UNIVERSITY OF
BIRMINGHAM

Research at Birmingham

Preservational bias controls the fossil record of pterosaurs

Dean, Christopher D.; Mannion, Philip D.; Butler, Richard

DOI:
10.1111/pala.12225

License:
Other (please specify with Rights Statement)

Document Version
Peer reviewed version

Citation for published version (Harvard):
Dean, CD, Mannion, PD & Butler, R 2016, 'Preservational bias controls the fossil record of pterosaurs', Palaeontology, vol. 59, no. 2, pp. 225-247. https://doi.org/10.1111/pala.12225

Link to publication on Research at Birmingham portal

Publisher Rights Statement:
This is the peer reviewed version of the following article: Preservational bias controls the fossil record of pterosaurs, which has been published in final form at 10.1111/pala.12225. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 01. Feb. 2019

Preservational bias controls the fossil record of pterosaurs

by CHRISTOPHER D. DEAN[1], PHILIP D. MANNION[1] & RICHARD J. BUTLER[2]

[1]*Department of Earth Science and Engineering, Imperial College London, London, United Kingdom; e-mail: christopher.dean09@imperial.ac.uk; philipdmannion@gmail.com*

[2]*School of Geography, Earth and Environmental Sciences, University of Birmingham, Edgbaston, Birmingham, United Kingdom; email: r.butler.1@bham.ac.uk*

Abstract: Pterosaurs, a Mesozoic group of flying archosaurs, have become a focal point for debates pertaining to the impact of sampling biases on our reading of the fossil record, as well as the utility of sampling proxies in palaeodiversity reconstructions. The completeness of the pterosaur fossil specimens themselves potentially provides additional information that is not captured in existing sampling proxies, and might shed new light on the group's evolutionary history. Here we assess the quality of the pterosaur fossil record via a character completeness metric based on the number of phylogenetic characters that can be scored for all known skeletons of 172 valid species, with averaged completeness values calculated for each geological stage. The fossil record of pterosaurs is observed to be strongly influenced by the occurrence and distribution of Lagerstätten. Peaks in completeness correlate with Lagerstätten deposits, and a recovered correlation between completeness and observed diversity is rendered non-significant when Lagerstätten species are excluded. Intervals previously regarded as potential extinction events are shown to lack Lagerstätten and exhibit low completeness values: as such, the apparent low diversity in these intervals might be at least partly the result of poor fossil record quality. A positive correlation between temporal patterns in completeness of Cretaceous pterosaurs and birds further demonstrates the prominent role that Lagerstätten deposits have on the preservation of smaller bodied organisms, contrasting with a lack of correlation with the completeness of large-bodied sauropodomorphs. However, we unexpectedly find a strong correlation between sauropodomorph and pterosaur completeness within the Triassic–Jurassic, but not the Cretaceous, potentially relating to a shared shift in environmental preference and thus preservation style through time. This study highlights the importance of understanding the relationship between various taphonomic controls when correcting for sampling bias, and provides additional evidence for the prominent role of sampling on observed patterns in pterosaur macroevolution.

Key words: Fossil record completeness; Lagerstätten, Mesozoic; Pterodactyloidea; Pterosauria; Sampling bias

PTEROSAURS were a group of Mesozoic flying archosaurs that went extinct at the Cretaceous/Paleogene (K/Pg) mass extinction 66 Ma, approximately 150 myr after their first appearance in the fossil record (Wellnhofer 1991; Unwin 2003, 2005; Barrett *et al.* 2008; Dalla Vecchia 2013; Witton 2013). They show high taxonomic (Unwin 2005; Butler *et al.* 2009, 2013; Witton 2013) and morphological diversity (Prentice *et al.* 2011; Foth *et al.* 2012; Butler *et al.* 2012), with a global distribution by at least the Early Jurassic (Barrett *et al.* 2008; Upchurch *et al.* 2015). A number of studies have evaluated evidence for fluctuations through time in pterosaur diversity, including discussion of their possible competitive replacement by birds (e.g. Unwin 1987, 2005; Buffetaut *et al.* 1996; Slack *et al.* 2006; Butler *et al.* 2009, 2012, 2013; Dyke *et al.* 2009; Benton *et al.* 2011; Andres 2012; Benson *et al.* 2014).

Whereas pterosaur remains have been found in hundreds of localities, and on all continents throughout the Mesozoic (Barrett *et al.* 2008; Upchurch *et al.* 2015), their often-fragmentary state means that many specimens do not yield sufficient morphological data to determine their taxonomic status. Pterosaur skeletons were adapted for flight and, as such, are light, thin-walled and highly fragile; preservation of complete specimens within high-energy depositional settings is rare (Wellnhofer 1991). Consequently, our knowledge of pterosaur anatomy and species richness is dominated by Lagerstätten (Buffetaut 1995; Butler *et al.* 2009), formations with unusually good fossil preservation (Seilacher 1970), of which several pterosaur-bearing units are known, including the Late Jurassic Solnhofen Limestone of Bavaria in Germany (e.g. Wellnhofer 1970, 1975), and the Early Cretaceous Jehol Group of China (e.g. Wang & Zhou 2006). Approximately a dozen of these formations account for ~50% of the total known species of pterosaurs (Benton *et al*. 2011). Previous palaeodiversity studies have shown that these Lagerstätten deposits can have a large influence on observed diversity for time periods in which they appear (Benson and Butler 2011), potentially distorting our understanding of patterns of diversity change through time. As such, their predominance as our main window onto pterosaur evolution might mask true diversity patterns throughout the Mesozoic: i.e. highly heterogeneous sampling might produce episodic peaks of observed diversity during periods containing Lagerstätten, whereas time intervals dominated by fragmentary or less diagnosable material might produce troughs in recorded diversity which do not necessarily reflect real diversity changes (Wellnhofer *et al*. 1991; Buffetaut 1995; Butler *et al*. 2009, 2013; Benton *et al*. 2011, 2013; Prentice *et al*. 2011; but see Dyke *et al*. 2009).

It has become increasingly apparent that biases in our sampling of the fossil record, resulting from geological, taphonomic and anthropogenic processes, might distort observed patterns of diversity (e.g. Raup 1976; Alroy *et al*. 2001; Smith 2001; Peters and Foote 2001; Peters 2005; Smith and McGowan 2005). Ameliorating for such biases is crucial to enable exploration of diversity through deep time. Sampling proxies, representations of bias introduced to the fossil record through anthropogenic or geological processes, have become a common, albeit controversial, method used to 'correct' for biased signals recorded in the geological record. Butler *et al.* (2009) presented a detailed examination of pterosaur species and genus level diversity through time, and attempted to account for sampling bias by comparing diversity with the numbers of geological formations preserving pterosaur remains (pterosaur-bearing formations: PBFs). A strong correlation was found between temporal variation in the number of PBFs and diversity over time for both taxonomic and phylogenetic diversity estimates (TDEs and PDEs respectively). Butler *et al.* (2009) suggested that these correlations could potentially be explained by one of two hypotheses: that either a significant portion of the observed diversity curve for pterosaurs is controlled by sampling variation, or that diversity is accurate, with diversity and sedimentary rock preservation covarying in response to a third driving factor (e.g. sea level). Butler *et al.* (2009) favoured the sampling bias hypothesis, but voiced caution with regard to their attempts to 'correct' diversity patterns using PBFs, because this approach did not account for the presence or absence of Lagerstätten.

However, the use of formations as a sampling proxy (for both raw numbers of formations and counting only fossiliferous formations) has been criticized by several authors (e.g. Crampton *et al*. 2003; Benton *et al*. 2011, 2013; Dunhill *et al*. 2012, 2013, 2014). Benton *et al.* (2011) argued that: (1) in many cases formation counts will always correlate with fossil counts, irrespective of the degree of sampling, rendering PBFs redundant with pterosaur diversity; and (2) absent results (either potential future sites of pterosaur recovery or sites which failed to produce pterosaurs but still represent a sampling attempt) are ignored: if sampling in these formations in insufficient, this can create an additional bias (see also Upchurch *et al*. 2011 for further debate). Butler *et al.* (2013) revisited the analyses of Butler *et al.* (2009), using new sampling proxies of pterosaur-bearing collections (PBCs), as well as dinosaur-bearing formations and collections (DBFs and DBCs respectively), the latter two representing attempts to quantify 'global' sampling and utilize sampling metrics independent of the pterosaur fossil record. DBFs and PBFs were found to correlate with one another at least within the Cretaceous, suggesting that PBFs at least partially reflect a 'global' sampling proxy (Butler *et al*. 2013). Butler *et al.* (2013) concluded that whereas 'true' pterosaur diversity might be impossible to ever fully reconstruct, current formation-based proxies enable an incomplete, but informative picture of broad scale patterns of species richness throughout the Mesozoic.

The completeness of the fossil material of a taxonomic group is likely to exert a direct influence on the observed species richness of that group through time, and thus represents an additional metric that might capture otherwise neglected aspects of sampling bias (Mannion and Upchurch 2010a; Brocklehurst *et al*. 2012; Walther and Fröbisch 2013; Brocklehurst and Fröbisch 2014; Cleary *et al*. 2015). Time bins with low average completeness values of specimens will potentially yield less taxonomically diagnosable specimens. Observed diversity could therefore be: (1) reduced, as collected remains cannot confidently be assigned to a species; or (2) increased, as a result of assigning partial remains of one taxon to multiple species (Mannion and Upchurch 2010a). Alternatively, periods of high completeness levels (including times in which Lagerstätten are present) should allow for clear recognition and identification of species, resulting in heightened observed diversity.

Only one previous study has attempted to examine the completeness of the pterosaur fossil record. Dyke *et al*. (2009) compiled a dataset of 66 genera, comprising 101 species, to calculate completeness over geological stage-level time bins, using three-categories of preservation, based on: (1) one bone; (2) more than one bone; and (3) more than one individual, each known from more than one bone. Although this approach utilized semi-quantitative values, these appear to be assigned as arbitrary metrics; the application of alternative preservation categories might have a notable impact on the resultant data, irrespective of actual fossil record quality or methodology. The low number of categories assigned (three unevenly assigned metrics relating to potential states of completeness) is coarse-grained and it is possible that underlying trends and subtle variations in fossil record quality might be obscured. In light of these potential weaknesses, as well as the recent spate of newly described pterosaur taxa (e.g. Manzig *et al.* 2014; Wang *et al.*, 2014a,b), a re-evaluation of the completeness of the pterosaur fossil record is timely.

Here, an extensive new dataset of pterosaur completeness is presented, utilizing a recently devised quantitative metric (Mannion and Upchurch 2010a), and comprising completeness values for 172 valid species of pterosaurs. The impact of completeness on our reading of the pterosaur fossil record is evaluated via statistical comparisons with various sampling proxies and sea level, as well as a 'corrected' diversity estimate. Pterosaur completeness is also compared to comparable datasets for Mesozoic birds and sauropodomorph dinosaurs in order to assess how completeness varies between different body plans and sizes through time. In addition, we make comparisons between our results and those of Dyke *et al.* (2009), to test for differences between these two approaches to assessing completeness. Four main aims form the focus of this study: (1) to determine the potential impact of pterosaur completeness on 'key' intervals of pterosaur history; (2) to determine whether completeness acts as a controlling mechanism on observed pterosaur diversity; (3) to assess the impact of Lagerstätten on the pterosaur fossil record through time; and (4) to compare the fossil record of small bodied (i.e. pterosaurs and birds) and large-bodied (sauropodomorphs) organisms to test for differences in preservational bias.

MATERIALS AND METHODS

Dataset

Data on all taxonomically diagnosable species of pterosaurs were compiled from Andres (2010), Butler *et al.* (2013), *The Paleobiology Database* (PaleoDB: www.paleobiodb.org), and a comprehensive review of the literature. Collated information was thoroughly scrutinized for potential synonyms and *nomina dubia.* The finalized dataset contains 135 genera, comprising 172 species described between 1812 and 1st January 2014, along with stratigraphic ranges, completeness data, and environment of deposition (see Appendix S1).

Although this study presents the most complete compilation of pterosaur taxic diversity to date, only minor changes are noticeable between this and the dataset of Butler *et al.* (2013). As such, our focus is on comparisons between diversity and fossil record quality, rather than a re-evaluation of aspects of pterosaur diversity, and results relating only to diversity or residual diversity estimates are not discussed.

Completeness metrics

The Character Completeness Metric (CCM) of Mannion and Upchurch (2010a) quantifies the potential phylogenetic data preserved in specimens; a percentage score is provided for the number of characters that can be coded for a specimen/taxon for phylogenetic analysis. This allows for adequate representation of individual skeletal elements that contain considerably more phylogenetically important information than others. CCM2 (*sensu* Mannion and Upchurch 2010a), combining the information of all specimens of a taxon to attain overall completeness, has been deemed more appropriate for comparisons between fossil record quality and diversity (Brocklehurst *et al.* 2012), and has thus been employed here. It is herein referred to just as the CCM. The phylogenetic data matrix presented in Butler *et al.* (2012 [based on that of Andres, 2010]) which samples 101 pterosaur species, scored for 183 characters (cranial = 107 characters, 59%; axial = 19 characters, 10%; appendicular = 57 characters, 31%) is utilized using the same methodology as Walther and Fröbisch (2013), whereby the number of scorable characters was counted (See Supporting Information). The completeness of the remaining 71 taxa was determined via the literature, using this same character list. Completeness of sauropodomorphs and Mesozoic birds was extracted from Mannion and Upchurch (2010a) and Brocklehurst *et al.* (2012).

