cretaceous dinosaurs, and is therefore constantly relatively flat (Fig. 4A). Campanian and Maastrichtian

270   diversity is constantly increasing at a faster rate than any other time interval, and consistently reveals a

271   diversity decline into the end-Cretaceous. This regional pattern is completely different to the subsampled

272   diversity estimates, which reveal increasing diversity from the Kimmeridgian to Tithonian, and relatively

273   low 'middle Cretaceous' diversity in the Aptian and Albian, which is emphasised by steadily increasing

274   Barremian and Cenomanian diversity (Fig. 4D). There is a clear radiation from the Santonian to the

275   Campanian, with Campanian diversity equal to that at the end of the Jurassic, and a shallow decline into

276   the Maastrichtian, which remains consistent throughout publication history. Raw European diversity is

277   much more constant due to a more intensive sampling history (Fig. 4B). There is a Middle Jurassic diversity

278   peak in the Bathonian, followed by a Callovian-Oxfordian trough, a second larger Kimmeridgian peak, and

279   then constant decline from the Tithonian to the Valanginian. Barremian diversity is increasing rapidly

280   through publication time, and is as high as the Kimmeridgian. Aptian and Albian diversity is low.

281   Campanian diversity is slowly increasing, but Maastrichtian diversity is increasing much faster through

282   time, emphasising the rate of diversity change towards the end of the Cretaceous. When subsampling is

283   applied, the only part of this pattern that remains is the large J/K boundary decline, but instead with a

284   small bounce back in diversity during the Valanginian (Fig. 4E). In Asia, raw Late Jurassic diversity is

285   generally lower than the Cretaceous (Fig. 4C). The Cretaceous sees three peaks in diversity during the

286   Aptian, Turonian and Campanian-Maastrichtian, with the latter interval being considerably more diverse

287   than any previous one. The signal becomes much patchier when subsampling is applied, with fairly

288   constant and low diversity through the Jurassic and into the Early Late Cretaceous (Fig. 4F). The

289   Campanian was historically less diverse than the Maastrichtian, but now this picture has changed in recent

290   years and now there appears to be a diversity loss from the Campanian to Maastrichtian, with Campanian

291   diversity increasing at a faster rate than that for the Maastrichtian. Raw South American diversity is very

292   unstable, with almost every interval in which dinosaurs are available to be sampled leading to rapidly

293   increasing diversity changes through publication time, especially in the 'middle' and Late Cretaceous. This

294   relative instability is emphasised when SQS is applied, as only a single interval is sampled well enough to

295   produce a signal. Here, the Cenomanian still shows a 7-fold increase in subsampled diversity over the last

296   20 years. As with ornithischians, African theropods are generally too poorly sampled at the stage level to

297   recognise any consistent patterns. There is a Cenomanian spike in the early Late Cretaceous, but how this

298   compares to the rest of the Cretaceous is obscured by patchy sampling. However, what is clear when

299   subsampling is applied is that there is a theropod radiation across the Early-Late Cretaceous boundary,

300   with diversity almost quadrupling in Africa.

301

302   ***Sauropodomorphs***

303   The North American sauropod record is very patchy, with the latest Jurassic showing a shift from rapidly

304   increasing raw diversity to a slight drop from the Kimmeridgian to Tithonian (Fig. 5A). Subsampling

305   appears to corroborate this pattern, and the shift becomes more dramatic through time (Fig. 5B). In

306   Europe, sauropods show a consistent and major decline in raw diversity from the Kimmeridgian to

307   Berriasian (Fig. 5B). Much of the rest of the Cretaceous is too poorly sampled, but there is evidence for

308   increasing diversity towards the end of the Cretaceous. Subsampling again reduces this signal due to an

309   overall poor sampling of Cretaceous sauropods in Europe, although there is strong evidence for the

310   sauropod decline beginning prior to the Jurassic/Cretaceous boundary (Fig. 5E). There is also evidence for

311   an Aptian to Albian diversity drop, and increasing diversity from the Campanian to Maastrichtian, all of

312   which appear to be stable patterns through publication history. In Asia, raw taxonomic diversity is

313   generally low compared to the Maastrichtian interval, in which diversity is relatively high and still rapidly

314   increasing (Fig. 5C). Subsampling greatly reduces the amount of signal, but we see increasing latest

315   Jurassic and latest Cretaceous diversity with publication history (Fig. 5F). Sauropodomorph dinosaurs are

316   generally better sampled than theropods and ornithischians in Africa, but still inconsistent enough

317   through publication and geological time to produce consistent patterns that we can meaningfully

318   interpret. When subsampling is applied, the few intervals in which a signal emerges reveal a fairly constant

319   level of diversity through the Jurassic and Cretaceous, and through publication time, with the notable

320   exception being greater diversity in the latest Jurassic. The South American sauropod record is also

321   relatively poor, mostly confined to the Late Cretaceous. Here, we see Campanian, Cenomanian, and

322   Maastrichtian diversity all increasing rapidly through publication time, with this consistently revealing a

323   Campanian to Maastrichtian decline. Subsampling does not support this pattern, with instead the only

324   notable recovered pattern being that Campanian diversity is about 1.5 times that of Cenomanian diversity,

325   with both increasing steadily and equally through publication time. Full subsampling results are provided

326   in Supplementary Information 6.

327

328   *Model-fitting results*

329   ***Dinosaur diversity and sea level***

330   Our results find varying strength of correlation between subsampled dinosaur diversity for each clade and

331   sea level. Similar to Tennant et al. (2016b) who used 10 million year time bins, our results using standard

332   pairwise correlation tests at the stage level find consistently weak and non-significant relationships

333   between sea level and diversity for each of ornithischians, sauropodomorphs, and theropods. However,

334   the strength of this relationship varies through time. In 1991, half of the variance in sauropodomorph

335   diversity could be explained by sea level changes. This relationship declined consistently and even became

336   negative up until 2005, and from 2007-2015 remains a weak but positive relationship, with no more than

337   around 20-30% of variance being accounted for by sea level. In Ornithischians, the amount of variance in

338   diversity that sea level explains is consistently between 0.3 and 0.5 through time based on a Spearman's

339   test. The strength of this correlation is much more variable using a Pearson's test, ranging from -0.16 to

340   0.45 through time, but in neither of these tests do any of the correlation results approach the traditional

341   threshold for statistical significance. The relationship between theropod diversity and sea level has

342   progressively weakened through time, and shifted from being weakly positive in the 1990s to weakly

343   negative or close to zero approaching modern times. If we sum all three clades together to represent total

344   dinosaur diversity, then a striking pattern emerges. There is a consistently negative and moderate to

345   strong correlation between summed diversity and sea level, with the strength of this correlation often

346   exceeding 0.5. While the strength of this has decreased during publication time, in 2015 at least a third of

347   total dinosaur diversity could be explained by sea level, but with a negative correlation.

348

349   ***Dinosaur diversity and palaeotemperature***

350   Similar to our sea level results, we find little to no evidence of any strong or consistent relationship

351   between the diversity of any dinosaur group and palaeotemperature. In sauropodomorphs, the

352   relationship is consistently weakly negative, become slightly weaker throughout publication history. In

353   ornithischians, the pattern is similar, but the strength of the negative correlation is much stronger,

354   averaging around -0.4-0.5 through time. In theropods, the relationship varies between weak negative and

355   weak positive, never exceeding around -0.3, but becoming gradually more positive with time. When all

356   dinosaurs are looked at together, the relationship is consistently moderately strongly negative. The

357   strength of this gradually weakens through time, and never approaches the traditional threshold for

358   statistical significance. Full model fitting results are given in Supplementary Information 7.

359

360   **Discussion**

361   *The influence of geology, sampling, and publication history on dinosaur diversity*

362   The impact of dinosaur diversity changes through publication time has direct consequences for our

363   interpretation of their evolutionary history and diversification. As research on dinosaurs continues in this

364   century and new taxa are described from existing fossiliferous formations, one implication of this is that

365   raw diversity is expected to become less correlated with rock availability as result of increasing sampling

366   effort (Benton 2015; Raup 1977; Wang & Dodson 2006), and represents a bibliographic form of publication

367   bias (Alroy 2000b; Jouve et al. 2017; Sepkoski Jr 1993). In this study, we tested whether by comparing

368   successive dinosaur diversity logistic curves we are approaching the end of the exponential phase of

369   dinosaur diversity increase, making our diversity analyses for this clade more stable and reliable for further

370   examination and interpretation. What we seem to be seeing is that for raw diversity estimates, we find

371   evidence for relatively stable patterns in spite of any 'bonanza effect' (Benton 2015; Raup 1977). The fact

372   that the curves remain relatively linearly consistent despite the non-random addition of new taxa also

373   provides support for the 'redundancy' hypothesis, that fossils and sampling are non-independent from

374   each other, when only raw data are considered (Benton 2015; Benton et al. 2011; Benton et al. 2013;

375   Dunhill et al. 2014). While others are now reaching the same conclusion, at least for the Mesozoic tetrapod

376   record, this further suggests that 'correcting' diversity estimates by using static proxies for sampling is not

377   an appropriate methodology (Benton 2015; Brocklehurst 2015; Sakamoto et al. 2017).

378   However, what is the explanation for the diversity patterns we obtained so far, and what does the

379   variation in these patterns tell us? Generally, a dinosaur bearing formation availability effect makes the

380   Kimmeridgian, Barremian, Albian, Aptian, Campanian, and Maastrichtian the most productive stages

381   (Barrett et al. 2009; Butler et al. 2011; Tennant et al. 2016b; Upchurch et al. 2011). By counting genus

382   density (number of genera per million year), three stages from these stand out: Kimmeridgian, Campanian

383   and Maastrichtian (Taylor 2006), with Asia being the most productive continent followed closely by North

384   America, then Europe, South America, Africa, Australasia and finally Antarctica. There is a major

385   recognised influence of the amount of rock available for palaeontologists to search for dinosaur fossils,

386   which have consequences for our interpretations of diversity patterns. This raises questions about the

387   extent to which many aspects of diversity curves could be artefacts caused by changes in global sea levels,

388   tectonics, and other geological processes related to preservational or geological megabiases (Heim &

389   Peters 2011; Peters & Foote 2001; Peters & Heim 2010; Peters & Heim 2011; Smith & Benson 2013; Smith

390   et al. 2001; Smith et al. 2012; Smith & McGowan 2007). As a way of exploring this, Barrett et al. (2009)

391   applied the "residuals" method (formerly designed by Smith & McGowan (2007) for marine fossil taxa)

392   to account for these sorts of structural biases, and demonstrated that many features of dinosaur diversity

393   curves are sampling artefacts that reflect changes in the amount of fossiliferous rocks and thus reflect

394   geological rather than biological signals. The influence of these geological biases appears to have been

395   largely mitigated in recent studies by considering a historically accurate account of sampling and

396   modelling variation through time (Alroy 2010a; Alroy 2010b; Alroy 2010c; Grossnickle & Newham 2016;

397   Mannion et al. 2015; Newham et al. 2014; Nicholson et al. 2015; Tennant et al. 2016a; Tennant et al.

398   2016b). Here, sampling heterogeneity in terms of both collection effort and rock availability can be

399   accounted for through subsampling methods, which appear to capture and alleviate at least part of the

400   geological signal.

