CATEGORICAL VERSUS GEOMETRIC MORPHOMETRIC APPROACHES TO CHARACTERIZING THE EVOLUTION OF MORPHOLOGICAL DISPARITY IN OSTEOSTRACI (VERTEBRATA, STEM GNATHOSTOMATA)

by HUMBERTO G. FERRÓN[1,2]* (iD), JENNY M. GREENWOOD[1], BRADLEY DELINE[3], CARLOS MARTÍNEZ-PÉREZ[1,2], HÉCTOR BOTELLA[2], ROBERT S. SANSOM[4], MARCELLO RUTA[5] (iD) *and* PHILIP C. J. DONOGHUE[1],* (iD)

[1]School of Earth Sciences, University of Bristol, Life Sciences Building, Tyndall Avenue, Bristol, BS8 1TQ, UK; humberto.ferron@bristol.ac.uk, carlos.martinez-perez@bristol.ac.uk, phil.donoghue@bristol.ac.uk
[2]Institut Cavanilles de Biodiversitat i Biologia Evolutiva, Universitat de València, C/ Catedràtic José Beltrán Martínez 2, 46980, Paterna, Valencia, Spain; hector.botella@uv.es, carlos.martinez-perez@uv.es
[3]Department of Geosciences, University of West Georgia, Carrollton, GA 30118, USA; bdeline@westga.edu
[4]School of Earth & Environmental Sciences, University of Manchester, Manchester, M13 9PT, UK; robert.sansom@manchester.ac.uk
[5]School of Life Sciences, University of Lincoln, Riseholme Hall, Lincoln, LN2 2LG, UK; mruta@lincoln.ac.uk
*Corresponding authors

Abstract: Morphological variation (disparity) is almost invariably characterized by two non-mutually exclusive approaches: (1) quantitatively, through geometric morphometrics; and (2) in terms of discrete, 'cladistic', or categorical characters. Uncertainty over the comparability of these approaches diminishes the potential to obtain nomothetic insights into the evolution of morphological disparity and the few benchmarking studies conducted so far show contrasting results. Here, we apply both approaches to characterizing morphology in the stem-gnathostome clade Osteostraci in order to assess congruence between these alternative methods as well as to explore the evolutionary patterns of the group in terms of temporal disparity and the influence of phylogenetic relationships and habitat on morphospace occupation. Our results suggest that both approaches yield similar results in morphospace occupation and clustering, but also some differences indicating that these metrics may capture different aspects of morphology. Phylomorphospaces reveal convergence towards a generalized 'horseshoe'-shaped cranial morphology and two strong trends involving major groups of osteostracans (benneviaspidids and thyestiids), which probably reflect adaptations to different lifestyles. Temporal patterns of disparity obtained from categorical and morphometric approaches appear congruent, however, disparity maxima occur at different times in the evolutionary history of the group. The results of our analyses indicate that categorical and continuous data sets may characterize different patterns of morphological disparity and that discrepancies could reflect preservational limitations of morphometric data and differences in the potential of each data type for characterizing more or less inclusive aspects of overall phenotype.

Key words: disparity, morphospace, categorical data, geometric morphometrics, Osteostraci.

QUANTIFICATION of morphology and morphological disparity is integral to assessing general macroevolutionary patterns in the fossil record, such as adaptive radiations, rates of evolution, responses to extinctions, biotic replacements and the existence of constraints on form (Foote 1997 and references therein). Consequently, there have been many attempts to evaluate the impact of potential biasing factors on morphological disparity estimates, including the possible influence of taphonomy (Webster & Hughes 1999; Barış 2017), taxonomic or geographic scale (Butler *et al.* 2012; Deline *et al.* 2012), environmental distribution (Hopkins 2014), community structure (Deline 2009), sampling of phenotypic characters (Foth *et al.* 2012; Hopkins 2017), choice of metric (Ciampaglio *et al.* 2001), character selection (Deline & Ausich 2017) and choice of methodology (Villier & Eble 2004; Hetherington *et al.* 2015; Hopkins 2017; Romano *et al.* 2017).