Sampling proxies and sea level

Four time series are utilised as proxies to account for sampling biases within the Mesozoic rock record: PBFs, PBCs, DBFs, and DBCs (see Table 1). Formations represent a prospective site of fossil extraction and, although potentially influenced by external factors (e.g. human naming and collection bias, geographical bias, extent of outcrop, ability to preserve fossils adequate for taxonomic assignment), and despite criticism (Dunhill 2012; Benton *et al*. 2011, 2013), have been argued to provide an appropriate proxy for estimating temporal variation in the amount of rock available for sampling (Peters 2005; Upchurch *et al*. 2011; Butler *et al*. 2013). Collections incorporate all specimens collected from a stratigraphic horizon tied to a single location. DBFs and DBCs are proxies that capture a more global estimate of terrestrial sampling: dinosaurs were abundant and geographically diverse throughout the Mesozoic, living in nearly all terrestrial and coastal environments. The use of formations and collections in which pterosaur fossils could, but have not yet, been recovered, reduces the possibility of 'redundancy' with pterosaur diversity estimates (Benton *et al*. 2011; Butler *et al*. 2013). Two further proxies were utilized for appropriate comparisons with sub-divided environmental data series (see below): terrestrial and marine tetrapod-bearing collections (TTBCs and MTBCs, respectively).

All data for PBFs, PBCs, DBFs and DBCs were obtained from Butler *et al.* (2013), originally sourced from PaleoDB downloads, augmented with data from new collections/formations reported since this publication. Numbers of TTBCs and MTBCs were downloaded from the PaleoDB on the 8th June 2013. The Mesozoic sea level curve was sourced from data in Butler *et al*. (2011), based on Haq *et al*. (1987) and Miller *et al*. (2005).

Time Bins

Pterosaur fossil record quality was assessed through calculating a mean average value of CCM scores and associated standard deviations for each Mesozoic stage-level time bin. Stage-level bins were chosen to facilitate comparisons with the diversity and sampling proxy data in Butler *et al.* (2013), the sea level data in Butler *et al*. (2011), and the completeness data in Dyke *et al.* (2009). This approach resulted in a discrepancy when making comparisons to the sub-stage level data for sauropodomorph completeness of Mannion and Upchurch (2010a). To enable comparisons between datasets, sauropodomorph data were averaged to produce artificial stage level bins, and resulting stages were assessed for any statistically significant outlying results. Most bins showed minimal or no change; the average alteration to completeness values across all time bins produced by this method is 1.39%. Original and averaged values are presented in Appendix S1. Comparisons between pterosaur and bird CCM were carried out for stages after the first appearance of birds (Tithonian onwards). Species appearing within multiple time bins were considered independently and included in the total count for each stage where they are present**.** All further time series data were assigned to stage-level time bins.

No valid pterosaur taxa are known from the Aalenian (early Middle Jurassic). Although this time interval might represent a genuine drop in diversity, pterosaurs clearly did not become entirely extinct, and inclusion of this zero-data time bin has the potential to artefactually influence the strength and significance of correlative tests. As such, we ran our analyses both including and excluding the Aalenian.

Subdivisions of pterosaurs

To examine additional parameters which might influence or record otherwise neglected aspects of pterosaur completeness through time, we subdivided our pterosaur dataset according to Lagerstätten, environment, and taxonomy, and then calculated additional time series of completeness and diversity for these subsets (see Appendix S1).

Lagerstätten. To assess the impact of Lagerstätten on both pterosaur completeness and diversity, we separated species into those that were recovered from Lagerstätten, and those that came from other formations. Although formations such as the Calcare di Zorzino of Italy (Dalla Vecchia 2003) could potentially be classified as Lagerstätten, our list remains identical to that of Butler *et al*. (2013) for ease of comparison, and can be found in Appendix S1.

Environments. Pterosaurs were additionally grouped according to whether they were deposited and preserved within marine/coastal or terrestrial settings. Marine environments include lagoonal, coastal and open marine deposits, whereas terrestrial environments included lacustrine, fluvial, and floodplain deposits (for a full list of environments see Appendix S1; Environments). For ease, we subsequently refer to pterosaurs as either terrestrial or marine. All data were sourced from the PaleoDB and the relevant literature.

Taxonomic groups. Pterosaurs were taxonomically divided into Pterodactyloidea and non-Pterodactyloidea (the paraphyletic assemblage of early-branching pterosaur taxa that have historically been referred to as 'rhamphorhynchoids'). The aim of this was to examine events around the Jurassic/Cretaceous (J/K) boundary, which has been suggested to mark a major extinction of non-pterodactyloids (e.g. Benson *et al*. 2014), as well as to consider any potential impact of palaeobiological changes on completeness metrics.

Pairwise statistical tests & Residual Diversity Estimates (RDE)

Jarque-Bera and Shapiro-Wilk tests were implemented to test for normality in all time series. As pterosaur CCM scores were shown to not be normally distributed, statistical comparisons between time series were therefore calculated using non-parametric tests (Spearman's rank correlation coefficient and Kendall's tau rank correlation coefficient). In addition to calculating raw taxic diversity, the residuals method of Smith and McGowan (2007) was used to produce species diversity curves that are 'corrected' for sampling bias. Using this approach, residual diversity estimates (RDEs) were calculated using all sampling proxies. Data were detrended, and generalized differencing (McKinney 1990) was applied where autocorrelation was detected. These analyses were implemented in Excel and PAST (Hammer *et al.* 2001). Abbreviations for common terms used in this paper can be found within Table 1.

Multiple Regression Modelling

For multivariate statistical approaches, stage duration was included as a non-optional predictor in models to account for stage length variation, and Lagerstätten were coded as present or absent using a binary variable in an attempt to account for their presence without discarding data (see Benson and Butler 2011). Both linear modelling (multiple regression models) and generalized least squares (GLS) models were applied to explore the possibility of multiple explanatory variables, which can otherwise be missed through pairwise comparisons. Multiple regression models allow for the sequential addition and removal of time series to seek the model that best explains the chosen metric, with autoregressive terms included to correct for serial correlation in time series data (Chatfield 2003). The order of the autoregressive term for each model was selected by comparing AICc values at autoregressive orders between zero and three.

GLS models take account of autocorrelation and permit specification of a non-standard variance-covariance matrix, allowing non-independence of points within data series according to specific models. The first-order autoregressive correlation model was used in this case, which seeks autocorrelation at up to one lag in either direction, and minimizes the error term (Box *et al.* 1994). Combinations of explanatory variables were compared using

AICc weights, as small sample sizes potentially reduce the efficiency of AIC for assessing the quality of model fitting (Burnham and Anderson 2001). R-squared, F-value or p-values were not computed for GLS models as their benefits are currently debated (e.g. Long and Freese 2006). All multivariate analyses were implemented in R version 3.0.2 (R Development Core Team 2015), using the packages lmtest version 0.9-22 (Zeilis and Hothorn 2002), nlme version 3.1-117 (Pinheiro *et al.* 2009), qPCR version 1.3-9 (Spiess and Ritz 2010), and tseries version 0.10-32 (Trapletti and Hornik 2009), and data were log10 transformed prior to analysis (only stage duration and Lagerstätten presence/absence were not transformed). The Jarque–Bera test indicated that the residuals from these regression analyses were all normally distributed.

RESULTS

Pterosaur completeness through time

Completeness scores and taxic diversity for each stage are plotted against time (Fig. 1). Initial completeness within the Late Triassic is high (~50%), followed by fluctuations through the Jurassic. After a Hettangian peak (84%), completeness declines to a Pliensbachian low (13%), prior to a peak in the Toarcian (62%). The remainder of the Jurassic exhibits varying levels of completeness, with a trough in the Bathonian (26%), and peaks in the Oxfordian (58%) and Tithonian (57%). Average completeness is generally lower in the Cretaceous than in preceding time intervals. Values fall dramatically across the J/K boundary (Berriasian = 15%), before a gradual rise to an Aptian peak (43%), and subsequent decrease in the Cenomanian (11%). A second Cretaceous peak is reached in the Coniacian–Santonian (~40%), before a decline in the Campanian (25%) that is continued into the Maastrichtian (14%).

Sampling proxies through time

PBFs, PBCs, DBFs and DBCs were plotted against time and compared with the pterosaur completeness curve (See Fig. 2A-C). PBF counts are low in the Triassic and Jurassic, averaging eight formations per stage, although a slight increase is observed towards the end of the Jurassic. PBCs exhibit strong, irregular peaks within the Toarcian and Kimmeridgian. Trends in PBCs and PBFs within the Cretaceous show a similar pattern to changes in completeness, mirroring the two broad peaks observed in completeness data; it should however be noted that these peaks appear to lag behind completeness values and, as such, correspond with intervals of lower completeness. This is also observable in the Late Jurassic, where troughs in completeness correlate with apparent peaks in PBF and PBC counts.

Whereas counts differ vastly between the two proxies, DBFs and DBCs exhibit a similar shape, and thus shall be discussed together. There is little clear similarity between pterosaur completeness and DBF/DBC curves throughout the Triassic or Jurassic; whereas completeness scores fluctuate, sampling proxy curves show a trend of gradual increase towards the Cretaceous. DBF/DBC counts seem to follow a similar shape to that of completeness within the Early and middle Cretaceous, although this appears as a one stage 'lag' behind completeness. A large discrepancy is observed in the Campanian and Maastrichtian between high values of DBFs/DBCs and steadily declining pterosaur completeness.

Correlations between completeness, diversity and sampling

All formation- and collection-based sampling proxies exhibit strong, positive and statistically significant correlations with species richness over geological time (see Table 2). Pterosaur completeness has a statistically significant positive correlation to both uncorrected diversity and residual diversity (calculated with PBFs, DBFs and DBCs) throughout the Mesozoic (see

Table 3) in all tests. The PBF-based residual diversity estimate produced the strongest correlation with pterosaur completeness, followed by DBF-based residuals, raw taxic diversity, and finally DBC-based residuals. Within the Triassic–Jurassic, a statistically significant, positive correlation is recovered between completeness and PBF-based residuals for all tests, and with taxic diversity and DBF-based residuals for the Kendall's tau test. Completeness shows a strong, positive correlation with raw taxic diversity, PBFs and DBFs in the Cretaceous, as well as with DBCs when using Spearman's Rank. PBCs are not shown to have any statistically significant correlations with completeness.

Comparisons with completeness of Sauropodomorpha and Aves

Completeness values for sauropodomorph dinosaurs and Mesozoic birds were plotted against time to facilitate comparisons (see Fig. 3). Whereas both pterosaur and sauropodomorph completeness show a general decrease throughout the Mesozoic, comparisons between the two curves reveal little in the way of similarities on a stage to stage basis. The avian completeness curve bears a striking similarity to that of Pterosauria throughout the Tithonian–Cretaceous; only one discrepancy exists, when the bird fossil record quality dips compared to that of pterosaurs in the Albian (6% and 30% average completeness respectively).

Statistical comparisons between average completeness scores for pterosaurs, birds and sauropodomorph dinosaurs are shown in Table 4. Pterosaurs exhibit higher average completeness than avian species within stages where both are present, but lower than sauropodomorphs. When split between the Triassic–Jurassic and Cretaceous, pterosaurs display the lowest completeness values within both time intervals (48% and 28% respectively; it should be noted that the first appearance of birds in the Tithonian does not allow for comparison of Triassic–Jurassic completeness between pterosaurs and birds). Pterosaur and avian completeness show a strong, positive and statistically significant correlation within the Cretaceous and across all time bins in which they are present for all statistical tests. Sauropodomorph completeness scores show no significant correlation to those of pterosaurs across the Mesozoic and for the Cretaceous by itself; however, a statistically significant positive correlation is found in analyses restricted to the Triassic–Jurassic component of our sauropodomorph and pterosaur completeness values.

Sensitivity tests

To examine the effect that stages with no data might have on correlation tests, analyses were also run with the Aalenian excluded (Table 5). For the Mesozoic, the results of correlation tests were the same as for the complete dataset, apart from with DBC-based residuals, which did not produce a statistically significant result. When we restricted analyses to the Triassic–Jurassic, only sauropodomorph completeness shows a statistically significant, positive correlation with pterosaur completeness values, compared to positive correlations between pterosaur completeness and taxic diversity, PBF residuals and DBF residuals with the Aalenian included. Cretaceous comparisons were largely unchanged, with only DBC-based residuals failing to recover a statistically significant result.

Lagerstätten

Time-binned CCM values for species from Lagerstätten and species from non-Lagerstätten deposits were each plotted against time (Fig. 4A). These were compared to overall pterosaur completeness scores to visualize the occurrence of deposits of exceptional preservation on temporal variation in the completeness of pterosaur remains (Fig. 4A). The removal of Lagerstätten deposits decreased average completeness for each stage that they appear within. The largest decreases appear within the Toarcian (from 62% to 26%) and Santonian (from 40% to 8%), corresponding to the Posidonienschiefer and Niobrara Chalk formations. Further reductions are visible in the Oxfordian, Barremian, Aptian and Campanian. A very strong, highly statistically significant correlation occurs between Lagerstätten-only completeness and taxic diversity (Fig. 4B) for the Mesozoic (for all tests). This pattern is also observed when the Triassic–Jurassic and Cretaceous are analysed separately, although the latter is marginally weaker. No significant correlation is observed between Lagerstätten-excluded completeness and diversity for the Mesozoic or for the Triassic–Jurassic and Cretaceous.