401   These relative changes in the amount of rock available for sampling, the number and abundance of

402   different taxa, and the historical sampling intensity of different rock formations have implications for the

403   patterns of palaeobiological change that we infer from them. We find that there are three main time

404   periods when great caution should be applied to interpreting further processes or patterns based on

405   dinosaur diversity, especially at a global level. These are: (1) the Middle-Late Jurassic interval for

406   theropods; (2) the Middle-Late Cretaceous interval for theropods; and (3) the Middle-Late Cretaceous

407   interval for ornithischians. These represent the times when diversity estimates are changing most rapidly

408   due to a combination of taxonomic revision and discovery-driven publication. While we cannot predict

409   the future of dinosaur discovery, or the fickle nature of publication, it seems prudent to suggest that we

410   are cautious in our interpretation of events in dinosaur macroevolution in these intervals. An interesting

411   extension of the present study, which explores historical publication bias, would be to test how the

412   historical context of sampling (e.g., outcrop area variation or availability through time, sampling intensity

413   through time) corresponds to our historical estimates of diversity.

414

415   *Discovery influences regional patterns of dinosaur diversity through time*

416   ***Ornithischians***

417   The Jurassic/Cretaceous (J/K) boundary decline remains constant and recognisable throughout

418   publication history, with this stability suggesting that either this is a real biological signal and not a

419   publication artefact (Tennant et al. 2016b). However, more focussed sampling needs to occur on J/K

420   boundary deposits to reveal the true global signal, as much of this pattern is based on fossils exclusively

421   from historically well-sampled European localities (Tennant et al. 2016c). Ornithischian diversity

422   decreases steadily through Early Cretaceous, with a possible radiation from the Campanian to

423   Maastrichtian, perhaps explained by an increase of recent occurrences of latest Cretaceous dinosaurian

424   findings mainly in Spain, Portugal, France and Romania (Csiki et al. 2010; Riera et al. 2009). However, many

425   of these latest Cretaceous European dinosaur faunas are not particularly well-resolved stratigraphically

426   compared to the well-studied North-American sections, which makes the timing of any regional extinction
427   here and comparison with North America and Asia difficult at the present. Despite the increasing
428   availability of Early Cretaceous DBFs in Africa in the last 20 years (e.g., Tunisia, Niger; (Anderson et al.
429   2007; Taquet & Russell 1999)), sampling here is still too limited to reveal any consistent patterns in
430   ornithischian diversity (Mannion et al. 2011; Tennant et al. 2016b; Upchurch et al. 2011).

431   Ornithischian diversity in Asia has been increasing steadily through publication time in the 'middle'
432   Cretaceous, filling in the gap from equivalent latitude European deposits at this time, and plausibly due
433   to the radiation of Parksosauridae and Ankylopollexia clades, two of the most dominant Late Cretaceous
434   dinosaurian clades around this time. Together with the North American record, this manifests as a great
435   global decline across the Early-Late Cretaceous interval, a pattern that was not recognised until more
436   recent years thanks to the discovery of more Konzentrat-Lagerstätten in Mongolia and China around this
437   time (Godefroit et al. 2008; Lambert et al. 2001; Upchurch et al. 2011). A perceived Late Cretaceous
438   diversity increase for Asian taxa, particularly hadrosauroids, could be also due to anthropogenic sampling
439   bias, following a renaissance in the discovery of Cretaceous Asian dinosaurs over the past two decades
440   (Barrett et al. 2009; Lloyd et al. 2008; Mo et al. 2016; Upchurch et al. 2011; Zhou & Wang 2010).

441   Globally, advanced ornithischian faunas appear to have hyperdiversified in the latest Cretaceous, but this
442   is classically explained by the oversampling of North American "Judithian" localities, including the Hell
443   Creek Formation and Dinosaur Park Formation, and their approximate temporal equivalents. Although a
444   small rise in diversity is recovered from the Campanian to the Maastrichtian in both Europe and Asia, this
445   is considerably less marked than the decline in North America, where subsampling reveals that
446   ornithischian diversity was actually declining from the Campanian to Maastrichtian (Brusatte et al. 2015).
447   This regional distinction could be due to the tie between ecomorphloglogical function and biological
448   diversity, as Asian hadrosauroids increased in morphological disparity during the latest Cretaceous,
449   whereas in North America large-bodied bulk-feeding ornithischians decreased in their disparity (Brusatte
450   et al. 2012; Campione & Evans 2011). Furthermore, this diversity decline in ornithischians coincided with
451   a decrease in their provinciality (i.e., beta diversity) and disruption of North American dinosaurian food
452   webs (Mitchell et al. 2012; Vavrek & Larsson 2010). In North America, several abiotic factors, including
453   extreme fluctuations of the Western Interior Sea, and the Laramide orogeny and proposed biogeographic
454   provincialism, may have affected the evolution of North America dinosaurs in distinct ways from species
455   on other continents (Arbour et al. 2016; Gates et al. 2012), meaning that the North American record is
456   unlikely to be representative of global diversity pattern (Simpson et al, 2010).

457   *Theropods*

458   As already shown elsewhere (Barrett et al., 2009, Brusatte et al., 2012), global theropod diversity trends

459   are overall very similar to that of Ornithischia, with significant diversity increases during the Late Jurassic

460   (Callovian-Oxfordian and Tithonian peaks punctuated by a Kimmeridgian crash), late Early Cretaceous

461   (Aptian), early Late Cretaceous (Cenomanian) and latest Cretaceous with a notable decline in diversity

462   during the Maastrichtian (with a less marked drop than that observed in the ornithischian before the K-

463   Pg). A notable exception to this congruence is the Late Triassic, where theropods exhibit a small peak in

464   diversity at a time when ornithischians are exceptionally rare (Butler et al. 2007; Irmis 2010; Irmis et al.

465   2007; Nesbitt et al. 2007). Middle and Late Jurassic diversity are high with a greater Late Jurassic radiation,

466   representing the radiation of major tetanuran and coelurosaurian clades, and a wealth of new discoveries

467   in recent years, particularly from Asia (Benson et al. 2014; Carrano et al. 2012; Tennant et al. 2016b;

468   Upchurch et al. 2011; Xu et al. 2011).

469   As for Ornithischia, the African theropod record is too patchy to show a consistent diversity record

470   (Upchurch et al. 2011), with some exceptions highlighted by our subsampling approach. There is a

471   Cenomanian boom mainly due to the multitaxic theropod dominated Kem Kem beds and other Albian-

472   Cenomanian ("middle" Cretaceous) equivalents in Northern Africa, but this signal might have been altered

473   by time averaging effects constraining a more timely diluted diversity in a single unit (Chiarenza & Cau

474   2016; Evers et al. 2015; Mannion & Barrett 2013). Asian diversity peaks in the Aptian, Turonian,

475   Campanian and Maastrichtian might be explained by a Lagerstätten 'bonanza' effect, especially

476   considering the high quality preservation deposits discovered and heavily sampled in the last 20 years

477   (e.g. Liaoning) (Godefroit et al. 2013; Lloyd et al. 2008; O'Connor & Zhou 2015; Tennant et al. 2016b;

478   Tennant et al. 2016c; Zhou & Wang 2010).

479   European diversity is more constant than in other regions, with a Bathonian peak followed by a Callovian-

480   Oxfordian trough, and a Kimmeridgian peak followed by a Tithonian to Valanginian drop. This can, at least

481   in part, be explained by an abundance of well-sampled Late Jurassic formations from the UK, Central

482   Europe (e.g. Solnhofen), Spain and Portugal (Benson et al. 2013; Tennant et al. 2016b; Tennant et al.

483   2016c; Upchurch et al. 2011). Barremian diversity is increasing rapidly through publication history, and is

484   now as high as calculated for the Kimmeridgian. As with the Late Jurassic, at least part of this signal

485   represents the influence of a Lagerstätten effect (e.g., Las Hoyas, Spain) (Buscalioni et al. 2008; Sánchez-

486   Hernández & Benton 2012; Upchurch et al. 2011), highlighting that single, well-sampled formations can

487   have a profound historical effect on our understanding of regional diversity patterns. The European

488   Aptian-Albian record is increasing slower through time compared to the Campanian-Maastrichtian.
489   However, this might possibly change in the future, as the ichnological record in southern Europe is quite
490   abundant for the Aptian-Albian interval, and suggests a currently unrecognised dinosaurian diversity
491   present there (Dalla Vecchia 2002; Meyer & Thuring 2003).

492   The North American pre-Campanian record is surely dwarfed by an oversampling of latest Cretaceous
493   dinosaur-bearing formations (e.g. Dinosaur Provincial Park, Hell Creek Formation). An increasingly more
494   even representation of latitudinally diverse localities from the Cenomanian-Campanian of Utah, Colorado,
495   New Mexico and Mexico (e.g. Wahweap Formation), may increase the magnitude of the high apparent
496   diversity drop through the Maastrichtian. Subsampling highlights a latest Jurassic peak in diversity (due to
497   the abundance of remains from the well-sampled Morrison Formation; (Foster 2003)), matching the
498   Campanian level of diversity. Conversely to Brusatte et al. (2015), who found no evidence for a progressive
499   Campanian-Maastrichtian decline in North American theropod faunas using similar SQS analyses
500   (implemented in R; see (Tennant et al. 2016a; Tennant et al. 2016b) and (Alroy 2010a; Alroy 2010c) for
501   comparative discussions), we find a slight decline that remains constant (around 20% loss) through
502   publication history, that likely relates to our usage of a more refined subsampling approach (Alroy 2010c).
503   Albian-Aptian diversity dwarfs the Barremian level due to the more-sampled localities from Montana to
504   Texas (Cifelli et al. 1999; Kirkland et al. 1997; Kirkland & Madsen 2007). Similarly to the pattern in Africa,
505   South America theropod diversity sparkles compared to other stages, as increasingly new taxa have been
506   discovered, often altering our knowledge of dinosaur phylogeny and biogeography from the 'middle'
507   Cretaceous of Patagonia and Brasil (Canale et al. 2009; Novas et al. 2013; Novas et al. 2005; Novas & Pol
508   2005).