Morphological disparity and morphospace occupation patterns have conventionally been based on two non-mutually exclusive approaches to characterizing morphology. Firstly, through geometric morphometrics, morphology is characterized quantitatively in terms of continuous variation in aspects of organismal shape and proportions (e.g. Stayton & Ruta 2006). This approach is most commonly used to assess morphospace occupation at lower taxonomic levels due to the need for homologous features in all specimens within a single study (Wills *et al.* 1994; Bookstein 1997). Therefore, while these methods are very sensitive to morphological variation, their effectiveness in characterizing overall phenotype diminishes with increasing taxonomic scale reflecting the concomitant reduction in the number of universally shared homologous features (Hetherington *et al.* 2015). Alternatively, morphology can be characterized in terms of discrete or categorical characters of the sort most commonly employed in the cladistic analysis of morphology. This approach to characterizing morphology is usually undertaken in analyses of broad taxonomic scope, where there are large numbers of categorical differences between taxa, but fewer universal homologous structures that might serve as a basis for geometric morphometrics. As such, the categorical characterization of morphology can capture more unique aspects of form (Briggs *et al.* 1992; Wills 1998). However, this approach is also much less sensitive to finer changes in shape and proportion. Despite the differences between these approaches, most of the small number of benchmarking studies that have been conducted have shown that continuous and categorical approaches to characterizing morphology yield broadly congruent patterns of disparity (Villier & Eble 2004; Anderson & Friedman 2012; Foth *et al.* 2012; Hetherington *et al.* 2015; Hopkins 2017; Romano *et al.* 2017; Schaeffer *et al.* 2019). Nevertheless, Mongiardino Koch *et al.* (2017) have shown that these two approaches yield contrasting patterns of disparity in their analysis of the scorpion genus *Brachistosternus*. This difference may reflect the relative power of continuous versus categorical approaches to characterizing morphology at different taxonomic scales. However, this study may also reflect the fact that benchmarking studies remain small in number and many more such studies are needed to obtain nomothetic insights into whether different approaches to characterizing morphology impact on the perception of the ensuing patterns of morphological disparity and the evolutionary processes that are inferred from them. To that end, we explore the evolution of morphological disparity in osteostracans, the extinct clade of jawless stem-gnathostomes that is generally perceived to be most closely related to jawed vertebrates (Janvier 1996; Donoghue & Keating 2014). As such, the evolution of morphological disparity within this clade is interesting in its own right, as a parallel to its sister-lineage of jawed

vertebrates. However, cranial shape characters contribute extensively to the systematics of the group (Sansom 2009*a*), making osteostracans an ideal focus for benchmarking the characterization of morphology for disparity analysis using discrete vs categorical characters. We characterize the morphology of this clade based on the distinctive and character-rich headshield, both in terms of geometric morphometric and discrete categorical 'cladistic' data. We compare perceptions of morphological disparity based on these data sets and, further, interpret the results in terms of their implications for understanding this temporal, phylogenetic and ecological context of the evolution of this key clade.

Osteostraci as a case study

Osteostracans constitute a clade of extinct jawless vertebrates that range from the Llandovery (lower Silurian) to the Frasnian (Upper Devonian) (Sansom 2009*a*). Osteostracans are a taxonomically and morphologically diverse group, characterized by a semicircular dermal headshield encompassing the cranial and pectoral regions, a number of cephalic fields, pineal foramen, nasohypophysial opening, and a postcranial body covered in thick scales (Janvier 1996). Cornual and/or rostral processes extending from the headshield are well developed in a number of independent lineages and paired fins are present in many taxa but lost in others (Janvier 1985; Sansom 2008), or they evolved convergently between jawed vertebrates and derived Osteostraci (Denison 1951). Most osteostracans have headshields that are strongly oblate dorsoventrally although some have a more approximately circular or prolate in cross-sectional profile (Janvier 1996) (Fig. 1A–B). Remarkably, despite being a potential source of valuable biological and ecological information (e.g. Janvier & Lawson 1985; Belles-Isles 1987; Bunker & Machin 1991; Afanassieva 1992; Mark-Kurik 1992; Janvier 1996; Morrissey *et al.* 2004; Davies 2009), such morphological disparity has not been analysed quantitatively. The headshield is the aspect of osteostracan anatomy that varies most through osteostracan phylogeny, as well as being the anatomical division most commonly preserved. Hence, it is the source of almost all characters that have been exploited in osteostracan phylogenetics (e.g. Janvier 1985; Sansom 2009*a*). Geometric morphometric characterization of the osteostracan headshield will, therefore, provide for a fair and effective comparison to existing categorical characterizations of osteostracan morphological variation, facilitating insights into the evolution of osteostracan morphological disparity and the equivalence of competing approaches to characterizing morphological disparity.