Two data series were examined to check for the influence of Lagerstätten on pterosaur completeness through deep time: Lagerstätten-only (Fig. 4C) and Lagerstätten-excluded (Fig. 4D). For many stages, completeness values are zero for the Lagerstätten-only data series, due to their heterogeneous appearance in the fossil record. However, when Mesozoic completeness is averaged only from stages containing fossil remains, Lagerstätten-only deposits show a much higher average completeness (59%). Additionally, differences in Lagerstätten completeness between periods can be assessed: for Lagerstätten-only, within the Triassic–Jurassic, an extremely high CCM (71%) is observed. When looking at the Cretaceous, Lagerstätten-only completeness is much lower (40%). With Lagerstätten excluded from the completeness scores, no significant correlations between completeness and other variables were observed for the Mesozoic as a whole, within all tests (Table 6). For the Triassic–Jurassic, a positive, statistically significant result was recorded when compared to sauropodomorph completeness scores; no significant correlations were observed for all other variables in both the Triassic–Jurassic and the Cretaceous.

Environmental variation

The completeness of pterosaurs from marine deposits closely follows overall pterosaur completeness from the Triassic until the Bathonian, at which point much lower completeness values are observed for marine pterosaurs (Fig. 5A). A sharp increase in marine pterosaur completeness occurs from the Callovian, reaching a peak in the Tithonian (63%), which corresponds with an increase in the taxic diversity of pterosaurs in marine deposits (Fig. 5B-C). This increase is followed by a decline in completeness across the J/K boundary, into the Berriasian. Marine and overall pterosaur completeness are similar for the rest of the Cretaceous, although marine values are generally lower; only within the Santonian and Campanian are there higher completeness values from marine specimens.

Terrestrial pterosaur completeness and overall completeness values are not similar for the Triassic and Early Jurassic, with terrestrial pterosaurs being less complete than the overall pattern (Fig. 5A). A peak in terrestrial completeness is observed from the Bajocian–Oxfordian, which matches closely with the overall pterosaur completeness curve; the end of this peak coincides with an increase in terrestrial taxic diversity. Following a decline within the Kimmeridgian and Tithonian, completeness values for terrestrial pterosaurs increase until the Albian. Whereas high terrestrial completeness values are observed in tandem with high terrestrial taxic diversity for the Barremian and Aptian (Fig. 5D), both the Hauterivian and Albian exhibit only high completeness, with low taxic diversity. Following a period of no recorded terrestrial specimens within the Turonian and Coniacian, completeness increases towards the K/Pg boundary, the opposite of that seen for the total pterosaur completeness curve.

Pterosaurs found within marine deposits are on average slightly more complete (31%) than those from terrestrial deposits (25%), both as an average across all stages and when only counting stages with pterosaurs recovered. This pattern is only partially supported when the Mesozoic is subdivided: marine completeness is much higher than that of the terrestrial realm in the Triassic–Jurassic (50% compared to 34%), but slightly lower within the Cretaceous (22% compared to 27%).

Marine completeness is statistically significantly, positively correlated with both marine taxic diversity and marine TBCs for the Mesozoic in all tests (Table 7). This same pattern is also observed within the Triassic–Jurassic, in addition to a statistically significant, positive correlation with sauropodomorph completeness. No significant correlations appear within the Cretaceous. Terrestrial completeness exhibits a strong, positive, statistically significant correlation with terrestrial taxic diversity for the Mesozoic in all tests; this is also observed within the Triassic–Jurassic. Once again, no significant correlation appears within the Cretaceous with any variable.

Taxonomic groups

The average completeness of non-pterodactyloid pterosaurs is much higher than that of Pterodactyloidea (45% versus 27%). Occurrences of non-Pterodactyloidea and Pterodactyloidea overlap only in the Late Jurassic (Figure 6); consequently, direct comparison of completeness values between the two groups is difficult. The first appearance of a possible pterodactyloid in the Oxfordian (Lü and Fucha, 2010) exhibits low completeness (*Archaeoistiodactylus linglongtaensis* [4.9%]; however, see Martill and Etches [2013] and Sullivan *et al*. [2014] regarding the affinities of this taxon). Subsequently, as pterodactyloid diversity increases within the following stages, their completeness values rise to a peak of 56% in the Tithonian, immediately prior to the J/K boundary. Non-pterodactyloid pterosaurs exhibit high completeness values throughout the Late Jurassic, although no taxa survive across the J/K boundary (Unwin 2003; Butler *et al*. 2013; Andres *et al*. 2014).

Positive and statistically significant correlations exist between pterodactyloid completeness scores and both pterodactyloid taxic diversity and bird completeness scores (Table 8). Non-Pterodactyloidea completeness is statistically significantly positively correlated with non-Pterodactyloidea taxic diversity (Kendall's Tau test only) and with sauropodomorph completeness.

Sea Level

There are no clear, shared patterns between pterosaur completeness and sea level changes: sea level shows an overall rising trend towards the K/Pg boundary, whereas the completeness of pterosaur specimens fluctuates widely (Figure 7). The correlation between sea level and pterosaur completeness through time is weak and statistically insignificant for all comparisons (See Tables 3, 5, 6, 7 & 8). Plots were also generated for groupings of marine and terrestrial completeness and diversity against sea level (see Appendix S1, All Categories). None of these showed evidence for correlation.

Multiple Regression Models

Multiple regression model fitting procedures show that a combination of taxic diversity and PBFs provides the best explanation for pterosaur completeness (Table 9). The relationship between completeness and diversity is strongly positive, but interestingly the relationship with formations is negative, suggesting that lower numbers of formations result in higher average completeness. However, it should be noted that the *p*-value for this model is above the 0.05 threshold, and thus the null model cannot be rejected.

Ranked by AICc value, the best fitting GLS model shows that a combination of DBFs and Lagerstätten presence/absence is the best predictor of pterosaur completeness (Table 10; for the full model, see Appendix S1: GLS Outputs). Out of the top five models, the time series that appear are DBFs, PBFs, PBCs and Lagerstätten presence/absence, with the latter appearing in four out of five. In the top twenty models, Lagerstätten presence/absence and PBFs are the most commonly observed variables. The bottom five models all contain many time series, and no discernable pattern is observed. However, in the bottom twenty models, stage length, PBCs and taxic diversity are the most commonly observed data (See Appendix S1).

Diversity: Additional GLS models were generated to test for the combination of variables which best explains pterosaur taxic diversity (Table 11; for the full model, see Appendix S1: GLS Outputs). By AICc value, the top GLS model shows a combination of Lagerstätten presence/absence and PBFs as the best explanation for pterosaur diversity, with the former also appearing in every one of the top twenty models.

DISCUSSION

Lagerstätten and comparisons with Dyke et al. (2009)

The strong heterogeneity of the sampling of the pterosaur fossil record likely introduces extensive sampling bias to observed diversity patterns. How Lagerstätten deposits influence patterns of completeness or observed diversity through time is poorly understood. A prior investigation by Dyke *et al.* (2009) concluded that the numbers of pterosaur species were not skewed towards specific stratigraphic intervals, although completeness of specimens through time was influenced by Lagerstätten distribution.

Multiple lines of evidence collected in this study indicate that Lagerstätten have a strong impact on the fossil record of pterosaurs. Results from the GLS models strongly support this hypothesis and imply a joint influence on diversity and completeness from Lagerstätten and formation numbers. No significant correlation is found between completeness and PBFs, PBCs, DBFs or DBCs; whilst this indicates that there is no relationship between these proxies and completeness, it does not contradict the existence of a pervasive Lagerstätten influence. A single Lagerstätte will represent one formation with numerous species and high completeness values. This formation might occur in a time interval in which other PBFs are scarce. The apparent discrepancy between low numbers of formations occurring at the same time as high apparent diversity and completeness results in absence of correlation between metrics. The same is true for our results from multiple regression modelling, specifically the negative relationship recovered between completeness and PBFs, suggesting fewer sampled formations leads to higher completeness values (although these results should be treated with caution due to lack of statistical significance).

Additionally, when looking at visual evidence of the influence of Lagerstätten, peaks in both average completeness and taxic diversity seem to correlate closely, or at least be influenced by the presence or absence of Lagerstätten. This is especially evident when looking at falls in diversity and completeness across stage boundaries where Lagerstätten are absent. These lines of evidence suggest that Lagerstätten are a dominating influence on completeness throughout the pterosaur fossil record, and confirm the hypothesis that Lagerstätten effects on observed diversity are severe for small and/or delicate organisms (Brocklehurst *et al.* 2012).

Within the last decade, there has been a marked increase in attempts to understand and address the potential impacts that uneven sampling of the fossil record might have on species richness through time (e.g. Alroy *et al*. 2001; Peters and Foote 2001; Smith 2001; Smith and McGowan 2007; Benson *et al.* 2010; Mannion *et al.* 2011). Correlation between residual diversity estimates and completeness recovered here could indicate that evolutionary factors related to increased diversity and abundance would have had a direct impact on the probability of recovering complete specimens. This study recovered a statistically significant positive correlation between pterosaur completeness and uncorrected raw diversity, as well as with residual diversity estimates produced using PBFs, DBFs and DBCs. As PBF-based residuals showed the strongest and most statistically significant result, it is possible that observed pterosaur completeness might be driven at least in part by genuine variation in species richness throughout time. Time periods during which pterosaur diversity was high would likely result in a wider geographical distribution and higher abundance, and as such an increased probability of high quality specimen preservation. Correlation with uncorrected diversity indicates an underlying influence from sampling bias within the fossil record; however, this appears to exert less control than the residual diversity estimate, implying that ’true‘ diversity (i.e. ‘sampling-corrected’ diversity) plays a more dominant role on completeness.

Multivariate analyses produce partly contrasting results. Diversity did not appear in the top five models when completeness was the independent variable, and vice versa. This disparity with the results from pairwise comparisons to residual diversity estimates is likely due to the fact that residual ‘corrected’ diversity does not include a correction for the presence or absence of Lagerstätten; this is further exacerbated by the increased influence of Lagerstätten-type deposits on small organisms. As such, it appears that diversity and completeness are unlikely to exert much influence upon one another, and the correlation recorded in pairwise comparison tests is a secondary effect of either Lagerstätten presence/absence and/or PBFs driving both in tandem.

Dyke *et al.* (2009) suggested two substantial areas of bias within the pterosaur fossil record: high completeness within the Tithonian (latest Jurassic) and fragmentary fossil material during the middle Cretaceous (Aptian–Cenomanian). Although comparison with the completeness curve presented here yields remarkably similar Tithonian and middle Cretaceous peaks and troughs, considerably more variation in fossil record quality is observed within the current study, with numerous fragmentary remains within the Jurassic (Pliensbachian and Callovian) and during the latest Cretaceous (Campanian–Maastrichtian).

Little further inference on patterns can be made, as significant changes in taxonomy have made comparison between Dyke *et al* (2009)'s results and ours problematic. Furthermore, the disparity between the opposing methods of using bone counts and character counts casts doubt on the few similarities that can be observed. As such, a better comparison with this study is likely to be produced when using the SCM method of measuring completeness, where a completeness score is based on the physical quantity and bulk of elements preserved, gauging an estimate as to how much of a complete skeleton is represented (see Mannion and Upchurch 2010a). However, despite this disparity between methods, our results challenge the idea that "the pterosaur fossil record is unbiased by sites of exceptional preservation (lagerstätte)" (Dyke *et al.* 2009, p890).

Potential Impact on Origins, Evolution and Extinctions

Completeness is high from the Norian to Hettangian (Late Triassic–earliest Jurassic), during the first appearance and radiation of pterosaurs (Andres 2012; Andres *et al*. 2014), potentially suggesting that specimen completeness is unlikely to be an important limiting factor in our understanding of taxonomic diversity at this time. Thus, the apparent rapid acquisition of high taxonomic diversity might be a genuine feature of the pterosaur fossil record, as implied by the interpretation of early pterosaur evolution as an adaptive radiation (Andres 2012). Whilst it could be argued that the Norian merely represents the first time bin with conditions suitable for pterosaur preservation, the first pterosaurs have all been found in marine formations, often in black shale lithologies, in Italy, Germany and Austria. Similar formations within the same geographical region and also preserving vertebrates are present within the Carnian (for example, the Polzberg Lagerstätte of the Reubgraben shales), yet yield no pterosaurs, providing a supportive taphonomic control. However, whereas currently recognized Norian diversity (eight species) approaches the average per stage for the Mesozoic (~9), species numbers throughout the Rhaetian–Pliensbachian are low, and fluctuations in completeness metrics might be driven by the relative preservation of individual, rarely discovered specimens. For example, the apparent peak in completeness during the Hettangian results from the discovery of a single species (*Dimorphodon macronyx* from the Lower Lias of England) within a marine formation that is well known for exceptionally preserved vertebrate specimens (Benton and Spencer 1995). The scarcity of pterosaur specimens from the Rhaetian–Pliensbachian, and the low completeness values of Sinemurian and Pliensbachian pterosaur species, makes it difficult to make inferences as to the impact of the Triassic/Jurassic mass extinction event on pterosaur evolution (see also Butler *et al*. 2013).