509   ***Sauropods***

510   Sauropodomorph diversity patterns share some characteristics of those for theropods and ornithischians,
511   despite having an overall poorer fossil record due to taphonomic differences (Dean et al. 2016; Mannion
512   & Upchurch 2010; Mannion & Upchurch 2011). This is compounded by a difficulty in assigning a large
513   number of taxa to specific stage bins, which excludes many of them from our analyses (SI 1). These
514   similarities include diversity peaks during the Late Jurassic, late Early Cretaceous and latest Cretaceous,
515   with a major decline during the early Late Cretaceous. Numerous sauropodomorph taxa were present
516   during the late Triassic to Middle Jurassic interval, while ornithischian diversity was relatively low
517   throughout this period. This divergence in diversity patterns has classically been interpreted as being due
518   to exclusive competition between the two main herbivorous dinosaurian subtaxa (Butler et al. 2009), with

519 an explosive radiation in ornithischians during the Early Cretaceous resulting from the apparent decline

520 in diversity of sauropodomorphs. In fact, the J/K boundary represents a major extinction event for

521 sauropodomorphs, reflecting the decline of non-neosauropods, diplodocoids and basal macronarians

522 (Mannion et al. 2013; Tennant et al. 2016b). Sauropodomorph faunas have a low diversity in the earliest

523 Cretaceous, coupled with a generally poor fossil record (Mannion & Upchurch 2010), but at a time when

524 we otherwise see rapid increases in theropod and ornithischian diversity and a prolonged phase of faunal

525 turnover (Tennant et al. 2016b; Upchurch & Mannion 2012). Sauropodomorph diversity levels are

526 maintained at an almost constant level during the final latest Cretaceous radiation, and did not decline

527 prior to the end-Cretaceous boundary. This find is somewhat contrary to that of Sakamoto et al. (2016)

528 who found that their decline was initiated in the Early Cretaceous, and that the diversification of

529 titanosaurs was at an insufficient rate to compensate for the overall loss of sauropodomorph lineages

530 throughout the rest of the Cretaceous. Sauropodomorphs appear to be overrepresented with respect to

531 what we might expect based on the number of DBFs for almost the entire duration of the Jurassic, whereas

532 the opposite is true for the Cretaceous (Mannion et al. 2011; Tennant et al. 2016b; Upchurch et al. 2011).

533 Sauropodomorph diversity patterns are distinct compared to theropods and ornithischians for being

534 inconsistent with them through time, although part of this might be due to a relatively patchier fossil

535 record. The general patter of diversity follow a steady increase from Middle to the end of Jurassic with

536 major decline through J/K boundary (Tennant et al. 2016b; Upchurch & Mannion 2012). The magnitude

537 of Late Cretaceous diversity increases in the last 20 years can at least partially be explained by the constant

538 discovery of new titanosaurian taxa especially from Gondwanan continents (Bandeira et al. 2016; de Jesus

539 Faria et al. 2015; Poropat et al. 2016; Vieira et al. 2014), and only recently a more appreciated diversity of

540 diplodocoids (e.g., dicraeosaurids, rebbachisaurids) from relatively poorly sampled regions such as Africa

541 (Ibrahim et al. 2016; Mannion & Barrett 2013; Wilson & Allain 2015).

542 Large-bodied sauropod diversity in the Tithonian is certainly influenced by oversampling of the North

543 American Morrison Formation, where there is an unusually high diversity and cranial disparity of

544 megaherbivores within a relatively resource-poor environment (Button et al. 2014). In addition,

545 sauropods appear to be better sampled than theropods and ornithischians in Africa, although their

546 records remain too inconsistent and patchy record to reveal any major patterns. When subsampling is

547 applied, Jurassic and Cretaceous levels of diversity are relatively low in all regions, with the notable

548 exception of the Tithonian. Here, high diversity remains in spite of our accounting for large collection

549 biases associated with Konzentrat-Lagerstätten (Alroy 2010a; Alroy 2010c), implying that the Late Jurassic

550 was when sauropodomorphs reached their true zenith in diversity. Asian raw diversity is constantly lower

551   than both theropods and ornithischians until the Maastrichtian, where it increases substantially based on

552   a series of recent discoveries from Pakistan and China (Junchang et al. 2013; Malkani 2010). This

553   phenomenon could be explained by a taphonomic size bias discriminating against the preservation of

554   larger-bodied animals in pre-Late Cretaceous Konservat-Lagerstätten, while they are more present

555   although more rare in the dense bone assemblages from the latest Cretaceous of Mongolia, China and

556   India (Kidwell, 2001).

557   There is a marked increase in European sauropod diversity from the Campanian to Maastrichtian due to

558   relatively good sampling of latest Cretaceous deposits in western Europe (Csiki-Sava et al. 2015; Upchurch

559   et al. 2011). Application of SQS reduces an already degraded European signal because of the very poor

560   sampling through much of the Cretaceous, but there is a notable diversity decline through the J/K

561   boundary, as with other dinosaurian groups (Tennant et al. 2016b; Upchurch & Mannion 2012), but

562   conversely to results when other methods (e.g., TRiPS) are applied and don't find any evidence for a

563   decline (Starrfelt & Liow 2016). Subsampling also reverses the latest Cretaceous decline, and instead this

564   reveals a small increase in diversity in the Maastrichtian. The sauropod record in South America is poor

565   and mostly confined to the Late Cretaceous, with the Cenomanian, Campanian and Maastrichtian as the

566   best represented stages and with diversity increasing rapidly through publication time, and revealing a

567   biodiversity decline in from the Campanian to Maastrichtian. However, this pattern is not supported after

568   application of subsampling, with the Campanian being more diverse than the Cenomanian, but with both

569   steadily increasing as new taxa get identified from emerging Patagonian and Brasilian deposits (Novas,

570   2007, 2009).

571   Here, it is worth noting the distinction between global and regional records. On a global level, we appear

572   to have strong evidence for a substantial sauropod diversity crash from the Campanian to Maastrichtian.

573   However, the Campanian signal is dominated by the South American record, which, although present, is

574   lost in the Maastrichtian due to failing to reach the SQS quorum. Instead, the 'global' signal in the

575   Maastrichtian being comprised of a medley of Asian, European, and North American records, neither of

576   which reveal a regional extinction. Therefore, the 'global extinction' of sauropods in the latest Cretaceous

577   is actually due to regionally heterogeneous sampling signals that are summed into a mis-leading 'global'

578   curve. Thus, when looking at diversity signals, interpretation of global patterns without considering

579   structural changes on a regional level is not recommended. Furthermore, the 'filtering' of data from the

580   fossil record should also be considered when interpreting these patterns, as there are two main stages of

581   data loss. Firstly, taxa from geological formations that do not fit exclusively into a single time bin will be

582   excluded from data pools; this will not be in a random fashion, and will depend on the temporal longevity

583   of formations, as well as the stratigraphic resolution to which they have been resolved. Future work

584   should investigate the degree to which these can be carefully integrated into diversity analyses, in order

585   to maximise the data potential and signal. Secondly, poorly sampled time intervals that fail to meet the

586   quorum will be excluded, which becomes an issue when summing regions together to look at 'global'

587   patterns, as above, and for performing additional statistical analysis. Therefore, a range of different

588   sampling thresholds should always be considered in order to explore the extent to which a diversity signal

589   can be obtained, as well as the impact this has on the magnitude of diversity changes between intervals.

590

591   *How do interpretations of the extrinsic drivers of dinosaur diversity change through time?*

592   Our model fitting results suggests one of two things. As Tennant et al. (2016b) suggest, use of pairwise

593   correlations with such small sample sizes (n is very small for subsampled curves) and for stage level data

594   when so many data are absent, may be inappropriate in testing for correlations, and instead a maximum

595   likelihood approach that performs better with smaller sample sizes is preferable to detect optimal models

596   among candidates. Alternatively, it suggests that sea level is a poor predictor of dinosaur diversity at the

597   stage level, and that dinosaur diversity and sea level are perhaps only related on broader temporal scales

598   (Butler et al. 2011; Haubold 1990; Tennant et al. 2016b). However, what we do see, at least for

599   sauropodomorphs and ornithischians, is that the strength of the relationship between sea level and

600   subsampled diversity is strongly contingent on the publication history of the group. This supports the

601   recent analysis of Jouve et al. (2017), who also found that small changes in the taxonomic composition of

602   a dataset can lead to divergent interpretations of the environmental regulators of diversity.

603

604   *Considering the implications for publication-based databases*

605   Often, only new species are published on as these are deemed 'more impactful' and therefore more likely

606   to be published, whereas new findings of already identified species are less likely to be published on as

607   they are deemed less publishable. Indeed, many palaeontologists are aware of the 'Top journals will

608   publish anything with feathers' phenomenon over the last 20 years, where many of the most high profile

609   papers published in our field, and certainly those with media attention, have been those that describe

610   new dinosaur taxa. The presently unknown and unquantifiable effect of this is that such publication bias

611   skews our interpretation of the species-abundance distribution by being more densely favoured towards

612   new species. This is especially a problem when it comes to using data sources like the Paleobiology
613   Database (which has 283 official publications at the time of writing this;
614   https://paleobiodb.org/classic/publications?a=publications – retrieved 18/05/2017), as well as
615   automation tools like PaleoDeepDive (Peters et al. 2014), where the source data are both largely, and in
616   some cases entirely, based on publication records. Museum collections also contain numerous
617   unpublished specimens, and digitisation of these and their continued integration with databases based
618   exclusively on publication records, like the Paleobiology Database, will be essential in generating more
619   reliable datasets in the future.

620   As we have shown, the interpretation of macroevolutionary patterns in dinosaurs is highly sensitive to
621   changes in the taxon-abundance curve, and we further distort this by relying on a biased source of data
622   for our analyses. In the future, palaeontologists can alleviate this potential issue in several ways. When
623   collection fossils in the field, collect all fossils, and not just those which are easiest to collect, or potentially
624   new to science. Where this is not possible, carefully document the records that were not able to be
625   collected. Make sure that all of this ends up in a publicly accessible institute for others to re-use, and if
626   possible, digitise the information and integrate it into databases like the Paleobiology Database. Finally,
627   when publishing specimens, we should always be comprehensive and consider the broader impact we
628   have when only selecting to publish on cherry-picked specimens, and exclude data based on perceived
629   value or 'newsworthiness'. There is beauty in the bone bed, and we have everything to gain in terms of
630   research quality by ensuring that our data are based on the present state of the fossil record. After all, it
631   is already biased enough without us contributing further to it with our publication habits.