MATERIAL AND METHOD

Geometric morphometric analysis

Our study was confined principally to species of established genera, although a small number of taxa with resolved phylogenetic affinity have also been included, enabling comparison with the most complete osteostracan phylogeny (Sansom 2009*a*) (Fig. 2). The study was conducted at genus level, with each genus represented by a single specimen. The holotype specimen of the type species for the genus was used except where this specimen was poorly preserved or unavailable for characterization;

in such circumstances, another better-represented species was used. A total of 29 specimens were included in the geometric morphometric analysis (Fig. 2). Some specimens exhibit minor deformation, principally due to dorsoventral compression. No attempt was made to correct for deformation as this would lead to the inclusion of human error and preliminary studies have suggested that biological signal is still well preserved when deformation is minor (Angielczyk & Sheets 2007). Images of specimens for digitization were obtained from photographs of original specimens (see Ferrón *et al.* 2020, appendix S1, file S1). When complete specimens could be pieced together from counterparts, images of these counterparts
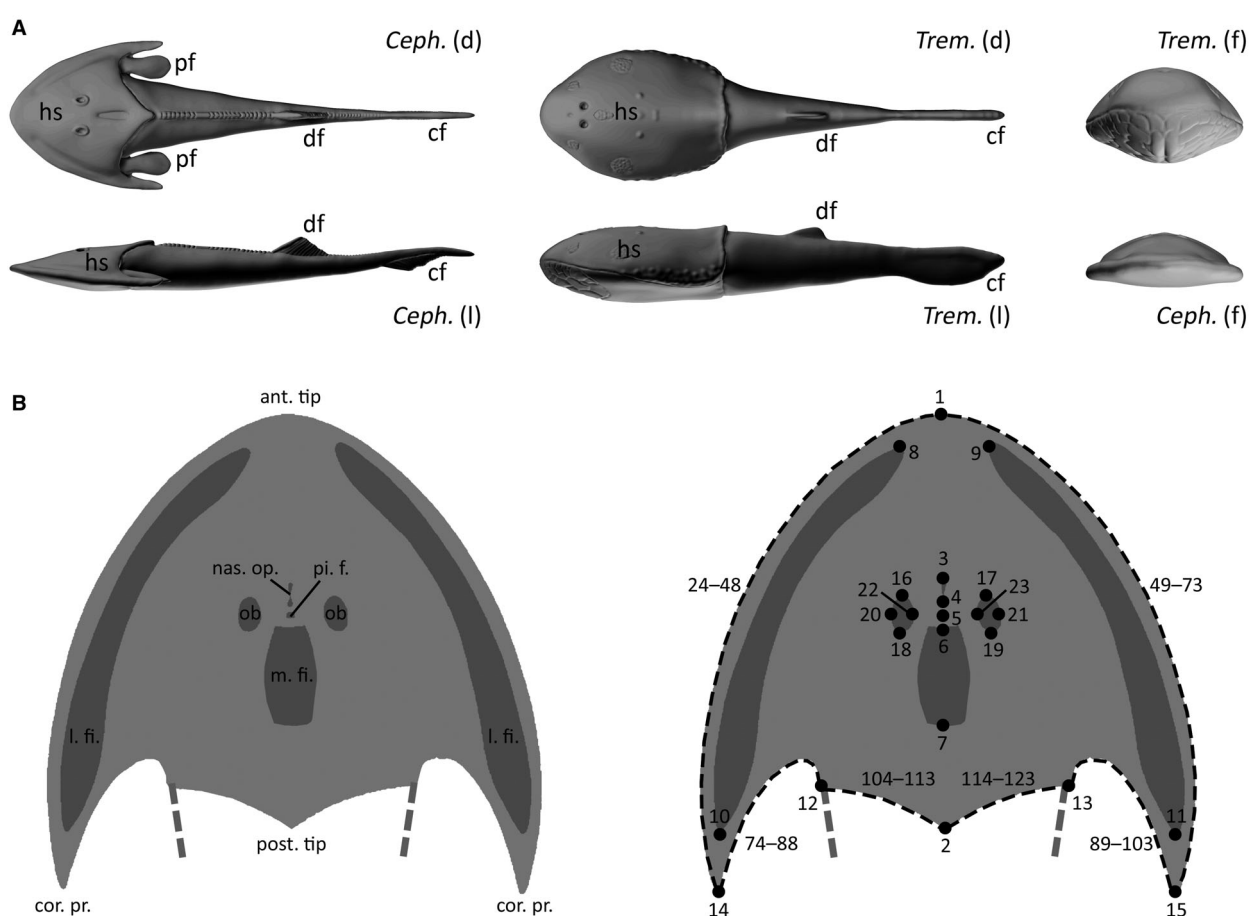
FIG. 1. A, drawings illustrating the general morphology of two distinct well-known osteostracans (*Cephalaspis* and *Tremataspis*) in dorsal (d), lateral (l) and frontal (f) views; position of the pectoral fins (pf), dorsal fin (df), caudal fin (cf) and the headshield (hs) is indicated. B, descriptive diagrams showing the terminology of commonly referred anatomical characters of osteostracan headshield (left) and the landmark configuration used in the geometric morphometric analysis (right). *Landmarks*: 1, anterior tip of the headshield or the rostral process (ant. tip.); 2, posterior tip of the headshield (post. tip.); 3–4, most anterior and most posterior medial points of the nasohypophysial opening (nas. op.) respectively; 5, pineal foramen (pi. f.); 6–7, most anterior and most posterior medial points of the median field (m. fi.) respectively; 8–9, most anterior points of the lateral fields (l. fi.); 10–11, most posterior points of the lateral fields (l. fi.); 12–13, points of connection between the headshield and the body; 14–15, most distal points of the cornual processes (cor. pr.); 16–23, most anterior, posterior, lateral and medial points of the eye orbits (ob.); 24–73, landmarks situated between Landmarks 1 and 14 and between Landmarks 1 and 15; 74–103, landmarks situated between Landmarks 14 and 12 and between Landmarks 15 and 13; 104–123, landmarks situated between Landmarks 12 and 2 and between Landmarks 13 and 2.