The J/K boundary has previously been noted as a period of potential extinction and faunal turnover for both marine and terrestrial groups (e.g. Raup & Sepkoski 1984; Upchurch *et al*. 2011; Benson and Druckenmiller 2014), including pterosaurs (e.g. Butler *et al*. 2009; Benson *et al*. 2014). Although the precise timing of their extinction is currently unknown, unambiguously dated remains of non-pterodactyloids are not observed after the J/K, and there is also an apparent decline in Pterodactyloidea species richness in the earliest Cretaceous. However, the impact of fluctuations in completeness levels through this interval has not previously been discussed in detail. There is a dramatic decline from 57% to 15% within average completeness, as well as a decline within Pterodactyloidea completeness from 54% to 15%, between the Tithonian and Berriasian. When split into preservational environments, pterosaurs appear to show varying completeness trends over the same interval: a decline is evident for marine completeness and diversity (with completeness falling from 63% to 11%), whereas terrestrial deposits show a slight increase in diversity and completeness (with completeness rising from 10% to 20%). One potential cause of this drop in completeness could be the lack of Lagerstätten reported in the Berriasian, in comparison to the Tithonian.

However, the same pattern of reduced completeness within the Berriasian is also observed when Lagerstätten are removed (showing a drop from 40% to 15%). Two explanations are possible: (1) the apparent low diversity of the earliest Cretaceous might be an artefact of a decreased preservation potential and the absence of Lagerstätten, perhaps resulting from global palaeoenvironmental change; or (2) the scrappy material by which currently known earliest Cretaceous pterosaurs are diagnosed could reflect a genuine scarcity of pterosaurs following a J/K extinction event. At present, these explanations cannot be readily distinguished and might not be mutually exclusive; however, given the results discussed above, it is likely that the first provides the more probable explanation.

A post-Albian reduction in pterosaur diversity has been equated with either a middle Cretaceous extinction event, or a long-term Late Cretaceous decline (Unwin 2005; Butler *et al.* 2009, 2013; Benson *et al*. 2014). A sudden decline in taxic diversity is observed during the Cenomanian, accompanied by troughs in completeness in the Cenomanian and Turonian. When split into environments, it is evident that the majority of this decline in completeness can be attributed to declines in the completeness of terrestrial pterosaurs. Lagerstätten deposits are also absent from this interval. Smith (2001) and Benson *et al.* (2013) have argued that a reduction in available fossiliferous rocks masks true diversity at this time. Although both a reduction in diversity and the loss of small-bodied pterosaurs clearly does occur through this time interval, low completeness values might make this decline appear more rapid and severe than was actually the case.

The record of pterosaurs during the Late Cretaceous has received particular interest due to the concurrent radiation of birds, with claims of a correlated Late Cretaceous decline within pterosaur diversity (Unwin 1987, 2005; Slack *et al.* 2006, Benson *et al.* 2014). A steady decline in pterosaur completeness is recovered from the Santonian onwards; taxic diversity first increases, then decreases through this interval. When split into environmental groupings, marine completeness is observed to dramatically drop during this time interval, whereas marine taxic diversity remains relatively level; terrestrial completeness increases towards the Maastrichtian, whereas corresponding terrestrial diversity shows a peak followed by a trough. However, a single species (*Quetzalcoatlus northropi*) produces an anomalously high value for pterosaur completeness within the Maastrichtian; when removed, completeness within the stage drops from 14% to 5%. Overall, these results indicate the poor quality of the fossil record used for assessing pterosaur diversity within the Maastrichtian, especially within the marine realm, and, as such, we urge caution when attempting to compare pterosaur taxic diversity in this stage with that of other vertebrate groups. In general, the importance of addressing environmental differences in taphonomy should be stressed when discussing the dynamics of extinction events within groups of low species and specimen counts.

Comparative Completeness

A statistically significant, strong positive correlation was recovered when comparing pterosaur and avian completeness within stage bins throughout the Mesozoic, but not with sauropodomorphs. This result is in agreement with the hypothesis proposed by Brocklehurst *et al.* (2012), whereby completeness is driven by differing taphonomical mechanisms between large, robust organisms and smaller fragile animals. Radically different body plans result in different methods of preservation and thus different completeness values. Both birds and pterosaurs exhibit similar flight-adapted bodies, with thin-walled and fragile bones; as such, similar modes of preservation within low energy depositional environments are common. Both are often found as part and counter-part split blocks in Lagerstätten deposits, yielding exceptionally preserved specimens; for example, Chinese Lagerstätten deposits, such as those in the Jehol Group, are famous for containing exquisitely preserved bird and pterosaur specimens (Wang *et al*. 2005; Wang & Zhou 2006). The control of Lagerstätten on the completeness of these groups can be clearly observed in the single discrepancy between bird and pterosaur completeness during the Albian, where birds are found to exhibit a much lower completeness. This can be explained by the Crato Formation, which acts as a pterosaur Lagerstätten, but for which bird specimens are only just starting to be discovered (Carvalho *et al*., 2015). As such, similar fossil record quality between pterosaur and avian species is to be expected.

In contrast, the large, robust bones of sauropodomorphs are likely to be preserved under different taphonomic settings. Despite the overall lack of correlation throughout the Mesozoic, and the expectation of a taphonomic difference between pterosaurs and sauropodomorphs, a positive correlation is present between pterosaur and sauropodomorph completeness for the Triassic–Jurassic. Similar results are apparent when sauropodomorph completeness is compared against non-Pterodactyloidea completeness, and the completeness of pterosaurs found in marine environments from those stages (and remain when the Aalenian is excluded). A potential cause for this surprising result might relate to environmental preferences. It has previously been suggested that sauropods displayed differing habitat preferences between non-titanosaurs and titanosaurs, with the former dominating during the Jurassic and the latter during the Cretaceous (Butler and Barrett 2008; Mannion and Upchurch 2010b). Although both were fully terrestrial groups, non-titanosaurs seem to have spent a greater amount of time in coastal environments than titanosaurs, which were more restricted to inland settings. Pterosaurs have also been inferred to have inhabited increasingly terrestrial environments through time (Butler *et al.* 2013); this appears to coincide with an increase in diversity and disparity towards the end of the Jurassic (Prentice *et al.* 2011; Butler *et al.* 2012; Foth *et al.* 2012) and the emergence of Pterodactyloidea (Andres *et al*. 2014). These results, in combination with an observed increase in terrestrial completeness values across the J/K boundary, suggest the possibility of an environmental-taphonomic shift, whereby sauropodomorphs and pterosaurs often frequented coastal environments during the Triassic and Jurassic, before showing increased terrestrialisation in the Cretaceous. This shared environment preference during the Triassic and Jurassic potentially led to similar patterns of fossil completeness. Coastal areas will likely preserve fossils within a narrow range of conditions, such as lagoonal or estuarine depositional environments. Similar preservation mechanisms will therefore be recorded for both groups during this time, resulting in the observed statistically significant correlation. Increased exploration of a wider range of habitats may have occurred within both sauropods and pterosaurs in the Early Cretaceous. Terrestrial environments naturally exhibit a wider range of depositional settings (Behrensmeyer *et al.* 1992), many of which would be unsuitable for pterosaur preservation, but adequate for sauropodomorphs. As such, this might account for the divergence in correlation during the Cretaceous.

Several counterpoints are presented to this environmental shift hypothesis. Firstly, the observed pattern is not supported as statistically significant in all comparisons, and visual representations of the data show little in the way of correlation between the two data series throughout the Triassic–Jurassic. The general lack of co-occurrences of pterosaurs and sauropodomorphs in a single locality (see the PaleoDB) also suggests that the two groups were not tightly linked ecologically. Additionally, pterosaur diversity and completeness curves for both marine and terrestrial environments (Fig. 5A-D) exhibit little evidence for an increase within terrestrial diversity at the J/K boundary, whereas an increase in marine taxic diversity is observed within the Albian. However, divergence between marine completeness and taxic diversity during this stage is potentially due to taxonomic inflation caused by the Cambridge Greensand Formation, consisting of highly fragmentary remains (Unwin 2001). Furthermore, the potential exists for discovering additional pterosaur-yielding terrestrial Lagerstätten within the Triassic and Jurassic, which currently remain unknown. Another point to consider is the comparison between the nature of the characters used in calculation of the CCM scores for the two groups. The phylogenetic matrix used to calculate the CCM scores of pterosaurs is heavily biased towards cranial characters (making up 59% of the total characters), whereas the sauropodomorph characters are assigned more evenly across the skeleton (Mannion and Upchurch 2010a). As 56 species of pterosaur are known solely from cranial material, with 13 of these appearing in the Triassic–Jurassic and 43 in the Cretaceous (making up 22% and 39% of the pterosaurs found in those time bins respectively), this could confound relationships between these groups. These factors make the testing of this hypothesis difficult, and increased collection of both sauropodomorphs and pterosaurs for the Jurassic–Cretaceous will be needed to fully explore this issue.

CONCLUSIONS

1. Completeness values for 172 species of pterosaur were generated using a Character Completeness Metric (CCM), and tested for correlations with 'raw' and 'corrected' diversity, a variety of sampling proxies, and completeness of other vertebrate groups, to address several critical questions relating to the fossil record of pterosaurs.
2. The earliest fossil record of pterosaurs is shown to be characterized by relatively high levels of completeness, and thus data quality is high; however, taxon counts from the Rhaetian–Bajocian (latest Triassic–early Middle Jurassic) are generally low. A decline in completeness values is recovered across the J/K boundary, with low completeness also characterizing the middle Cretaceous, meaning that apparent extinctions at these times might at least partly reflect sampling artefacts. A taxic decline throughout the Late Cretaceous is observed; however, the Maastrichtian exhibits extremely low completeness, and thus taxic diversity is potentially unreliable.

3. CCM values of pterosaurs and birds exhibit a strong positive correlation throughout the Mesozoic, in contrast to sauropodomorph dinosaurs, supporting claims of taphonomic mode differences between large and small organisms affecting completeness. However, there is a correlation between pterosaur and sauropodomorph CCM when only the Triassic–Jurassic is examined. The absence of such a pattern in the Cretaceous might reflect a move within both groups to increased terrestrial habitation after the J/K boundary.
4. The fossil record of pterosaurs is strongly and pervasively affected by Lagerstätten deposits and heterogeneous sampling, which consequently drives both observed pterosaur taxic diversity and completeness through time.

Acknowledgements

We are grateful for the efforts of all those who have collected pterosaur data, and to all those who have entered this into the Paleobiology Database, especially Matthew Carrano. Reviewer comments from Neil Brocklehurst and Alistair McGowan and comments from the editor Roger Benson improved an earlier version of this work. PDM's research was funded by an Imperial College London Junior Research Fellowship. R.J.B. received funding from the European Union's Horizon 2020 research and innovation programme 2014–2018 under grant agreement 637483 (ERC Starting Grant: TERRA). This is Paleobiology Database official publication number XXX.

SUPPORTING INFORMATION

Data for this study are available in the [Dryad Digital Repository]:

http://dx.doi.org/10.5061/dryad.td570

REFERENCES

Alroy, J., Marshall, C. R., Bambach, R. K., Bezusko, K., Foote, M., Fürsich, F. T., Hansen, T.A., Holland, S. M., Ivany, L. C., Jablonski, D., Jacobs, D. K., Jones, D. C., Kosnik, M. A., Lidgard, S., Low,S., Miller, A. I., Novack-Gottshall, P. M., Olszewski, T. D., Patzkowsky, M. E., Raup, D. M., Roy, K., Sepkoski, J. J. Jr., Sommers, M. G., Wagner, P. J., & Webber, A. 2001. Effects of sampling standardization on estimates of Phanerozoic marine diversification. *Proceedings of the National Academy of Sciences*, **98(11),** 6261-6266.

Andres, B. B. 2010. Systematics of the Pterosauria, PhD Dissertation. Yale University. 366 pp.

--2012. The early evolutionary history and adaptive radiation of the Pterosauria. *Acta Geologica Sinica (English Edition)*, 86(6), 1356-1365.

Andres, B. B., Clark, J., & Xu, X. 2014. The earliest pterodactyloid and the origin of the group. *Current Biology*, **24(9)**, 1011-1016.

Barrett, P. M., Butler, R. J., Edwards, N. P., and Milner. A. R. 2008. Pterosaur distribution in time and space: an atlas. Zitteliana. 61-107.

Behrensmeyer, A. K. 1992. *Terrestrial ecosystems through time: evolutionary paleoecology of terrestrial plants and animals*. University of Chicago Press. 568pp.

Benson, R. B., & Butler, R. J. 2011. Uncovering the diversification history of marine tetrapods: ecology influences the effect of geological sampling biases.*Geological Society, London, Special Publications*, **358 (1)**, 191-208.

Benson, R. B., & Druckenmiller, P. S. 2014. Faunal turnover of marine tetrapods during the Jurassic–Cretaceous transition. *Biological Reviews*, **89 (1)**, 1-23.

Benson, R. B., Frigot, R. A., Goswami, A., Andres, B., & Butler, R. J. 2014. Competition and constraint drove Cope's rule in the evolution of giant flying reptiles. *Nature communications*, **5**, 8 pp.

Benson, R. B., Mannion, P. D., Butler, R. J., Upchurch, P., Goswami, A., & Evans, S. E. 2013. Cretaceous tetrapod fossil record sampling and faunal turnover: implications for biogeography and the rise of modern clades.*Palaeogeography, Palaeoclimatology, Palaeoecology*, **372,** 88-107.

Benton, M. J., Dunhill, A.M., Lloyd G. T., and Marx, F. G., 2011. Assessing the quality of the fossil record: insights from vertebrates. *Geological Society, London, Special Publications,* **358**(1), 63-94.

Benton, M. J., Ruta, M., Dunhill, A. M., and Sakamoto, M. 2013. The first half of tetrapod evolution, sampling proxies, and fossil record quality. *Palaeogeography, Palaeoclimatology, Palaeoecology,* **372** (0), 18-41.