632

633   Conclusions

634   In this study, we tested diversity trends through time for three major clades of Dinosauria (Ornithischia,
635   Sauropodomorpha and Theropoda), reducing primary dataset of body fossils occurrences by progressively
636   removing publications at each 2 year intervals, up until 1991. By analysing both empirical and subsampled
637   curves, we have been able to see how publication history influences different estimates of dinosaur
638   diversity. We found that while increases in raw diversity through publication history generally do not
639   affect the overall shape of each curve, global diversity is growing much faster in all three groups during
640   the latest Cretaceous, late Early Cretaceous, and for theropods and sauropods during the Late Jurassic.
641   When we apply subsampling, we see a different structure emerge, that there are major discrepancies

642   between the 1991 and 2015 curves for theropods in the Late Jurassic, and for ornithischians and

643   theropods in the late Early Cretaceous. Sauropodomorphs, on the other hand, show almost no changes in

644   global subsampled diversity through publication history, although at least part of this is likely due to their

645   overall different taphonomy and fossil record quality. Almost without exception, these seemingly

646   continuous 'global' diversity patterns are the product of summing together different, and invariably

647   patchier, continental signals with vastly different trends. In ornithischians, a J/K boundary decline is based

648   almost exclusively on European fossils, and a perceived global reduction in their diversity in the latest

649   Cretaceous is the result of an overpowering North American signal. Similarly, 'global' subsampled

650   theropod diversity is prevalently based on the European record, with Asia and North America contributing

651   substantially more after the earliest Cretaceous. Theropod diversity in the latest Cretaceous is changing

652   the most rapidly of any other time, and therefore we should be careful not to over-interpret patterns

653   coming up to the end-Cretaceous mass extinction without considering the historical context. For

654   sauropods, an apparent 'global' decline in the latest Cretaceous is actually the result of well-sampled

655   Campanian deposits in North America being followed by relatively poorly sampled Maastrichtian deposits

656   from Laurasia. Gondwanan dinosaurian faunas are still relatively poorly sampled despite intensive

657   exploration in the last 20 years, and we expect the influence of discovery in Africa and South America to

658   become more important in the future. The consequences this appears to have on our interpretation of

659   the extrinsic drivers of 'global' dinosaur diversity are fairly minimal, although we do find that for

660   sauropodomorphs and ornithischians, the strength of correlations is dependent on their publication

661   history. The results of this study should be of interest to those who use occurrence-based compilations

662   like the Paleobiology Database that rely heavily on the published literature. Both the addition of new taxa,

663   and new occurrences of existing taxa, are clearly important in establishing stable and re-usable diversity

664   curves for further research. By ignoring the publication history, and potential biases involved in this, we

665   open ourselves up to mis-interpreting the patterns and processes involved in their evolution, and we

666   should be careful to exercise more caution as a discipline in the selectivity of our publishing behaviour.

667

668   **Acknowledgements**

669   JPT is funded by a National Environmental Research Council PhD studentship. AAC is funded by an Imperial

670   College London Janet Watson Departmental PhD Scholarship. We are grateful for the combined efforts of

671   all those who have collected Jurassic–Cretaceous dinosaur data, and to those who have entered these

672   data into the Paleobiology Database, especially J. Alroy, M. T. Carrano, R. B. J. Benson, and R. J. Butler. We

673   also thank J. Alroy for providing the Perl script used to perform SQS analyses. This is Paleobiology Database

674   official publication number XXX. JPT would like to thank Lisa Matthias (aka 'Engelchen') for putting up

675   with, and also feeding, him during the writing of this paper. We would both like to thank Serjoscha Evers

676   for helping to come up with this idea in the pub during Progressive Palaeontology 2016, and for

677   subsequent discussions, also mostly in the pub.

678

679   **Author contributions**

680   Both authors contributed to the design of the study, the analyses, drafting the figures, and writing this

681   manuscript.

682

683   **References**

684   Alroy J. 2000a. New methods for quantifying macroevolutionary patterns and processes. *Paleobiology*
685   26:707-733.
686   Alroy J. 2000b. Successive approximations of diversity curves: ten more years in the library. *Geology*
687   28:1023-1026.
688   Alroy J. 2003. Global databases will yield reliable measures of global biodiversity. *Paleobiology* 29:26-29.
689   Alroy J. 2010a. Fair sampling of taxonomic richness and unbiased estimation of origination and
690   extinction rates. In: J. Alroy GH, ed. *Quantitative Methods in Paleobiology*: The Paleontological
691   Society, 55-80.
692   Alroy J. 2010b. Geographical, environmental and intrinsic biotic controls on Phanerozoic marine
693   diversification. *Palaeontology* 53:1211-1235.
694   Alroy J. 2010c. The shifting balance of diversity among major marine animal groups. *Science* 329:1191-
695   1194.
696   Anderson PE, Benton MJ, Trueman CN, Paterson BA, and Cuny G. 2007. Palaeoenvironments of
697   vertebrates on the southern shore of Tethys: The nonmarine Early Cretaceous of Tunisia.
698   *Palaeogeography, Palaeoclimatology, Palaeoecology* 243:118-131.
699   Arbour VM, Zanno LE, and Gates T. 2016. Ankylosaurian dinosaur palaeoenvironmental associations
700   were influenced by extirpation, sea-level fluctuation, and geodispersal. *Palaeogeography,*
701   *Palaeoclimatology, Palaeoecology* 449:289-299. 10.1016/j.palaeo.2016.02.033
702   Bandeira KL, Simbras FM, Machado EB, de Almeida Campos D, Oliveira GR, and Kellner AW. 2016. A New
703   Giant Titanosauria (Dinosauria: Sauropoda) from the Late Cretaceous Bauru Group, Brazil. *PLOS*
704   *ONE* 11:e0163373.
705   Barrett PM, McGowan AJ, and Page V. 2009. Dinosaur diversity and the rock record. *Proceedings of the*
706   *Royal Society Series B: Biological Sciences* 276:2667-2674. DOI 10.1098/rspb.2009.0352
707   Benjamini Y, and Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful
708   approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)*
709   57:289-300.
710   Benson RB, Butler RJ, Alroy J, Mannion PD, Carrano MT, and Lloyd GT. 2016a. Near-stasis in the long-
711   term diversification of Mesozoic tetrapods. *PLoS Biol* 14:e1002359.

712 Benson RB, Campione NE, Carrano MT, Mannion PD, Sullivan C, Upchurch P, and Evans DC. 2014. Rates
713   of dinosaur body mass evolution indicate 170 million years of sustained ecological innovation on
714   the avian stem lineage. *PLOS Biology* 12:e1001853.
715 Benson RB, and Mannion PD. 2011. Multi-variate models are essential for understanding vertebrate
716   diversification in deep time. *Biology Letters* 8:127-130.
717 Benson RB, and Upchurch P. 2013. Diversity trends in the establishment of terrestrial vertebrate
718   ecosystems: interactions between spatial and temporal sampling biases. *Geology* 41:43-46.
719 Benson RBJ, and Butler RJ. 2011. Uncovering the diversification history of marine tetrapods: ecology
720   influences the effect of geological sampling biases. In: McGowan AJ, and Smith AB, eds.
721   *Comparing the Geological and Fossil Records: Implications for Biodiversity Studies*: Geological
722   Society of London, Special Publications, 191-208.
723 Benson RBJ, Butler RJ, Alroy J, Mannion PD, Carrano MT, and Lloyd GT. 2016b. Near-stasis in the long-
724   term diversification of Mesozoic tetrapods. *PLOS Biology* 14:e1002359.
725   10.1371/journal.pbio.1002359
726 Benson RBJ, Butler RJ, Lindgren J, and Smith AS. 2010. Mesozoic marine tetrapod diversity: mass
727   extinctions and temporal heterogeneity in geological megabiases affecting vertebrates.
728   *Proceedings of the Royal Society Series B: Biological Sciences* 277:829-834. DOI
729   10.1098/rspb.2009.1845
730 Benson RBJ, Mannion PD, Butler RJ, Upchurch P, Goswami A, and Evans SE. 2013. Cretaceous tetrapod
731   fossil record sampling and faunal turnover: Implications for biogeography and the rise of
732   modern clades. *Palaeogeography, Palaeoclimatology, Palaeoecology* 372:88-107. DOI
733   10.1016/j.palaeo.2012.10.028
734 Benton MJ. 2015. Palaeodiversity and formation counts: redundancy or bias? *Palaeontology* 58:1003-
735   1029.
736 Benton MJ, Dunhill AM, Lloyd GT, and Marx FG. 2011. Assessing the quality of the fossil record: insights
737   from vertebrates. In: McGowan AJ, and Smith AB, eds. *Comparing the Geological and Fossil
738   Records: Implications for Biodiversity Studies*. London: Geological Society of London, 63-94.
739 Benton MJ, Ruta M, Dunhill AM, and Sakamoto M. 2013. The first half of tetrapod evolution, sampling
740   proxies, and fossil record quality. *Palaeogeography, Palaeoclimatology, Palaeoecology* 372:18-
741   41. DOI 10.1016/j.palaeo.2012.09.005
742 Box GE, and Jenkins GM. 1976. *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.
743 Brocklehurst N. 2015. A simulation-based examination of residual diversity estimates as a method of
744   correcting for sampling bias. *Palaeontologia Electronica* 18:1-15.
745 Brocklehurst N, Day MO, Rubidge BS, and Fröbisch J. 2017. Olson's Extinction and the latitudinal
746   biodiversity gradient of tetrapods in the Permian. Proc R Soc B: The Royal Society. p 20170231.
747 Brocklehurst N, Upchurch P, Mannion PD, and O'Connor J. 2012. The completeness of the fossil record
748   of Mesozoic birds: implications for early avian evolution. *PLOS ONE* 7.
749 Brusatte SL, Butler RJ, Barrett PM, Carrano MT, Evans DC, Lloyd GT, Mannion PD, Norell MA, Peppe DJ,
750   and Upchurch P. 2015. The extinction of the dinosaurs. *Biological Reviews* 90:628-642.
751 Brusatte SL, Butler RJ, Prieto-Márquez A, and Norell MA. 2012. Dinosaur morphological diversity and the
752   end-Cretaceous extinction. *Nature Communications* 3:804. 10.1038/ncomms1815

753 https://www.nature.com/articles/ncomms1815#supplementary-information
754 Buscalioni A, Fregenal M, Bravo A, Poyato-Ariza F, Sanchíz B, Báez A, Moo OC, Closas CM, Evans S, and
755   Lobón JM. 2008. The vertebrate assemblage of Buenache de la Sierra (Upper Barremian of
756   Serranía de Cuenca, Spain) with insights into its taphonomy and palaeoecology. *Cretaceous
757   Research* 29:687-710.