were superimposed (see Ferrón *et al.* 2020, appendix S1, for a list of sources and information on the nature of deformation of these specimens). A total of 123 landmarks of type I, II and III were digitized using TpsDig v.2.26 (Rohlf 2016*a*) (Fig. 1B). The choice of landmarks was (by definition) constrained by homology (Bookstein 1997), but also by preservation. Therefore, only frequently preserved characters were included as landmarks in the study. Landmarks of type III were equally interpolated along the specimen outlines in six separate open curves. The number of landmarks chosen to represent each curve reflects the relative length and complexity of each portion of the headshield and was determined visually. TpsRelw v.1.65 (Rohlf 2016*b*) was used to fit landmark coordinates of all specimens by generalized Procrustes superimposition to remove variation in rotational, scale and translational differences between specimens so that only geometric information was left.

Categorical characters

Our categorical characterization of osteostracan morphology is based on the cladistic data set of Sansom (2009*a*) which comprises 112 characters and 65 taxa (see Ferrón *et al.* 2020, appendix S1). We analysed these data in two ways: (1) the complete data set including cranial and postcranial characters; and (2) a subset of characters that relate to the cephalothoracic 'headshield', including external shape characters, but also neurocranial and histological features. Analyses of this subset of cranial characters facilitates a direct comparison with the analyses of the geometric morphometric data which are similarly limited to the headshield.