Brocklehurst, N. and J. Fröbisch (2014). Current and historical perspectives on the completeness of the fossil record of pelycosaurian-grade synapsids. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **399**, 114-126.

Benton, M. J., Spencer, P. S., & Broadhurst, F. M. 1995. Fossil Reptiles of Great Britain. *Biological Journal of the Linnean Society*, **55(4)**, 385pp.

Brocklehurst, N., Upchurch, P., Mannion, P. D., and O'Connor, J. 2012. The completeness of the fossil record of Mesozoic birds: implications for early avian evolution. *PloS one*, **7**(6): e39056.

Buffetaut, E. 1995. The importance of "Lagerstätten" for our understanding of the evolutionary history of certain groups of organisms: the case of pterosaurs. *II International Symposium on Lithographic Limestones. Extended Abstracts*. Ediciones de la Universidade Autónoma de Madrid. 49-52.

Buffetaut, E., Clarke, J. B., & Le Loeuff, J. 1996. A terminal Cretaceous pterosaur from the Corbières (southern France) and the problem of pterosaur extinction. *Bulletin de la Société géologique de France*, **167(6)**, 753-759.

Burnham, K. P., & Anderson, D. R. 2001. Kullback-Leibler information as a basis for strong inference in ecological studies. *Wildlife research*, **28** (2), 111-119.

Butler, R. J., & Barrett, P. M. 2008. Palaeoenvironmental controls on the distribution of Cretaceous herbivorous dinosaurs. *Naturwissenschaften*, **95(11)**, 1027-1032.

Butler, R. J., Barrett, P. M., Nowbath, S. and Upchurch, P. 2009. Estimating the effects of sampling biases on pterosaur diversity patterns: implications for hypotheses of bird/pterosaur competitive replacement. *Paleobiology***, 35** (3): 432-446.

Butler, R. J., Benson, R. B., and Barrett, P. M. 2013. Pterosaur diversity: untangling the influence of sampling biases, Lagerstätten, and genuine biodiversity signals. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **372**: 78-87.

Butler, R. J., Brusatte, S. L., Andres, B., and Benson, R. B. 2012. How do geological sampling biases affect studies of morphological evolution in deep time? A case study of pterosaur (Reptilia: Archosauria) disparity. *Evolution*, **66** (1), 147-162.

Carvalho, I. d. S., Novas, F. E., Agnolín, F. L., Isasi, M. P., Freitas, F. I., & Andrade, J. A. 2015. A Mesozoic bird from Gondwana preserving feathers. *Nature Communications*, *6*.

Cleary, T. J., Moon, B. C., Dunhill, A. M., & Benton, M. J. 2015. The fossil record of ichthyosaurs, completeness metrics and sampling biases. *Palaeontology*, **58** (3), 521-536.

Crampton, J.S., Beu, A.G., Cooper, R.A., Jones, C.M., Marshall, B., Maxwell, P.A., 2003. Estimating the rock volume bias in palaeodiversity studies. *Science,* **301**, 358–360.

Dalla Vecchia, F. M. 2003. A review of the Triassic pterosaur record. *Rivista del Museo Civico di Scienze Naturali "E. Caffi" di Bergamo*, **22**, 13-29.

Dunhill, A. M., Hannisdal, B., and Benton, M. J. 2014. Disentangling rock record bias and common-cause from redundancy in the British fossil record. *Nature communications*, **5.**

Dunhill, A. M., Benton, M. J., Newell, A. J., and Twitchett, R. J. 2013. Completeness of the fossil record and the validity of sampling proxies: a case study from the Triassic of England and Wales. *Journal of the Geological Society*, **170** (2): 291-300.

Dunhill, A. M., Benton, M. J., Twitchett, R. J., and Newell, A. J. 2012. Completeness of the fossil record and the validity of sampling proxies at outcrop level. *Palaeontology*, **55** (6), 1155-1175.

Dyke, G. J., McGowan, A., Nudds, R. and D. Smith, D. 2009. The shape of pterosaur evolution: evidence from the fossil record. *Journal of evolutionary biology*, **22**(4), 890-898.

Foth, C., Brusatte, S. L., & Butler, R. J. 2012. Do different disparity proxies converge on a common signal? Insights from the cranial morphometrics and evolutionary history of Pterosauria (Diapsida: Archosauria). *Journal of evolutionary biology*, ***25*(5)**, 904-915.

Haq, B.U., Hardenbol, J., Vail, P.R., 1987. Chronology of fluctuating sea levels since the Triassic. *Science*, **235**, 1156–1167.

Long, J. S., & Freese, J. 2006. *Regression models for categorical dependent variables using Stata*. Stata press.

Lü, J. & Fucha, X. 2010. A new pterosaur (Pterosauria) from Middle Jurassic Tiaojishan Formation of western Liaoning, China. *Global Geology*, **13**, 113–118.

McKinney, M. L. 1990. Classifying and analyzing evolutionary trends. *In* Evolutionary Trends, McNamara K. J., ed. Belhaven, London, pp. 28-58.

Mannion, P. D. and Upchurch, P. 2010a. Completeness metrics and the quality of the sauropodomorph fossil record through geological and historical time. *Paleobiology*, **36** (2), 283-302.

-- 2010b. A quantitative analysis of environmental associations in sauropod dinosaurs. *Journal Information*, **36(2)**, 253-282

Mannion, P. D., Upchurch, P., Carrano, M. T., & Barrett, P. M. 2011. Testing the effect of the rock record on diversity: a multidisciplinary approach to elucidating the generic richness of sauropodomorph dinosaurs through time.*Biological Reviews*, **86** (1), 157-181.

Manzig, P. C., Kellner, A. W., Weinschütz, L. C., Fragoso, C. E., Vega, C. S., Guimarães, G. B., Godoy, L.C., Liccardo, A., Ricetti, J. H. Z., & de Moura, C. C. 2014. Discovery of a rare pterosaur bone bed in a Cretaceous desert with insights on ontogeny and behavior of flying reptiles. *PloS one*, **9** (8), e100005.

Martill, D. M., & Etches, S. 2013. A new monofenestratan pterosaur from the Kimmeridge Clay Formation (Upper Jurassic, Kimmeridgian) of Dorset, England. *Acta Palaeontologica Polonica,* **58** (2): 285–294

Miller K. G., Kominz, M. A., Browning, J. V., Wright, J. D., Mountain, G. S., Katz, M. E.., Sugarman, P. J., Cramer, B. S., Christie-Blick, N., & Pekar, S. F. 2005. The Phanerozoic record of global sea-level change. *Science,* **310**: 1293–1298.

Peters, S. E. 2005. Geologic constraints on the macroevolutionary history of marine animals. *Proceedings of the National academy of Sciences of the United States of America*, **102** (35), 12326-12331.

Peters, S. E., & Foote, M. 2001. Biodiversity in the Phanerozoic: a reinterpretation. *Paleobiology*, **27** (4), 583-601.

J. Pinheiro, Bates, D., Debroy, S., Sarkar, D., R Development Core Team. 2009. nlme: Linear and Nonlinear Mixed Effects Models**.** R package version 3.1-93 (2009)

Prentice, K. C., Ruta, M. and Benton, M. J., 2011. Evolution of morphological disparity in pterosaurs. *Journal of Systematic Palaeontology*, **9** (3), 337-353.

R Core Team. 2013. R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Raup, D. M. 1976. Species diversity in the Phanerozoic: an interpretation. *Paleobiology*, **2** (4), 289-297.

Raup, D. M. & Sepkoski, J. J. Jr. 1984. Periodicity of extinctions in the geologic past. *Proceedings of the National Academy of Sciences of the United States of America,* **81**, 801–805.

Seilacher, A. 1970. Begriff und bedeutung der Fossil-Lagerstätten. Neues Jahrbuch für Geologie und Paläontologie, Monatshefte, **1970**(1), 34-39.

Slack, K.E., Jones, C. M., Ando, T., Harrison, G. A., Fordyce, R. E., Arnason, U., & Penny, D., 2006. Early penguin fossils, plus mitochondrial genomes, calibrate avian evolution. *Molecular biology and evolution*, **23(6)**. 1144–55.

Smith, A. B. 2001. Large–scale heterogeneity of the fossil record: Implications for Phanerozoic biodiversity studies. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, **356** (1407), 351-367.

-- 2007. Marine diversity through the Phanerozoic: problems and prospects. *Journal of the Geological Society*, **164** (4), 731-745.

Smith, A. B. and McGowan, A. J. 2005. Cyclicity in the fossil record mirrors rock outcrop area. *Biology Letters,* **1** (4), 443-445.

Spiess, A.N., & Ritz, C. 2010. qPCR: modelling and analysis of real-time PCR data. R Package Version 1.2-7 (2010) http://CRAN.R-project.org/package=qpcR

Sullivan, C., Wang, Y., David W. E. Hone, D. W. E., Wang, Y., Xu X. & Zhang, F. 2014. The vertebrates of the Jurassic Daohugou Biota of northeastern China. *Journal of Vertebrate Paleontology,* **34** (2), 243–280

Trapletti, A., & Hornik, K. 2009. tseries: time series analysis and computational finance. R Package Version 0.10-22 (2009)

Unwin, D. M. 2003. On the phylogeny and evolutionary history of pterosaurs.*Geological Society, London, Special Publications*, *217*(1), 139-190.

-- 2005. *The pterosaurs from deep time*. Pi Press, New York. 352 pp.

Upchurch, P., Andres, B., Butler, R. J., & Barrett, P. M. 2015. An analysis of pterosaurian biogeography: implications for the evolutionary history and fossil record quality of the first flying vertebrates. *Historical Biology*, **27**, 697–717.

Upchurch, P., Mannion, P.D., Benson, R. B. J., Butler, R. J. and Carrano, M. T. 2011. Geological and anthropogenic controls on the sampling of the terrestrial fossil record: a case study from the Dinosauria. *Geological Society, London, Special Publications*, **358** (1), 209-240.

Walther, M. and Fröbisch, J. 2013. The quality of the fossil record of anomodonts (Synapsida, Therapsida). *Comptes Rendus Paleovol,* **12,** 495-504.

Wang, X., Kellner, A. W. A., Jiang, S., Wang, Q., Ma, Y., Paidoula, Y., Cheng, X., Rodrigues, T., Meng, X., Zhang, J., Li, N. & Zhou, Z. 2014a. Sexually dimorphic tridimensionally preserved pterosaurs and their eggs from China. *Current Biology,* **24** (12), 1323-1330

Wang, X., Rodrigues, T., Jiang, S., Cheng, X., & Kellner, A. W. 2014b. An Early Cretaceous pterosaur with an unusual mandibular crest from China and a potential novel feeding strategy. *Scientific reports*, **4**, 9 pp.

Wang, X. & Zhou, Z., 2006. Pterosaur assemblages of the Jehol Biota and their implication for the Early Cretaceous pterosaur radiation. *Geological Journal*, **41(3-4)**, 405–418.

Wellnhofer, P. 1970. Die Pterodactyloidea (Pterosauria) der Oberjura-Plattenkalke Siiddeutschlands. Bayerische Akademie der Wissenschaften, Mathematisch-Wissenschaftlichen Klasse, Abhandlungen 141.

-- 1975. Die Rhamphorhynchoidea (Pterosauria) der Oberjura-Plattenkalke Suiddeutschlands. Palaeontographica, A 148: 1-33, 148:132-186, 149:1-30.

-- 1991. The illustrated encyclopedia of pterosaurs. Crescent Books, 192 pp.

Witton, M. P. 2013. *Pterosaurs: natural history, evolution, anatomy*. Princeton University Press, 304pp.

Zeilis, A., Hothorn, T. 2002. Diagnostic checking in regression relationships. R News, 2 (2002), 7–10 pp.

Tables and Figure Captions:

Table 1: All abbreviated terms discussed in the paper.

PBF	Pterosaur-Bearing Formations
PBC	Pterosaur-Bearing Collections
DBF	Dinosaur-Bearing Formations
DBC	Dinosaur-Bearing Collections
SCM	Skeletal Completeness Metric
CCM	Character Completeness Metric
RDE	Residual Diversity Estimate
TTBC	Terrestrial Tetrapod-Bearing Collections
MTBC	Marine Tetrapod-Bearing Collections

Table 2: Spearman’s Rank and Kendall’s Tau correlation coefficients and uncorrected probability values (*p*) of statistical comparisons between taxic diversity of pterosaurs and various sampling proxies.

Taxic Div vs:	Spearman's rs	Kendall's tau
PBFs	0.51231 (p=0.00884)	0.34 (p=0.0172)
PBCs	0.56 (p=0.00360)	0.38 (p=0.00776)
DBFs	0.42154 (p=0.0358)	0.3 (p=0.0356)
DBCs	0.43846 (p=0.0283)	0.31333 (p=0.0281)

Table 3: Spearman’s Rank and Kendall’s Tau correlation coefficients and uncorrected probability values (*p*) of statistical comparisons between mean CCM values of pterosaurs and various sampling proxies, for the Mesozoic, the Triassic/Jurassic, and the Cretaceous.