758   Butler RJ, Barrett PM, Kenrick P, and Penn MG. 2009. Testing co‐evolutionary hypotheses over
759   geological timescales: interactions between Mesozoic non‐avian dinosaurs and cycads.
760   *Biological Reviews* 84:73-89.
761   Butler RJ, Benson RBJ, and Barrett PM. 2013. Pterosaur diversity: Untangling the influence of sampling
762   biases, Lagerstatten, and genuine biodiversity signals. *Palaeogeography, Palaeoclimatology,*
763   *Palaeoecology* 372:78-87. DOI 10.1016/j.palaeo.2012.08.012
764   Butler RJ, Benson RBJ, Carrano MT, Mannion PD, and Upchurch P. 2011. Sea level, dinosaur diversity and
765   sampling biases: investigating the 'common cause' hypothesis in the terrestrial realm.
766   *Proceedings of the Royal Society of London Series B: Biological Sciences* 278:1165-1170. DOI
767   10.1098/rspb.2010.1754
768   Butler RJ, Smith RM, and Norman DB. 2007. A primitive ornithischian dinosaur from the Late Triassic of
769   South Africa, and the early evolution and diversification of Ornithischia. *Proceedings of the Royal*
770   *Society of London B: Biological Sciences* 274:2041-2046.
771   Button DJ, Rayfield EJ, and Barrett PM. 2014. Cranial biomechanics underpins high sauropod diversity in
772   resource-poor environments. Proc R Soc B: The Royal Society. p 20142114.
773   Campione NE, and Evans DC. 2011. Cranial growth and variation in edmontosaurs (Dinosauria:
774   Hadrosauridae): implications for latest Cretaceous megaherbivore diversity in North America.
775   *PLOS ONE* 6:e25186.
776   Canale JI, Scanferla CA, Agnolin FL, and Novas FE. 2009. New carnivorous dinosaur from the Late
777   Cretaceous of NW Patagonia and the evolution of abelisaurid theropods. *Naturwissenschaften*
778   96:409-414.
779   Carrano MT, Alroy J, Mannion PD, Benson R, and Butler RJ. 2015. Taxonomic occurrences of Jurassic to
780   Cretaceous Tetrapoda recorded in , Fossilworks, the Evolution of Terrestrial Ecosystems
781   database, and the Paleobiology Database, https://paleobiodb.org/.
782   Carrano MT, Benson RB, and Sampson SD. 2012. The phylogeny of Tetanurae (Dinosauria: Theropoda).
783   *Journal of Systematic Palaeontology* 10:211-300.
784   Chiarenza AA, and Cau A. 2016. A large abelisaurid (Dinosauria, Theropoda) from Morocco and
785   comments on the Cenomanian theropods from North Africa. *PeerJ* 4:e1754.
786   Cifelli RL, Nydam RL, Gardner JD, Weil A, Eaton JG, Kirkland JI, and Madsen SK. 1999. Medial Cretaceous
787   vertebrates from the Cedar Mountain Formation, Emery County, Utah: the Mussentuchit local
788   fauna. *Vertebrate paleontology in Utah* 99:219-242.
789   Csiki-Sava Z, Buffetaut E, Ősi A, Pereda-Suberbiola X, and Brusatte SL. 2015. Island life in the Cretaceous-
790   faunal composition, biogeography, evolution, and extinction of land-living vertebrates on the
791   Late Cretaceous European archipelago. *ZooKeys* 469:1.
792   Csiki Z, Grigorescu D, Codrea V, and Therrien F. 2010. Taphonomic modes in the Maastrichtian
793   continental deposits of the Haţeg Basin, Romania—Palaeoecological and palaeobiological
794   inferences. *Palaeogeography, Palaeoclimatology, Palaeoecology* 293:375-390.
795   Dalla Vecchia F. 2002. Cretaceous dinosaurs in the Adriatic–Dinaric carbonate platform (Italy and
796   Croatia): paleoenvironmental implications and paleogeographical hypotheses. *Memorie della*
797   *Società Geologica Italiana* 57:89-100.
798   de Jesus Faria CC, Riga BG, dos Anjos Candeiro CR, da Silva Marinho T, David LO, Simbras FM, Castanho
799   RB, Muniz FP, and da Costa PVLG. 2015. Cretaceous sauropod diversity and taxonomic
800   succession in South America. *Journal of South American Earth Sciences* 61:154-163.
801   Dean CD, Mannion PD, and Butler RJ. 2016. Preservational bias controls the fossil record of pterosaurs.
802   *Palaeontology* 59:225-247. 10.1111/pala.12225
803   Dunhill AM, Hannisdal B, and Benton MJ. 2014. Disentangling rock record bias and common-cause from
804   redundancy in the British fossil record. *Nature Communications* 5:4818.

805   Evers SW, Rauhut OW, Milner AC, McFeeters B, and Allain R. 2015. A reappraisal of the morphology and
806   systematic position of the theropod dinosaur *Sigilmassasaurus* from the "middle" Cretaceous of
807   Morocco. *PeerJ* 3:e1323.
808   Foster JR. 2003. *Paleoecological Analysis of the Vertebrate Fauna of the Morrison Formation (Upper
809   Jurassic), Rocky Mountain Region, USA: Bulletin 23*: New Mexico Museum of Natural History and
810   Science.
811   Gardner G, Harvey A, and Phillips G. 1980. Algorithm AS 154: An algorithm for exact maximum likelihood
812   estimation of autoregressive-moving average models by means of Kalman filtering. *Journal of
813   the Royal Statistical Society Series C (Applied Statistics)* 29:311-322.
814   Gates TA, Prieto-Márquez A, and Zanno LE. 2012. Mountain building triggered Late Cretaceous North
815   American megaherbivore dinosaur radiation. *PLOS ONE* 7:e42135.
816   Godefroit P, Demuynck H, Dyke G, Hu D, Escuillié F, and Claeys P. 2013. Reduced plumage and flight
817   ability of a new Jurassic paravian theropod from China. *Nature Communications* 4:1394.
818   Godefroit P, Shulin H, Tingxiang Y, and Lauters P. 2008. New hadrosaurid dinosaurs from the uppermost
819   Cretaceous of northeastern China. *Acta Palaeontologica Polonica* 53:47-74.
820   Gradstein FM, Ogg JG, Schmitz M, and Ogg G. 2012. *The Geologic Time Scale 2012*: Elsevier.
821   Grossnickle DM, and Newham E. 2016. Therian mammals experience an ecomorphological radiation
822   during the Late Cretaceous and selective extinction at the K–Pg boundary. Proc R Soc B: The
823   Royal Society. p 20160256.
824   Haubold H. 1990. Dinosaurs and fluctuating sea levels during the Mesozoic. *Historical Biology* 4:75-106.
825   Heim NA, and Peters SE. 2011. Covariation in macrostratigraphic and macroevolutionary patterns in the
826   marine record of North America. *Geological Society of America Bulletin* 123:620-630.
827   Ibrahim N, Dal Sasso C, Maganuco S, Fabbri M, Martill DM, Gorscak E, and Lamanna M. 2016. Evidence
828   of a derived titanosaurian (Dinosauria, Sauropoda) in the "Kem Kem beds" of Morocco, with
829   comments on sauropod paleoecology in the Cretaceous of Africa. *Cretaceous Period: Biotic
830   Diversity and Biogeography New Mexico Museum of Natural History and Science Bulletin*, 149-
831   159.
832   Irmis RB. 2010. Evaluating hypotheses for the early diversification of dinosaurs. *Earth and Environmental
833   Science Transactions of the Royal Society of Edinburgh* 101:397-426.
834   Irmis RB, Parker WG, Nesbitt SJ, and Liu J. 2007. Early ornithischian dinosaurs: the Triassic record.
835   *Historical Biology* 19:3-22.
836   Jouve S, Mennecart B, Douteau J, and Jalil N-E. 2017. Biases in the study of relationships between
837   biodiversity dynamics and fluctuation of environmental conditions. *Palaeontologia Electronica*
838   20:1-21.
839   Junchang L, Laiping Y, Hui Z, and Xuefang W. 2013. A new somphospondylan sauropod (Dinosauria,
840   Titanosauriformes) from the Late Cretaceous of Ganzhou, Jiangxi Province of southern China.
841   *Acta Geologica Sinica (English Edition)* 87:678-685.
842   Kirkland JI, Britt B, Burge DL, Carpenter K, Cifelli R, DeCourten F, Eaton J, Hasiotis S, and Lawton T. 1997.
843   Lower to middle Cretaceous dinosaur faunas of the central Colorado Plateau: a key to
844   understanding 35 million years of tectonics, sedimentology, evolution, and biogeography.
845   *Brigham Young University Geology Studies* 42:69-104.
846   Kirkland JI, and Madsen SK. 2007. The Lower Cretaceous Cedar Mountain Formation, eastern Utah.
847   Lambert O, Godefroit P, Li H, Shang C-Y, and Dong Z-M. 2001. A new species of Protoceratops
848   (Dinosauria, Neoceratopsia) from the Late Cretaceous of Inner Mongolia (PR China). *Bulletin-
849   Institut royal des sciences naturelles de Belgique Sciences de la terre* 71:5-28.
850   Lloyd GT. 2012. A refined modelling approach to assess the influence of sampling on palaeobiodiversity
851   curves: new support for declining Cretaceous dinosaur richness. *Biology Letters* 8:123-126. DOI
852   10.1098/rsbl.2011.0210