*Disparity quantification from categorical and morphometric data*

The disparity of osteostracans was explored from temporal, phylogenetic and ecological perspectives considering stratigraphic ranges, phylogenetic relationships, major taxa and habitat interpretations established in Sansom (2009*a*) and Sansom *et al.* (2015). Disparity was quantified from both categorical and continuous character data sets in two different ways: (1) the preordination distance, which is the average squared distance between taxa based on the original data; and (2) the ratio of generalized variance, which is the ratio of the taxa within a time bin to the entire data set (in this case the generalized variance is the determinant of the covariance matrix from the first three axes of the principal coordinate analysis).

Distance matrices were obtained from both categorical and geometric morphometric data sets (considering Gower and Euclidean distances, respectively) using the package `cluster` (Maechler *et al.* 2019) implemented in R (R Core Team 2019). Distance matrices were then subjected to principal coordinate analysis in the R package `ape` (Paradis & Schliep 2018). The same procedure was repeated with a categorical data subset considering only the 29 genera included in the continuous data analysis. This allowed us to determine how much of the total morphological disparity was represented by the subset of genera used in the morphometric analysis and also to assess whether discrete and continuous characterization of form yielded similar results. In parallel, Mantel tests were implemented using the R package `vegan` (Oksanen *et al.* 2013) to establish the strength and significance of linear correlations among the distance matrices derived from each of the data sets.

Morphospaces and phylomorphospaces were constructed using the R packages `phytools` (Revell 2012) and `ggplot2` (Wickham 2016). We followed two alternative approaches to creating phylomorphospaces based on post and preordination ancestral state estimation (OASE) (for a review, see Lloyd 2018). In the case of the preordination procedure, ancestral state reconstruction was achieved using stochastic character state mapping (Huelsenbeck *et al.* 2003) using the R packages `phytools` (Revell 2012) and `geomorph` (Adams & Otárola-Castillo 2013; Adams *et al.* 2019) for categorical and geometric morphometric data respectively. We used the phylogenetic tree from Sansom (2009*a*) after time-calibration, using the minimum branch length method (Laurin 2004), in the R package `paleotree` (Bapst 2012). Tip ages were established from stratigraphic ranges published in Sansom *et al.* (2015).