	MESOZOIC		TRIASSIC/JURASSIC		CRETACEOUS	
CCM vs:	*Spearman's rs*	*Kendall's tau*	*Spearman's rs*	*Kendall's tau*	*Spearman's rs*	*Kendall's tau*
Taxic Diversity	**0.53 (*p*=0.00643)**	**0.413 (p=0.00378)**	0.456 (p=0.117)	**0.410 (p=0.0509)**	**0.650 (p=0.0220)**	**0.515 (p=0.0197)**
PBFs	-0.00538 (p=0.980)	-0.00667 (p=0.963)	-0.0110 (p=0.972)	-0.0256 (p=0.903)	0.0490 (p=0.880)	0.0606 (p=0.784)
PBCs	0.213 (p=0.306)	0.18 (p=0.207)	0.280 (p=0.354)	0.256 (p=0.222)	0.168 (p=0.602)	0.182 (p=0.411)
DBFs	0.287 (p=0.164)	0.207 (p=0.148)	0.467 (p=0.108)	0.333 (p=0.113)	0.238 (p=0.457)	0.182 (p=0.411)
DBCs	0.282 (p=0.173)	0.233 (p=0.102)	0.412 (p=0.162)	0.359 (p=0.0876)	0.273 (p=0.391)	0.182 (p=0.411)
PBF Residuals	**0.604 (p=0.0013924)**	**0.453 (p=0.0014918)**	-0.0110 (p=0.972)	-0.0256 (p=0.903)	**0.727 (p=0.00736)**	**0.576 (p=0.00917)**
PBC Residuals	0.211 (p=0.312)	0.127 (p=0.375)	0.280 (p=0.354)	0.256 (p=0.222)	0.238 (p=0.457)	0.091 (p=0.681)
DBF Residuals	**0.539 (p=0.00541)**	**0.427 (p=0.00279)**	0.467 (p=0.108)	0.333 (p=0.113)	**0.720 (p=0.00824)**	**0.515 (p=0.0197)**
DBC Residuals	**0.469 (p=0.018)**	**0.327 (p=0.0221)**	0.412 (p=0.162)	0.359 (p=0.0876)	**0.615 (p=0.0332)**	0.394 (p=0.0746)
Sea Level	0.0269 (p=0.898)	0.0333 (p=0.815)	0.00549 (p=0.986)	-0.0256 (p=0.903)	-0.182 (p=0.572)	-0.0606 (p=0.784)

Results in bold highlight statistical significance at p=0.05

Table 4: Spearman’s Rank and Kendall’s Tau correlation coefficients and uncorrected probability values (*p*) of statistical comparisons between mean CCM values of pterosaurs, birds and sauropodomorphs, for the Mesozoic, Triassic/Jurassic, and the Cretaceous.

	MESOZOIC		TRIASSIC/JURASSIC		CRETACEOUS	
CCM vs:	*Spearman's rs*	*Kendall's tau*	*Spearman's rs*	*Kendall's tau*	*Spearman's rs*	*Kendall's tau*
Sauropod CCM	0.148 (p=0.479)	0.133 (p=0.350)	**0.670 (p=0.0122)**	**0.513 (p=0.0147)**	-0.510 (p=0.0899)	-0.212 (p=0.337)
Aves CCM	**0.758 (p=0.00267)**	**0.641 (p=0.00229)**	N/A	N/A	**0.804 (p=0.00161)**	**0.727 (p=0.000997)**

Results in bold highlight statistical significance at p=0.05

Table 5: Spearman’s Rank and Kendall’s Tau correlation coefficients and uncorrected probability values (*p*) of statistical comparisons between mean CCM values of pterosaurs and various independent variables when the Aalenian is removed for the Mesozoic, Triassic/Jurassic, and the Cretaceous

AALENIAN REMOVED	MESOZOIC		TRIASSIC/JURASSIC		CRETACEOUS	
CCM vs:	*Spearman's rs*	*Kendall's tau*	*Spearman's rs*	*Kendall's tau*	*Spearman's rs*	*Kendall's tau*
Taxic Diversity	**0.412 (p=0.0453)**	**0.304 (p=0.0372)**	0.322 (p=0.308)	0.303 (p=0.170)	**0.587 (p=0.0446)**	0.424 (p=0.0549)
PBFs	-0.183 (p=0.393)	-0.116 (p=0.427)	-0.224 (p=0.484)	-0.182 (p=0.411)	-0.0559 (p=0.863)	-0.0303 (p=0.891)
PBCs	0.142 (p=0.50883)	0.130 (p=0.372)	0.133 (p=0.680)	0.121 (p=0.583)	0.14 (p=0.665)	0.152 (p=0.493)
DBFs	0.163 (p=0.448)	0.130 (p=0.372)	0.252 (p=0.43)	0.152 (p=0.493)	0.175 (p=0.587)	0.212 (p=0.337)
DBCs	0.133 (p=0.535)	0.123 (p=0.399)	0.28 (0.379)	0.273 (p=0.217)	0.161 (p=0.618)	0.121 (p=0.583)
Sauropod CCM	0.027 (p=0.901)	0.029 (p=0.843)	**0.657 (p=0.0202)**	**0.455 (p=0.0397)**	-0.476 (p=0.118)	-0.212 (p=0.337)
Aves CCM	**0.841 (p=0.000319)**	**0.667 (p=0.00151)**	N/A	N/A	**0.839 (p=0.000643)**	**0.697 (p=0.00161)**
Sea Level	0.0548 (p=0.799)	0.0362 (p=0.804)	0.126 (p=0.697)	0.0909 (0.681)	-0.23077 (p=0.47053)	-0.152 (p=0.493)
PBF Residuals	**0.461 (p=0.0234)**	**0.333 (p=0.0225)**	0.54)	0.273 (p=0.217)	**0.73427 (p=0.0065435)**	**0.515 (p=0.197)**
PBC Residuals	0.183 (p=0.393)	0.159 (p=0.275)	0.147 (p=0.649)	0.152 (p=0.493)	0.18881 (p=0.55674)	0.152 (p=0.493)
DBF Residuals	**0.422 (p=0.0401)**	**0.319 (p=0.0291)**	0.364 (p=0.225)	0.333 (p=0.131)	**0.580 (p=0.0479)**	**0.394 (p=0.0746)**
DBC Residuals	0.335 (p=0.11)	0.232 (p=0.1124)	0.203 (p=0.527)	0.182 (p=0.411)	0.392 (p=0.208)	0.242 (p=0.273)

Results in bold highlight statistical significance at p=0.05

Table 6: Spearman's Rank and Kendall's Tau correlation coefficients and uncorrected probability values (*p*) of statistical comparisons between CCM values of pterosaurs and various proxies within Lagerstätten only and Lagerstätten removed subgroups for the Mesozoic, Triassic/Jurassic, and the Cretaceous.

	MESOZOIC			
	LAGERSTÄTTEN ONLY		LAGERSTÄTTEN REMOVED	
CCM vs:	*Spearman's rs*	*Kendall's tau*	*Spearman's rs*	*Kendall's tau*
Taxic Diversity	**0.738 (p=0.0000258)**	**0.547 (p=0.000128)**	0.328 (p=0.109)	0.253 (p=0.0759)
Lagerstatten presence/absence Tax. Div.	**0.936 (p=0.00000000000633)**	**0.84 (p=0.00000000397)**	0.219 (p=0.292)	0.16 (p=0.262)
Sauropod CCM	-0.00308 (p=0.988)	-0.0133 (p=0.926)	0.182 (p=0.383)	0.147 (p=0.304)
Aves CCM	**0.615 (p=0.0252)**	**0.385 (p=0.0672)**	0.418 (p=0.156)	0.385 (p=0.0672))
Sea Level	0.236 (p=0.256)	0.18 (p=0.207)	-0.0654 (p=0.756)	-0.02 (p=0.744)

Results in bold highlight statistical significance at p=0.05

	TRIASSIC/JURASSIC			
	LAGERSTÄTTEN ONLY		LAGERSTÄTTEN REMOVED	
CCM vs:	*Spearman's rs*	*Kendall's tau*	*Spearman's rs*	*Kendall's tau*
Taxic Diversity	**0.615 (p=0.0252)**	**0.385 (p=0.0672)**	0.302 (p=0.316)	0.231 (p=0.272)
Lagerstatten presence/absence Tax. Div.	**0.967 (p=0.0000000706)**	**0.897 (p=0.0000195)**	0.275 (p=0.364)	0.179 (p=0.393)
Sauropod CCM	0.445 (p=0.128)	0.333 (p=0.113)	0.489 (p=0.0899)	0.387 (p=0.0672)
Aves CCM	N/A	N/A	N/A	N/A
Sea Level	0.214 (p=0.482)	0.103 (p=0.626)	-0.0824 (p=0.789)	-0.103 (p=0.626)

Results in bold highlight statistical significance at p=0.05

	CRETACEOUS			
	LAGERSTÄTTEN ONLY		LAGERSTÄTTEN REMOVED	
CCM vs:	*Spearman's rs*	*Kendall's tau*	*Spearman's rs*	*Kendall's tau*
Taxic Diversity	**0.755 (p=0.00451)**	**0.636 (p=0.00398)**	0.294 (p=0.354)	0.212 (p=0.337)
Lagerstatten presence/absence Tax. Div.	**0.944 (p=0.00000393)**	**0.848 (p=0.000123)**	0.105 (p=0.746)	0.0909 (p=0.681)
Sauropod CCM	-0.538 (p=0.0709)	-0.394 (p=0.0746)	-0.035 (p=0.914)	0.0303 (p=0.891)
Aves CCM	**0.622 (p=0.0307)**	**0.424 (p=0.0549)**	0.378 (p=0.226)	0.364 (p=0.0998)
Sea Level	0.119 (p=0.713)	0.121 (p=0.583)	-0.189 (p=0.557)	-0.121 (0.583)

Results in bold highlight statistical significance at p=0.05

Table 7: Spearman’s Rank and Kendall’s Tau correlation coefficients and uncorrected probability values (*p*) of statistical comparisons between CCM values of pterosaurs and various proxies within marine and terrestrial subgroups for the Mesozoic, Triassic/Jurassic, and the Cretaceous.

	MESOZOIC				TRIASSIC/JURASSIC				CRETACEOUS			
	TERRESTRIAL		MARINE		TERRESTRIAL		MARINE		TERRESTRIAL		MARINE	
CCM vs:	*Spearman's rs*	*Kendall's tau*	*Spearman's rs*	*Kendall's tau*	*Spearman's rs*	*Kendall's tau*	*Spearman's rs*	*Kendall's tau*	*Spearman's rs*	*Kendall's tau*	*Spearman's rs*	*Kendall's tau*
Taxic Diversity	**0.587 (p=0.00204)**	**0.407 (p=0.00438)**	0.247 (p=0.234)	0.12 (p=0.4)	**0.555 (p=0.049)**	**0.385 (p=0.0672)**	0.269 (p=0.374)	0.154 (p=0.464)	0.531 (p=0.0754)	0.394 (p=0.0746)	0.028 (p=0.931)	0.0303 (p=0.891)
Env. Taxic Diversity	**0.585 (p=0.00215)**	**0.44 (p=0.00205)**	**0.686 (p=0.000153)**	**0.527 (p=0.000224)**	**0.714 (p=0.00609)**	**0.538 (p=0.0104)**	**0.615 (p=0.0252)**	**0.513 (p=0.0147)**	0.413 (p=0.183)	0.364 (p=0.0998)	0.028 (p=0.931)	0.0303 (p=0.891)
Group TBCs	-0.0454 (p=0.829)	-0.0333 (p=0.815)	**0.536 (p=0.00573)**	**0.373 (p=0.0089)**	-0.209 (p=0.494)	-0.231 (p=0.272)	**0.626 (p=0.022)**	**0.462 (p=0.0281)**	0.413 (p=0.183)	0.303 (p=0.170)	0.294 (p=0.354)	0.212 (p=0.337)
Sauropod CCM	-0.00615 (p=0.977)	-0.00667 (p=0.963)	0.283 (p=0.170)	0.187 (p=0.191)	-0.011 (p=0.972)	-0.0769 (p=0.714)	**0.571 (p=0.0413)**	**0.410 (p=0.0509)**	-0.266 (p=0.404)	-0.152 (p=0.493)	-0.21 (p=0.513)	-0.182(p=0.411)
Aves CCM	0.209 (p=0.494)	0.154 (p=0.464)	0.473 (p=0.103)	0.308 (p=0.143)	N/A	N/A	N/A	N/A	0.238 (p=0.457)	0.182 (p=0.411)	0.476 (p=0.118)	0.333 (p=0.131)
Sea Level	0.162 (p=0.438)	0.107 (p=0.455)	-0.115 (p=0.583)	-0.113 (p=0.427)	0.368 (p=0.216)	0.256 (p=0.222)	-0.346 (p=0.247)	-0.333 (p=0.113)	-0.112 (p=0.729)	-0.0606 (p=0.784)	-0.042 (p=0.897)	-0.0909 (p=0.681)

Results in bold highlight statistical significance at p=0.05

Table 8: Spearman’s Rank and Kendall’s Tau correlation coefficients and uncorrected probability values (*p*) of statistical comparisons between CCM values of pterosaurs and various proxies within Pterodactyloidea and non-Pterodactyloidea subgroups.

	MESOZOIC			
	PTERODACTYLOIDEA		NON PTERODACTYLOIDEA	
CCM vs:	*Spearman's rs*	*Kendall's tau*	*Spearman's rs*	*Kendall's tau*
Taxic Diversity	0.389 (p=0.152)	0.352 (p=0.0671)	0.522 (p=0.0673)	**0.462 (p=0.00281)**
Group Diversity	**0.729 (p=0.00207)**	**0.6 (p=0.00182)**	0.121 (p=0.694)	0.128 (p=0.542)
Sauropod CCM	-0.482 (p=0.0688)	-0.314 (p=0.102)	**0.714 (p=0.00609)**	**0.564 (p=0.00727)**
Aves CCM	**0.758 (p=0.00267)**	**0.564 (p=0.00727)**	N/A	N/A
Sea Level	-0.114 (p=0.685)	-0.0476 (p=0.805)	0.0714 (p=0.817)	0.0256 (p=0.903)

Results in bold highlight statistical significance at p=0.05

Table 9: Multiple regression results showing sampling proxy effects on pterosaur completeness.