853 Lloyd GT, Davis KE, Pisani D, Tarver JE, Ruta M, Sakamoto M, Hone DWE, Jennings R, and Benton MJ.
854   2008. Dinosaurs and the Cretaceous Terrestrial Revolution. *Proceedings of the Royal Society of*
855   *London Series B: Biological Sciences* 275:2483-2490. DOI 10.1098/rspb.2008.0715
856 Malkani MS. 2010. New Pakisaurus (Pakisauridae, Titanosauria, Sauropoda) remains, and Cretaceous
857   Tertiary (KT) boundary from Pakistan. *Sindh University Research Journal-SURJ (Science Series)* 42.
858 Mannion PD, and Barrett PM. 2013. Additions to the sauropod dinosaur fauna of the Cenomanian (early
859   Late Cretaceous) Kem Kem beds of Morocco: Palaeobiogeographical implications of the mid-
860   Cretaceous African sauropod fossil record. *Cretaceous Research* 45:49-59.
861 Mannion PD, Benson RBJ, Carrano MT, Tennant JP, Judd J, and Butler RJ. 2015. Climate constrains the
862   evolutionary history and biodiversity of crocodylians. *Nature Communications* 6:8438.
863   10.1038/ncomms9438
864 Mannion PD, and Upchurch P. 2010. Completeness metrics and the quality of the sauropodomorph fossil
865   record through geological and historical time. *Paleobiology* 36:283-302.
866 Mannion PD, and Upchurch P. 2011. A re-evaluation of the 'mid-Cretaceous sauropod hiatus' and the
867   impact of uneven sampling of the fossil record on patterns of regional dinosaur extinction.
868   *Palaeogeography, Palaeoclimatology, Palaeoecology* 299:529-540.
869 Mannion PD, Upchurch P, Barnes RN, and Mateus O. 2013. Osteology of the Late Jurassic Portuguese
870   sauropod dinosaur *Lusotitan atalaiensis* (Macronaria) and the evolutionary history of basal
871   titanosauriforms. *Zoological Journal of the Linnean Society* 168:98-206.
872 Mannion PD, Upchurch P, Carrano MT, and Barrett PM. 2011. Testing the effect of the rock record on
873   diversity: a multidisciplinary approach to elucidating the generic richness of sauropodomorph
874   dinosaurs through time. *Biological Reviews* 86:157-181. DOI 10.1111/j.1469-185X.2010.00139.x
875 Martin JE, Amiot R, Lecuyer C, and Benton MJ. 2014. Sea surface temperature contributes to marine
876   crocodylomorph evolution. *Nature Communications* 5:4658. 10.1038/Ncomms5658
877 Meyer C, and Thuring B. 2003. The First Iguanodontid Dinosaur Tracks from the Swiss Alps
878   (Schrattenkalk Formation, Aptian). *Ichnos* 10:221-228. 10.1080/10420940390256186
879 Miller KG, Kominz MA, Browning JV, Wright JD, Mountain GS, Katz ME, Sugarman PJ, Cramer BS,
880   Christie-Blick N, and Pekar SF. 2005. The Phanerozoic record of global sea-level change. *Science*
881   310:1293-1298. DOI 10.1126/science.1116412
882 Mitchell JS, Roopnarine PD, and Angielczyk KD. 2012. Late Cretaceous restructuring of terrestrial
883   communities facilitated the end-Cretaceous mass extinction in North America. *Proceedings of*
884   *the National Academy of Sciences* 109:18857-18861.
885 Mo J, Buffetaut E, Tong H, Amiot R, Cavin L, Cuny G, Suteethorn V, Suteethorn S, and Jiang S. 2016. Early
886   Cretaceous vertebrates from the Xinlong Formation of Guangxi (southern China): a review.
887   *Geological Magazine* 153:143-159.
888 Nakagawa S. 2004. A farewell to Bonferroni: the problems of low statistical power and publication bias.
889   *Behavioral Ecology* 15:1044-1045.
890 Nesbitt SJ, Irmis RB, and Parker WG. 2007. A critical re‐evaluation of the Late Triassic dinosaur taxa of
891   North America. *Journal of Systematic Palaeontology* 5:209-243.
892 Newham E, Benson R, Upchurch P, and Goswami A. 2014. Mesozoic mammaliaform diversity: The effect
893   of sampling corrections on reconstructions of evolutionary dynamics. *Palaeogeography,*
894   *Palaeoclimatology, Palaeoecology* 412:32-44.
895 Nicholson DB, Holroyd PA, Benson RB, and Barrett PM. 2015. Climate-mediated diversification of turtles
896   in the Cretaceous. *Nature Communications* 6:1-8. 10.1038/ncomms8848
897 Nicholson DB, Holroyd PA, Valdes P, and Barrett PM. 2016. Latitudinal diversity gradients in Mesozoic
898   non-marine turtles. *Royal Society Open Science* 3:160581.
899 Novas FE, Agnolín FL, Ezcurra MD, Porfiri J, and Canale JI. 2013. Evolution of the carnivorous dinosaurs
900   during the Cretaceous: the evidence from Patagonia. *Cretaceous Research* 45:174-215.

901 Novas FE, de Valais S, Vickers-Rich P, and Rich T. 2005. A large Cretaceous theropod from Patagonia,
902   Argentina, and the evolution of carcharodontosaurids. *Naturwissenschaften* 92:226-230.
903 Novas FE, and Pol D. 2005. New evidence on deinonychosaurian dinosaurs from the Late Cretaceous of
904   Patagonia. *Nature* 433:858-861.
905 O'Connor J, and Zhou Z. 2015. Early evolution of the biological bird: perspectives from new fossil
906   discoveries in China. *Journal of Ornithology* 156:333-342.
907 Peters SE, and Foote M. 2001. Biodiversity in the Phanerozoic: a reinterpretation. *Paleobiology* 27:583-
908   601.
909 Peters SE, and Heim NA. 2010. The geological completeness of paleontological sampling in North
910   America. *Paleobiology* 36:61-79.
911 Peters SE, and Heim NA. 2011. Macrostratigraphy and macroevolution in marine environments: testing
912   the common-cause hypothesis. In: McGowan AJ, and Smith AB, eds. *Comparing the Geological*
913   *and Fossil Records: Implications for Biodiversity Studies*. London: Geological Society of London,
914   95-104.
915 Peters SE, Zhang C, Livny M, and Ré C. 2014. A machine reading system for assembling synthetic
916   paleontological databases. *PLOS ONE* 9:e113523.
917 Poropat SF, Mannion PD, Upchurch P, Hocknull SA, Kear BP, Kundrát M, Tischler TR, Sloan T, Sinapius
918   GH, and Elliott JA. 2016. New Australian sauropods shed light on Cretaceous dinosaur
919   palaeobiogeography. *Scientific Reports* 6.
920 Prokoph A, Shields GA, and Veizer J. 2008. Compilation and time-series analysis of a marine carbonate
921   δ_{18}O, δ_{13}C, ^{87}Sr/^{86}Sr and δ^{34}S database through Earth history. *Earth-Science Reviews* 87:113-133.
922   http://dx.doi.org/10.1016/j.earscirev.2007.12.003
923 R Development Core Team. 2013. R: A language and environment for statistical computing. R
924   Foundation for Statistical Computing. Vienna, Austria.
925 Raup DM. 1972. Taxonomic diversity during the Phanerozoic. *Science* 177:1065-1071.
926 Raup DM. 1975. Taxonomic diversity estimation using rarefaction. *Paleobiology* 1:333-342.
927 Raup DM. 1976. Species diversity in the Phanerozoic: an interpretation. *Paleobiology* 2:289-297.
928 Raup DM. 1977. Systematists follow the fossils. *Paleobiology* 3:328-329.
929 Riera V, Oms O, Gaete R, and Galobart À. 2009. The end-Cretaceous dinosaur succession in Europe: the
930   Tremp Basin record (Spain). *Palaeogeography, Palaeoclimatology, Palaeoecology* 283:160-171.
931 Robeck HE, Maley CC, and Donoghue MJ. 2000. Taxonomy and temporal diversity patterns. *Paleobiology*
932   26:171-187. Doi 10.1666/0094-8373(2000)026<0171:Tatdp>2.0.Co;2
933 Sakamoto M, Benton MJ, and Venditti C. 2016. Dinosaurs in decline tens of millions of years before their
934   final extinction. *Proceedings of the National Academy of Sciences of the United States of America*
935   113:5036-5040. 10.1073/pnas.1521478113
936 Sakamoto M, Venditti C, and Benton MJ. 2017. 'Residual diversity estimates' do not correct for sampling
937   bias in palaeodiversity data. *Methods in Ecology and Evolution* 8:453-459. 10.1111/2041-
938   210X.12666
939 Sánchez-Hernández B, and Benton MJ. 2012. Filling the ceratosaur gap: A new ceratosaurian theropod
940   from the Early Cretaceous of Spain. *Acta Palaeontologica Polonica* 59:581-600.
941 Sepkoski Jr JJ. 1982. Mass extinctions in the Phanerozoic oceans: a review. *Geological Society of America*
942   *Special Papers* 190:283-290.
943 Sepkoski Jr JJ. 1992. A compendium of fossil marine animal families. *Contributions in Biology and*
944   *Geology Milwaukee Public Museum* 83:1-156.
945 Sepkoski Jr JJ. 1993. Ten years in the library: new data confirm paleontological patterns. *Paleobiology*
946   19:43-51.

947   Sepkoski Jr JJ. 1996. Patterns of Phanerozoic extinction: a perspective from global data bases. In:
948   Walliser OH, ed. *Global Events and Event Stratigraphy in the Phanerozoic*: Springer Berlin
949   Heidelberg, 35-51.
950   Sepkoski Jr JJ, Bambach RK, Raup DM, and Valentine JW. 1981. Phanerozoic marine diversity and the
951   fossil record. *Nature* 293:435-437.
952   Sepkoski Jr JJ, Jablonski D, and Foote M. 2002. A compendium of fossil marine animal genera. *Bulletins*
953   *of American Paleontology* 363:1-563.
954   Smith AB, and Benson RBJ. 2013. Marine diversity in the geological record and its relationship to
955   surviving bedrock area, lithofacies diversity, and original marine shelf area. *Geology* 41:171-174.
956   Doi 10.1130/G33773.1
957   Smith AB, Gale AS, and Monks NE. 2001. Sea-level change and rock-record bias in the Cretaceous: a
958   problem for extinction and biodiversity studies. *Paleobiology* 27:241-253.
959   Smith AB, Lloyd GT, and McGowan AJ. 2012. Phanerozoic marine diversity: rock record modelling
960   provides an independent test of large-scale trends. *Proceedings of the Royal Society of London*
961   *Series B: Biological Sciences* 279:4489-4495. DOI 10.1098/rspb.2012.1793
962   Smith AB, and McGowan AJ. 2007. The shape of the Phanerozoic marine palaeodiversity curve: How
963   much can be predicted from the sedimentary rock record of western Europe? *Palaeontology*
964   50:765-774. DOI 10.1111/j.1475-4983.2007.00693.x
965   Starrfelt J, and Liow LH. 2016. How many dinosaur species were there? Fossil bias and true richness
966   estimated using a Poisson sampling model. *Philosophical Transactions of the Royal Society B:*
967   *Biological Sciences* 371:20150219.
968   Taquet P, and Russell DA. 1999. A massively-constructed iguanodont from Gadoufaoua, Lower
969   Cretaceous of Niger. Annales De Paleontologie: Elsevier. p 85-96.
970   Taylor MP. 2006. Dinosaur diversity analysed by clade, age, place and year of description. Ninth
971   international symposium on Mesozoic terrestrial ecosystems and biota, Manchester, UK
972   Cambridge Publications Natural History Museum, London, UK. p 134-138.
973   Tennant JP, Mannion PD, and Upchurch P. 2016a. Environmental drivers of crocodyliform extinction
974   across the Jurassic/Cretaceous transition. *Proc R Soc B* 283:20152840.
975   Tennant JP, Mannion PD, and Upchurch P. 2016b. Sea level regulated tetrapod diversity dynamics
976   through the Jurassic/Cretaceous interval. *Nature Communications* 7.
977   Tennant JP, Mannion PD, Upchurch P, Sutton MD, and Price GD. 2016c. Biotic and environmental
978   dynamics through the Late Jurassic–Early Cretaceous transition: evidence for protracted faunal
979   and ecological turnover. *Biological Reviews*.
980   Upchurch P, Mannion P, Benson R, Butler R, and Carrano M. 2011. Geological and anthropogenic
981   controls on the sampling of the terrestrial fossil record: a case study from the Dinosauria. In:
982   McGowan AJ, and Smith AB, eds. *Comparing the Geological and Fossil Records: Implications for*
983   *Biodiversity Studies*. London: Geological Society of London, 209-240.
984   Upchurch P, and Mannion PD. 2012. Dinosaur evolution and extinction across the Jurassic/Cretaceous
985   boundary. *Actas de V Jornadas Internacionales sobre Paleontología de Dinosaurios y su Entorno,*
986   *Salas de los Infantes, Burgos*:135-155.
987   Vavrek MJ, and Larsson HC. 2010. Low beta diversity of Maastrichtian dinosaurs of North America.
988   *Proceedings of the National Academy of Sciences* 107:8265-8268.
989   Vieira WLS, Vieira KS, Nobrega RP, Montenegro PFGP, Pereira Filho GA, Santana GG, Alves RRN, Almeida
990   WO, and Vasconcellos A. 2014. Species richness and evidence of random patterns in
991   assemblages of South American Titanosauria during the Late Cretaceous (Campanian–
992   Maastrichtian). *PLOS ONE* 9:e108307.
993   Wang SC, and Dodson P. 2006. Estimating the diversity of dinosaurs. *Proceedings of the National*
994   *Academy of Sciences* 103:13601-13605.

995 Wilson JA, and Allain R. 2015. Osteology of Rebbachisaurus garasbae Lavocat, 1954, a diplodocoid
996   (Dinosauria, Sauropoda) from the early Late Cretaceous–aged Kem Kem beds of southeastern
997   Morocco. *Journal of Vertebrate Paleontology* 35:e1000701.
998 Xu X, You H, Du K, and Han F. 2011. An *Archaeopteryx*-like theropod from China and the origin of
999   Avialae. *Nature* 475:465-470.
1000 Zhou Z, and Wang Y. 2010. Vertebrate diversity of the Jehol Biota as compared with other lagerstätten.
1001   *Science China Earth Sciences* 53:1894-1907.

1002

1003

1004 **Figure captions**

1005 Figure 1: Raw global diversity for a) Theropoda, b) Sauropodomorpha, and c) Ornithischia based on our

1006 published knowledge in 1991 and 2015.

1007 Figure 2: Subsampled global diversity for a) Theropoda, b) Sauropodomorpha, and c) Ornithischia based

1008 on our published knowledge in 1991 and 2015.

1009 Figure 3: Raw and subsampled regional ornithischian diversity for North America (A, D), Europe (B, E), and

1010 Asia (C, F) based on our published knowledge in 1991 and 2015.

1011 Figure 4: Raw and subsampled regional theropod diversity for North America (A, D), Europe (B, E), and

1012 Asia (C, F) based on our published knowledge in 1991 and 2015.

1013 Figure 5: Raw and subsampled regional sauropodomorph diversity for North America (A, D), Europe (B,

1014 E), and Asia (C, F) based on our published knowledge in 1991 and 2015.

Figure 1

Figure 1

Figure 1: Raw global diversity for a) Theropoda, b) Sauropodomorpha, and c) Ornithischia based on our published knowledge in 1991 and 2015.
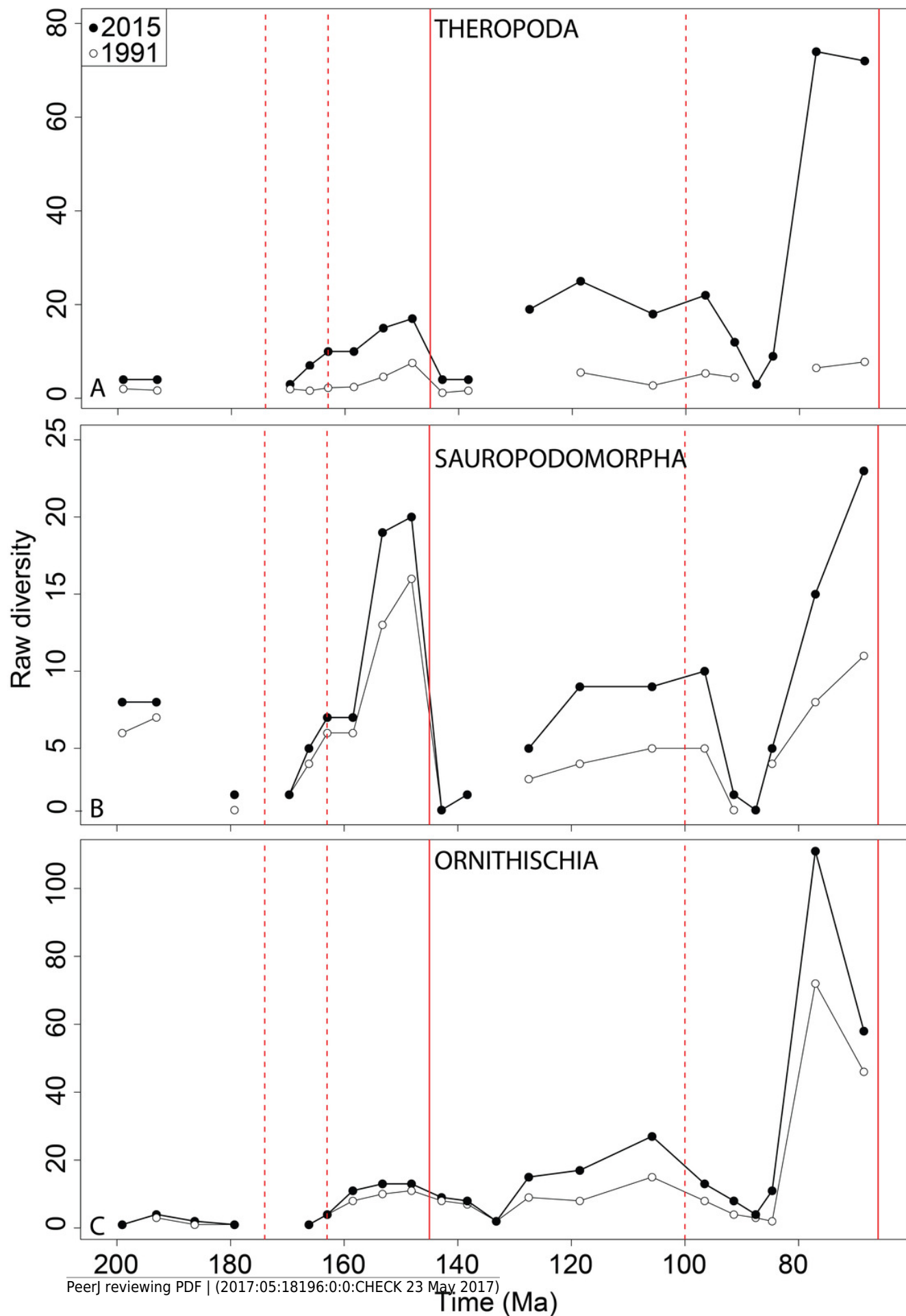
Figure 2