	Full				Best			
Dependent	*Independents*	*R²*	*p*	*AIC*	*Independents*	*R²*	*p*	*AIC*
CCM	Lagerstätten Presence/absence + PBFs + PBCs + DBFs + DBCs + Diversity + Stage Length	0.22	0.58	10.5	PBFs + Diversity	0.17	0.14	4.11

Details of full and best model selection process can be found in Supporting Information.

Table 10: Statistical comparison of the top twenty possible explanatory models for completeness of pterosaurs through the Mesozoic.

Model	*Parameters*	*df*	*Weighting*	*AICc*	*AIC*	*BIC*	*logLik*
112	DfL	5	0.11070917	-2.172124414	-3.372124414	2.518144738	6.686062207
119	LPf	5	0.079832147	-1.518163596	-2.718163596	3.172105555	6.359081798
99	LPcPf	6	0.068188103	-1.202851277	-3.308114435	3.760208547	7.654057217
120	PcPf	5	0.054518999	-0.755409579	-1.955409579	3.934859573	5.977704789
94	DfLPf	6	0.050394272	-0.598066097	-2.703329255	4.364993727	7.351664627
81	DcDfL	6	0.049432517	-0.559527974	-2.664791132	4.40353185	7.332395566
90	DfSL	6	0.02663663	0.677114286	-1.428148872	5.64017411	6.714074436
71	DDfL	6	0.026328567	0.700379821	-1.404883337	5.663439645	6.702441669
93	DfLPc	6	0.026161243	0.71313088	-1.392132278	5.676190704	6.696066139
79	DPcPf	6	0.0250594	0.799190877	-1.306072281	5.762250702	6.65303614
54	DcDfLPf	7	0.02335848	0.939768771	-2.393564563	5.85281225	8.196782281
63	DfLPcPf	7	0.023135291	0.958970553	-2.37436278	5.872014032	8.18718139
97	SLPf	6	0.022332941	1.0295634	-1.075699758	5.992623225	6.537849879
78	DLPf	6	0.022278074	1.034483014	-1.070780144	5.997542839	6.535390072
88	DcLPf	6	0.019942075	1.256025272	-0.849237886	6.219085096	6.424618943
98	SPcPf	6	0.019289483	1.32256884	-0.782694318	6.285628665	6.391347159
64	SLPcPf	7	0.017452414	1.522732971	-1.810600362	6.43577645	7.905300181
105	DPf	5	0.015889977	1.710311852	0.510311852	6.400581004	4.744844074
49	DLPcPf	7	0.014044312	1.957253944	-1.376079389	6.870297424	7.688039694
59	DcLPcPf	7	0.013834353	1.987379236	-1.345954097	6.900422715	7.672977049

Models comprise different combinations of diversity, dinosaur and pterosaur bearing collections and formations, stage length and Lagerstätten presence/absence that potentially explain the character completeness metric. Models are ranked in order of explanatory power according to the Akaike's second-order corrected information criterion (AICc), where a smaller value equals a superior fit. The full table can be found with Supporting Information. Abbreviations of parameters: **D,** diversity. **Df,** Dinosaur Bearing Formations. **Dc,** Dinosaur Bearing Collections. **S,** Stage Length. **L,** Lagerstätten presence/absence. **Pf,** Pterosaur Bearing Formations. **Pc,** Pterosaur Bearing Collections.

Table 11: Statistical comparison of the top twenty possible explanatory models for diversity of pterosaurs through the Mesozoic.

Model	*Parameters*	*df*	*Weights*	*AICc*	*AIC*	*BIC*	*logLik*
119	LPf	5	0.161913896	4.054284125	2.854284125	8.744553276	3.572857938
97	SLPf	6	0.100306361	5.011955262	2.906692104	9.975015086	4.546653948
118	LPc	5	0.088667634	5.258623631	4.058623631	9.948892783	2.970688184
99	LPcPf	6	0.072214036	5.669144623	3.563881465	10.63220445	4.218059267
96	SLPc	6	0.051663081	6.338926642	4.233663484	11.30198647	3.883168258
78	CLPf	6	0.044769643	6.625352899	4.520089741	11.58841272	3.739955129
94	DfLPf	6	0.041389992	6.782335268	4.677072111	11.74539509	3.661463945
88	DcLPf	6	0.037942491	6.956270247	4.851007089	11.91933007	3.574496455
93	DfLPc	6	0.034870645	7.125122796	5.019859638	12.08818262	3.490070181
64	SLPcPf	7	0.033949556	7.178661915	3.845328581	12.09170539	5.077335709
87	DcLPc	6	0.024699875	7.81481714	5.709553982	12.77787696	3.145223009
77	CLPc	6	0.024044448	7.868605272	5.763342114	12.8316651	3.118328943
47	CSLPf	7	0.021915304	8.054043071	4.720709737	12.96708655	4.639645131
61	DfSLPf	7	0.020836498	8.155001204	4.82166787	13.06804468	4.589166065
57	DcSLPf	7	0.020757362	8.162611565	4.829278232	13.07565504	4.585360884
60	DfSLPc	7	0.016659272	8.602479664	5.269146331	13.51552314	4.365426834
49	CLPcPf	7	0.015331366	8.76861193	5.435278597	13.68165541	4.282360701
63	DfLPcPf	7	0.015171737	8.789544867	5.456211534	13.70258835	4.271894233
59	DcLPcPf	7	0.014771424	8.843024555	5.509691222	13.75606803	4.245154389
46	CSLPc	7	0.012575983	9.164835738	5.831502405	14.07787922	4.084248798

Models comprise different combinations of completeness, dinosaur and pterosaur bearing collections and formations, stage length and Lagerstätten presence/absence that potentially explain pterosaur diversity. Models are ranked in order of explanatory power according to the Akaike's second-order corrected information criterion (AICc), where a smaller value equals a superior fit. The full table can be found with Supporting Information. Abbreviations of parameters: **C,** Completeness. **Df,** Dinosaur Bearing Formations. **Dc,** Dinosaur Bearing Collections. **S,** Stage Length. **L,** Lagerstätten presence/absence. **Pf,** Pterosaur Bearing Formations. **Pc,** Pterosaur Bearing Collections.

Figure 1:

Completeness and taxic diversity of pterosaurs throughout the Mesozoic. Light blue areas surrounding Character Completeness Metric represent one standard deviation around the mean.

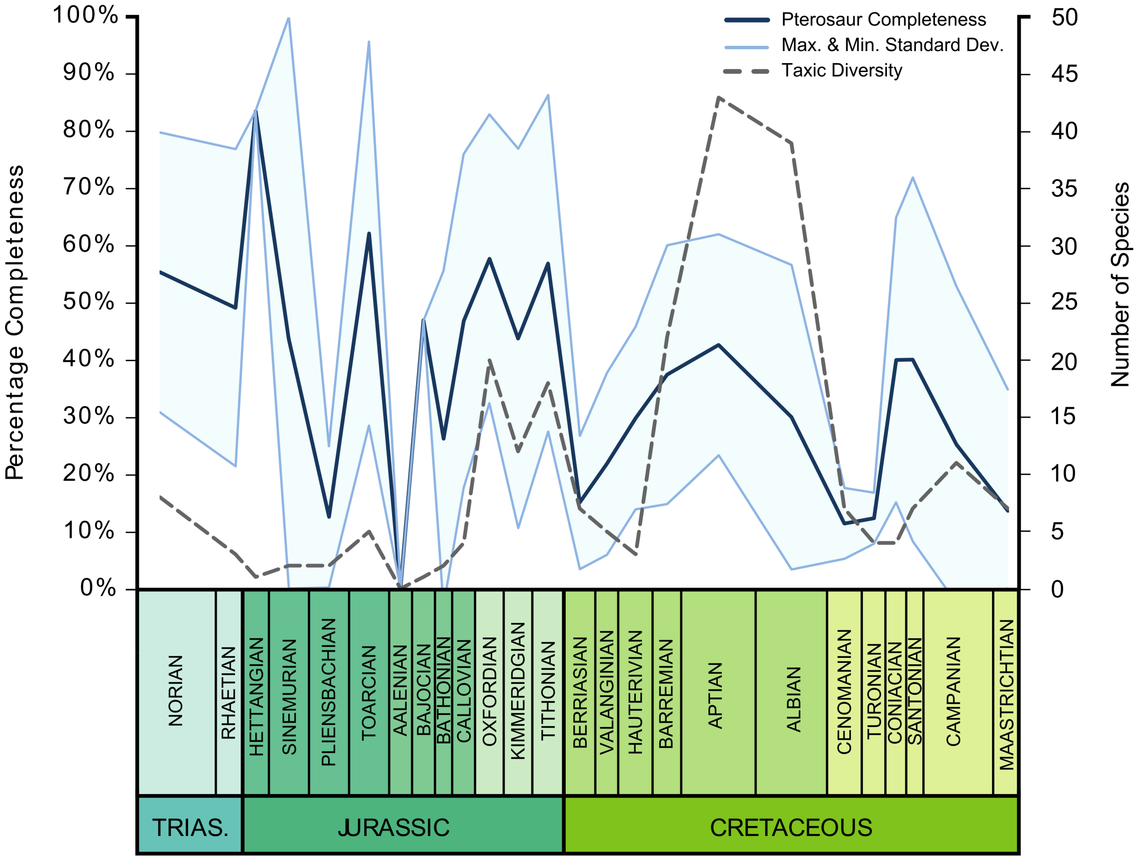

Figure 2:

Sampling proxies plotted throughout the Mesozoic. **A**, Pterosaur completeness, Pterosaur Bearing Formations and Pterosaur Bearing Collections. **B**, Pterosaur completeness and Dinosaur Bearing Formations. **C**, Pterosaur completeness and Dinosaur Bearing Collections.

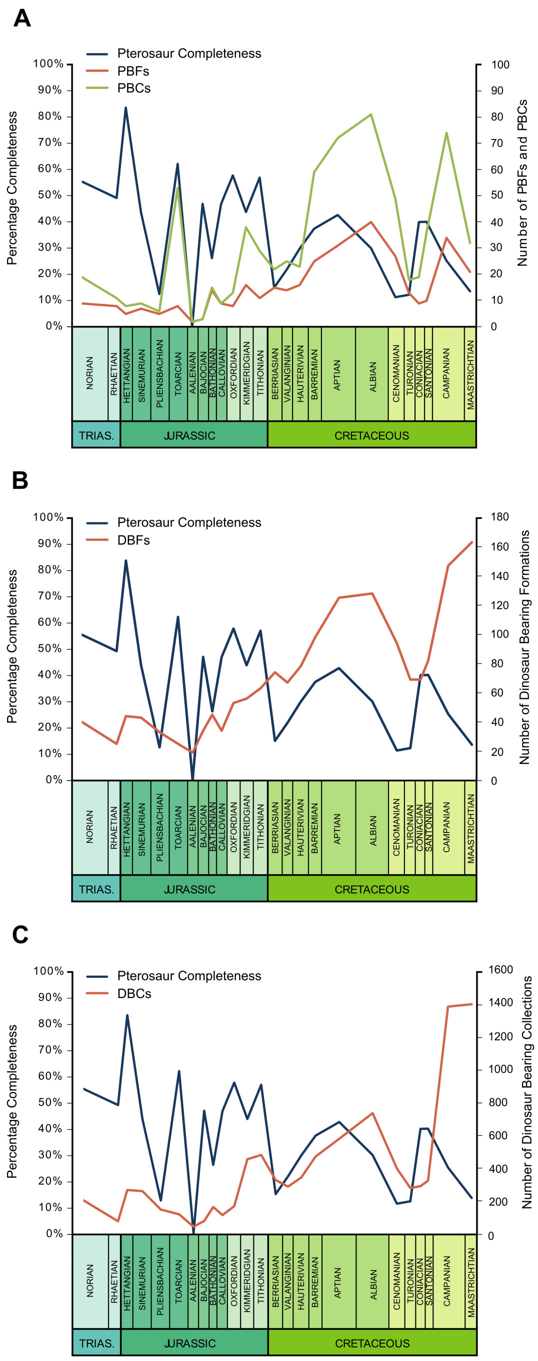

Figure 3:

Pterosaur, bird and sauropodomorph completeness plotted throughout the Mesozoic. Bird completeness is not available prior to the Tithonian.