Subsampled global diversity

Figure 2: Subsampled global diversity for a) Theropoda, b) Sauropodomorpha, and c) Ornithischia based on our published knowledge in 1991 and 2015.
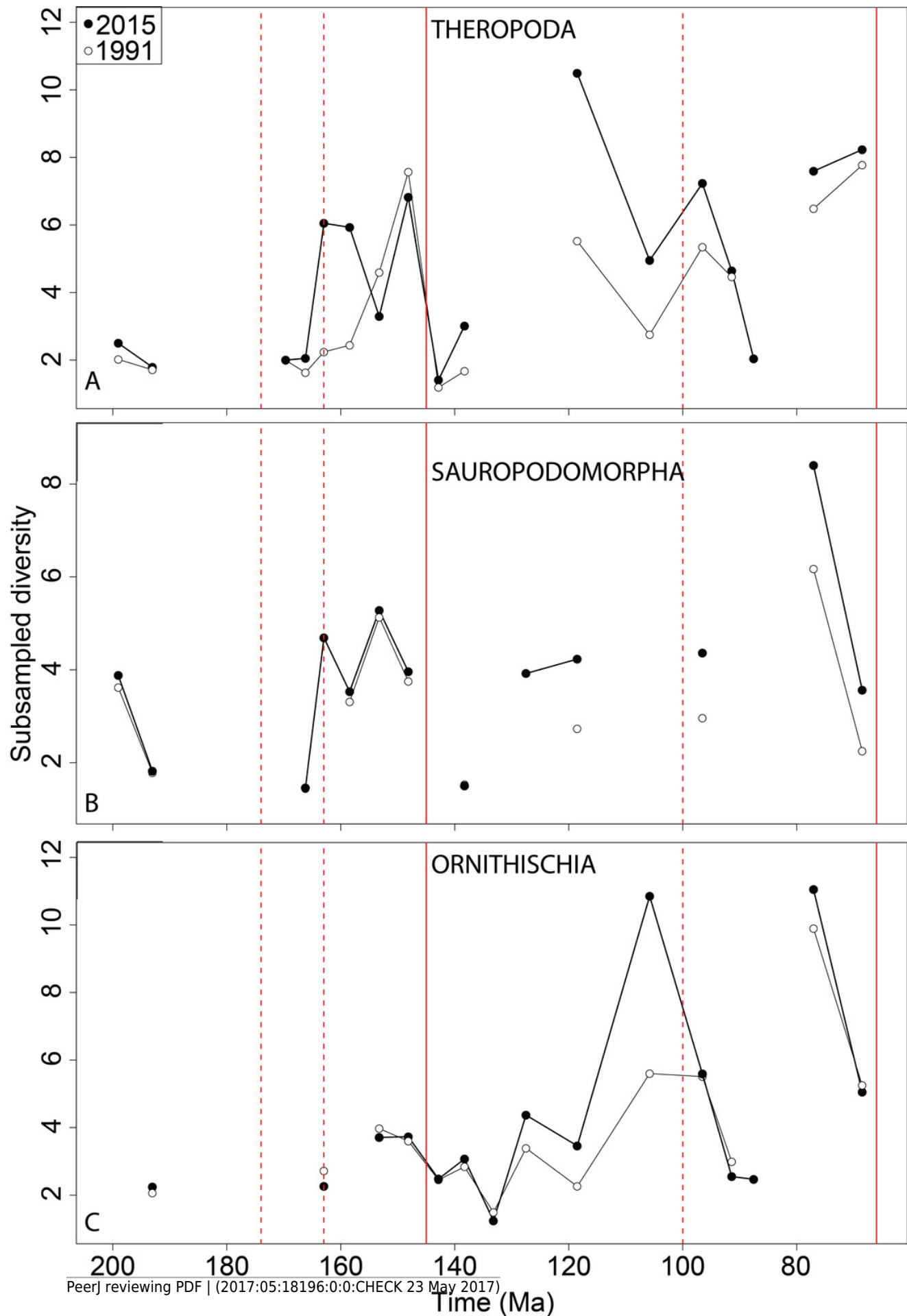
Figure 3

Regional ornithiscian diversity

Figure 3: Raw and subsampled regional ornithischian diversity for North America (A, D), Europe (B, E), and Asia (C, F) based on our published knowledge in 1991 and 2015.
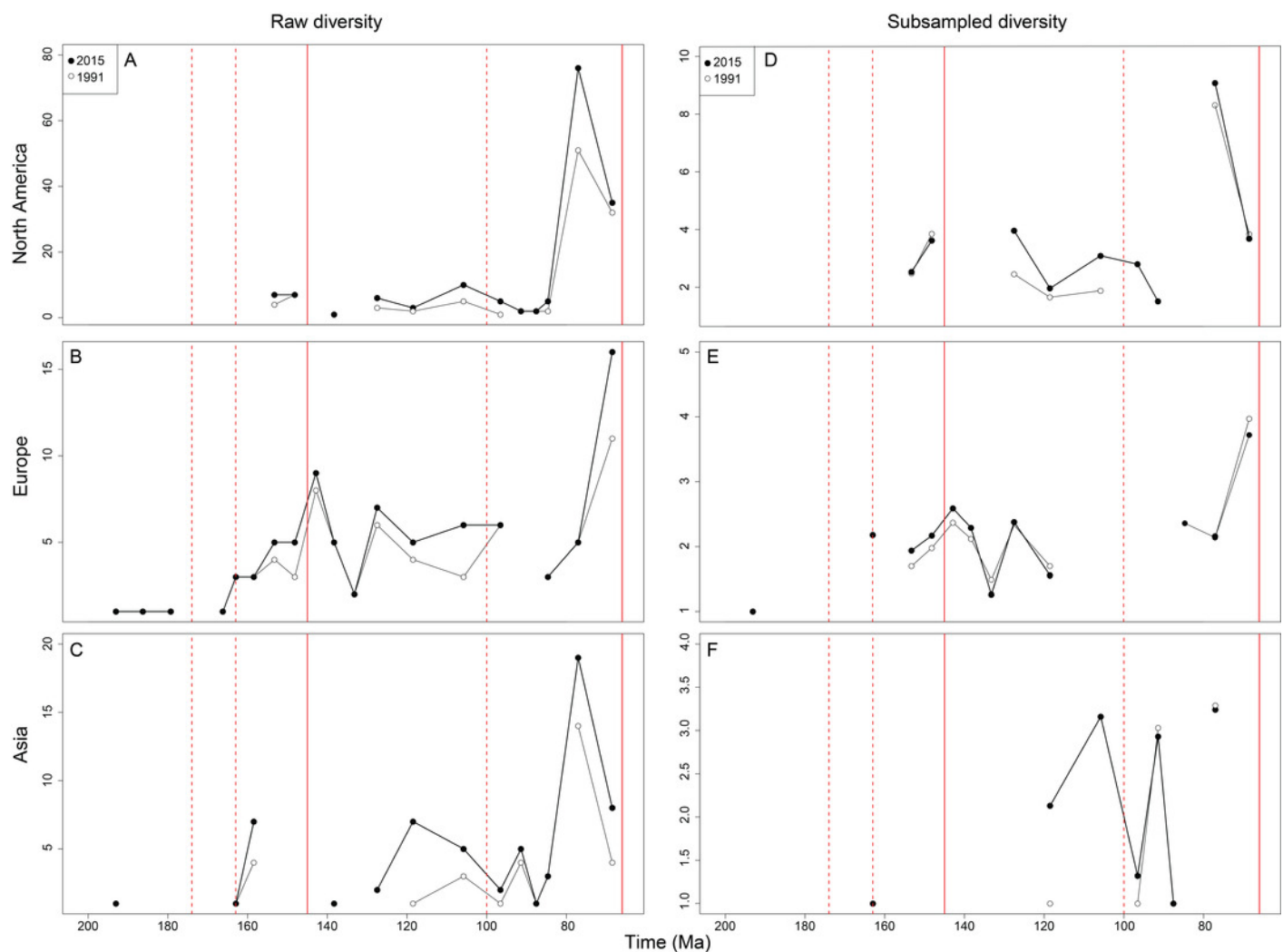
Figure 4

Subsampled theropod diversity

Figure 4: Raw and subsampled regional theropod diversity for North America (A, D), Europe (B, E), and Asia (C, F) based on our published knowledge in 1991 and 2015. ? �
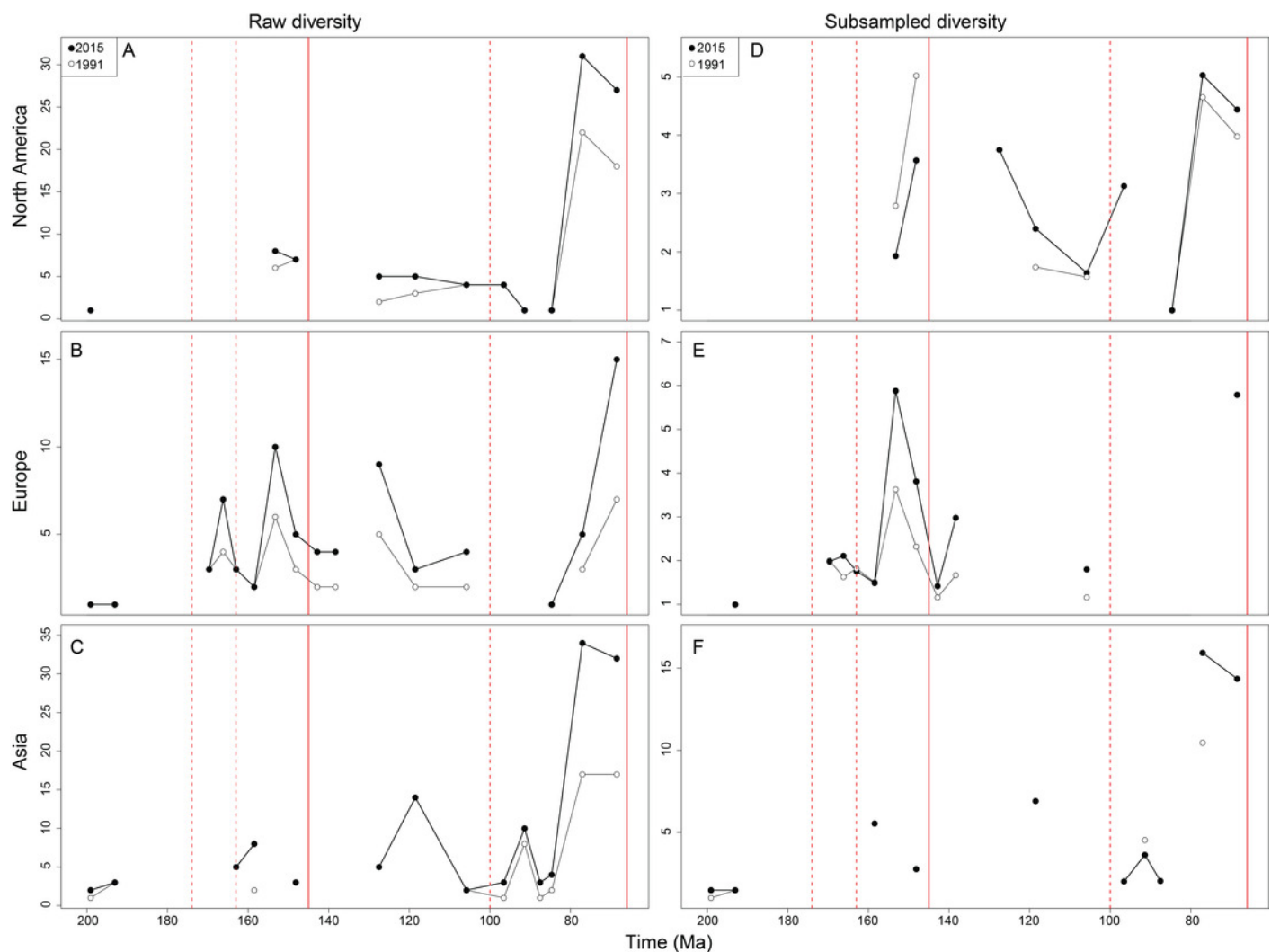
Figure 5

Regional sauropod diversity

Figure 5: Raw and subsampled regional sauropodomorph diversity for North America (A, D), Europe (B, E), and Asia (C, F) based on our published knowledge in 1991 and 2015.