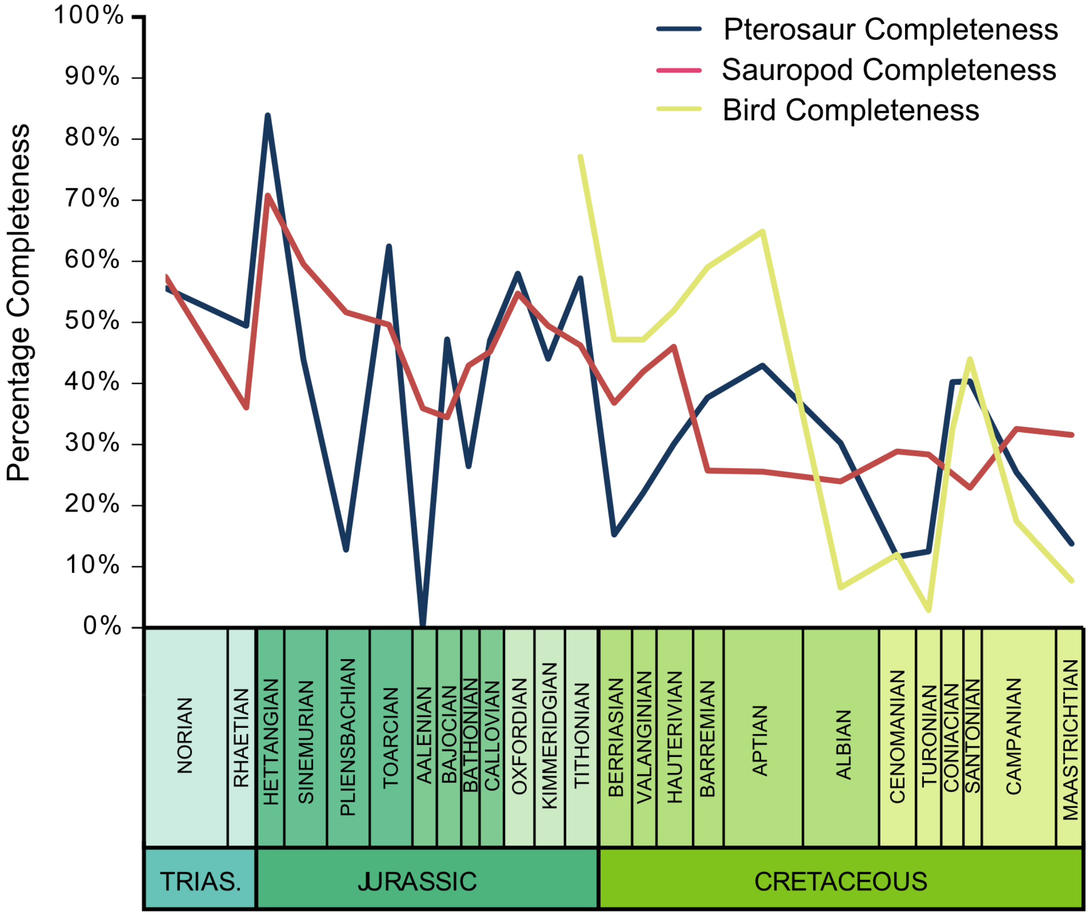

Figure 4:

Plots of Lagerstätten groupings plotted throughout the Mesozoic. **A,** Overall pterosaur completeness compared with subgrouping of Lagerstätten only and Lagerstätten removed completeness. **B,** Overall pterosaur diversity compared with subgrouping of Lagerstätten only and Lagerstätten removed diversity. **C,** Lagerstätten only completeness and diversity. **D,** Lagerstätten removed completeness and diversity.

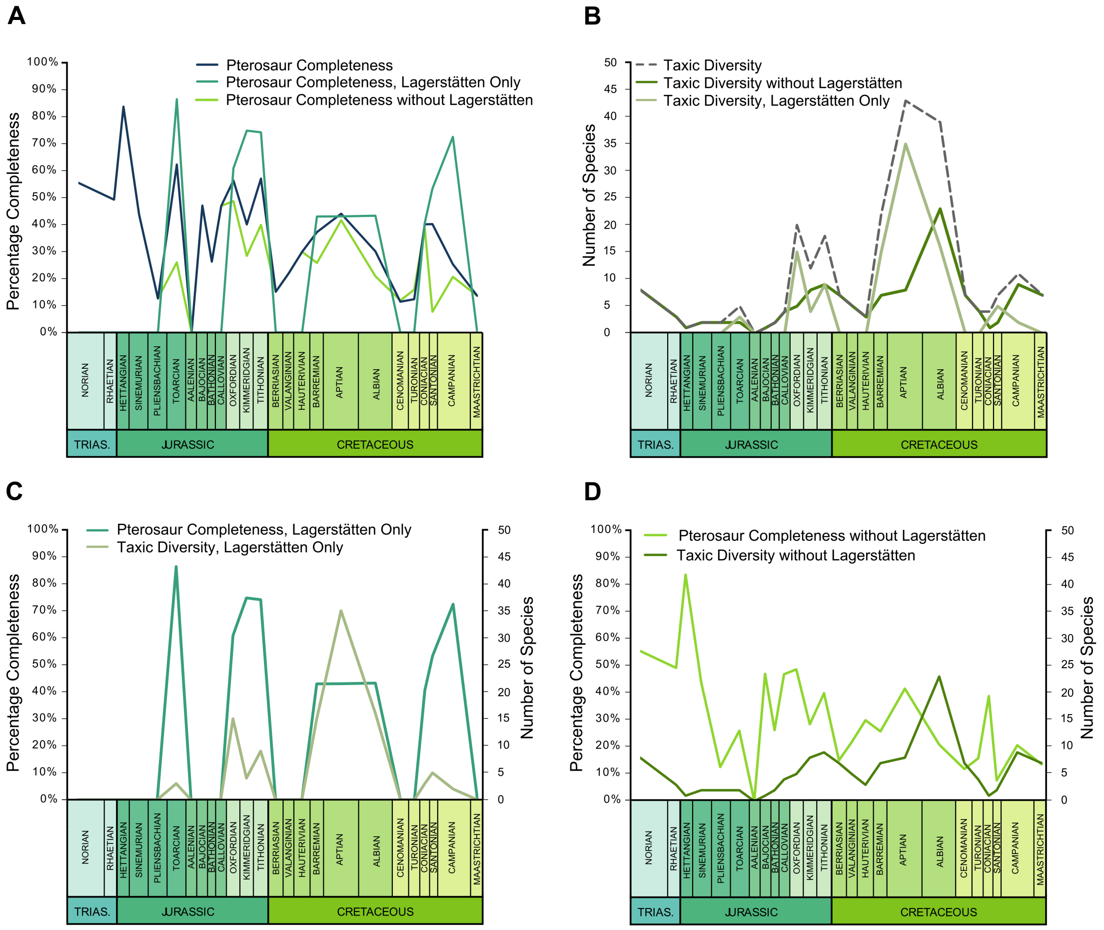

Figure 5:

Plots of pterosaurs found within marine and terrestrial environments for all stages of the Mesozoic. **A,** Overall pterosaur completeness compared with subgrouping of marine and terrestial completeness. **B,** Overall pterosaur diversity compared with subgrouping of marine and terrestial diversity. **C,** Marine completeness and diversity. **D,** Terrestrial completeness and diversity.

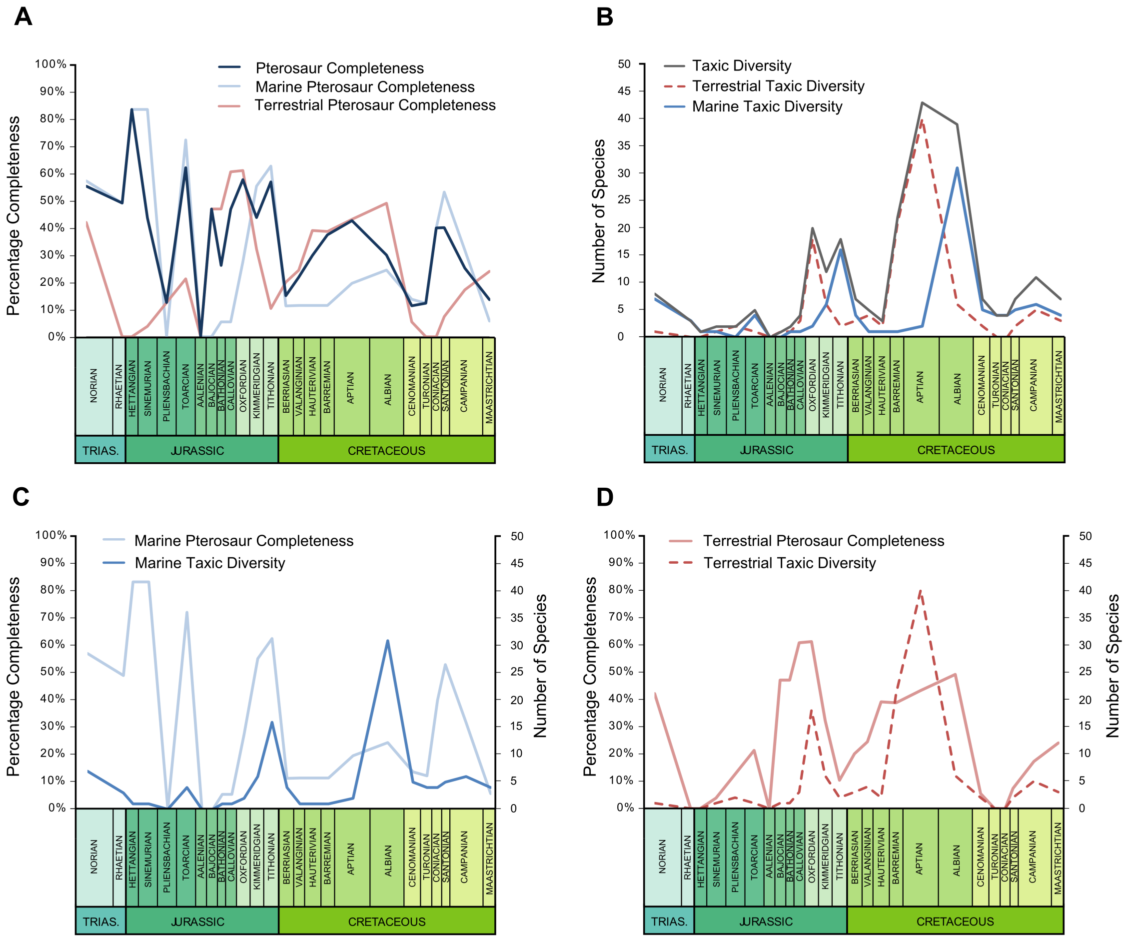

Figure 6:

Plots showing completeness and diversity in subgrouping of Pterodactyloidea and non-Pterodactyloidea throughout the Mesozoic.

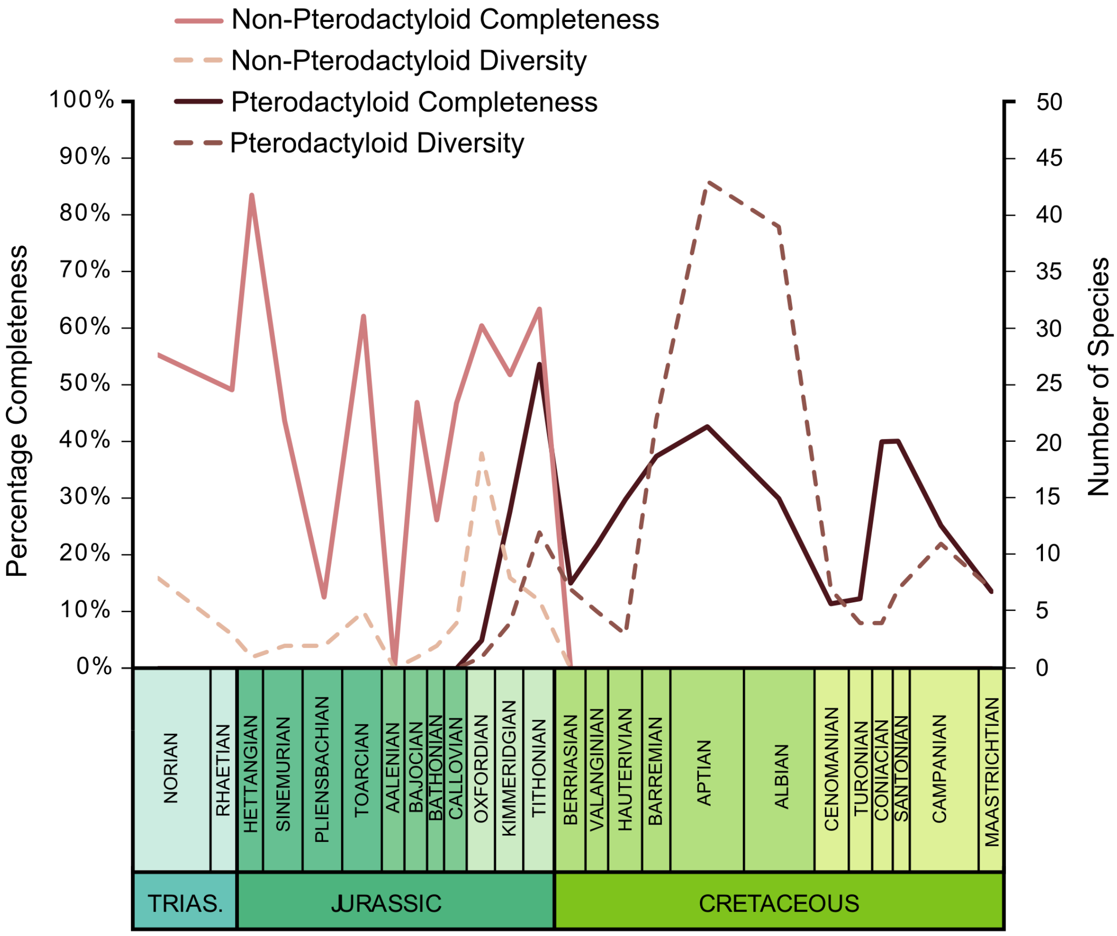

Figure 7:

Comparison between global sea level (Butler et al. 2013, after Haq et al. 1987) and pterosaur completeness for the Mesozoic.