Often comparisons of disparity between groups are more valuable than characterizations of the distribution of taxa within morphospace. To assess whether the different data types (categorical observations vs continuous measurements) yield similar conclusions about the relative disparity of groups we randomly selected 9 taxa and 12 taxa (with no overlap) and calculated the ratio of disparity considering both the categorical and geometric morphometric data. We selected groups of this size to mirror the largest clades within the data set (Benneviaspida and Thyestida). This subsampling routine was repeated 10 000 times. If both data sets capture similar morphological signals, the relative disparity between the two samples should be the same or similar for both metrics and there should be agreement on the more disparate subsample. Our null model was based on a simulation approach, involving morphological data evolving under Brownian motion on a phylogeny, following Schaeffer *et al.* (2019). We first generated 1000 phylogenetic trees using the topology from Sansom (2009*a*), time-calibrated 500 times using both the equal (Brusatte *et al.* 2008) and minimum branch length (Laurin 2004) dating approaches
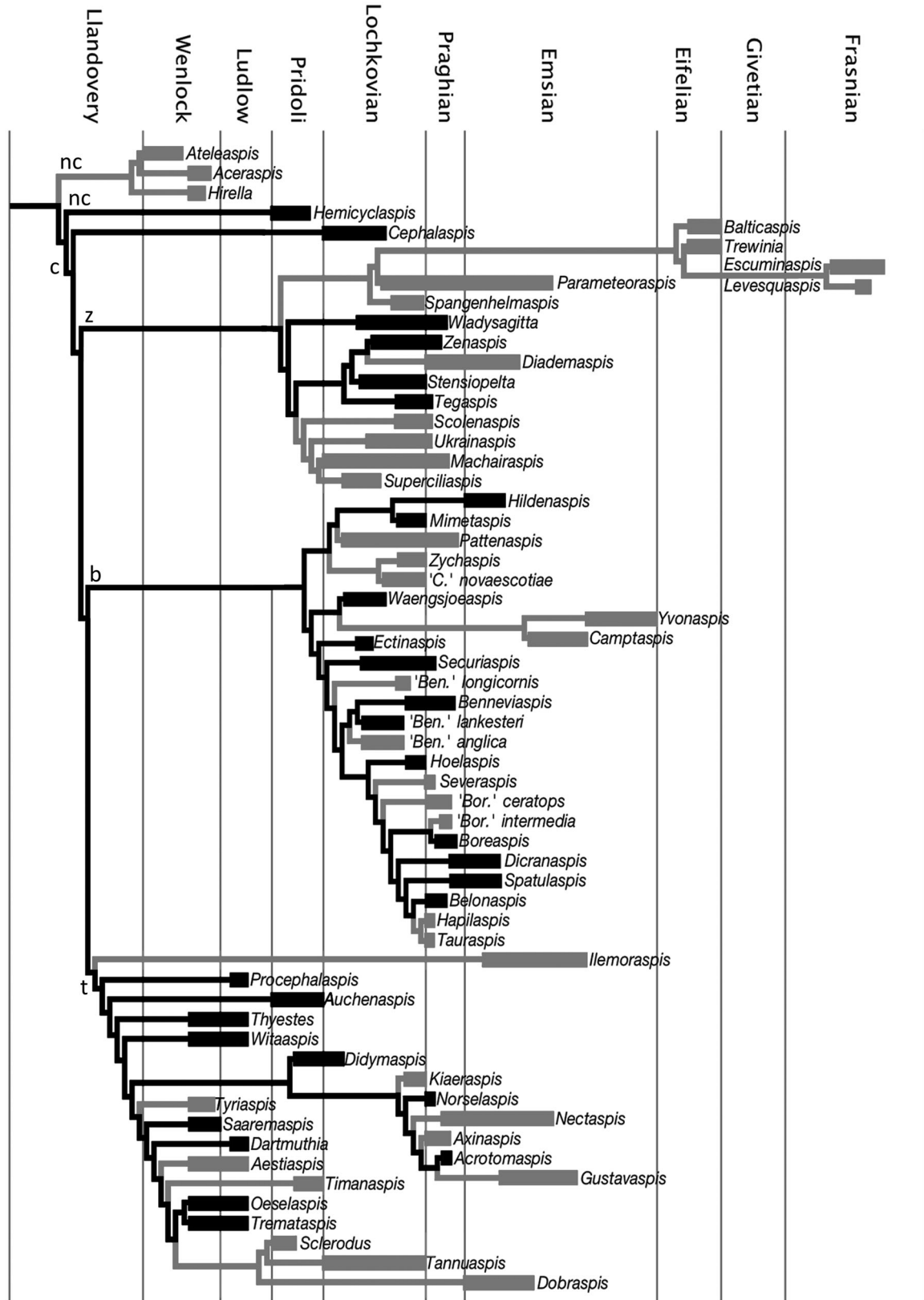
FIG. 2. Phylogenetic hypothesis of the Osteostraci (modified from Sansom *et al.* 2015), highlighting in black the species represented in the geometric morphometric analysis. *Major groups:* b, Benneviaspidida; c, Cornuata; nc, 'Non Cornuata'; t, Thyestiida; z, Zenaspida.

in `paleotree`. For each of these trees, tip ages were established by randomly sampling ages between each taxon's first and last stratigraphic occurrences. We then simulated both categorical and continuous morphology-like data on all of the 1000 dated phylogenies. We simulated discrete character data sets in the R package `dispRity` (Guillerme 2018) using the equal-rates (ER = Mk) model (Paradis & Schliep 2018). For the model parameters, we randomly sampled gamma distributions with shapes of 0.5, 1 or 2 and rates of 5, 10 or 20 (higher rates increased phylogenetic signal) following Schaeffer *et al.* (2019). We simulated geometric shape data, generating 29 'headshield-like' structures defined by 123 landmark coordinates for each tree in the R package `Evomorph` (Cabrera & Giri 2016). As inputs, we considered the consensus shape of the original analysis as the ancestral morphology and the Procrustes residuals from our original landmark configurations as co-variance data. The resulting categorical and continuous data sets were then subjected to similar subsampling routines to those described above and the results were compared with our empirical data. In this case, the disparity ratios were calculated on randomly selected groups as well as on the benneviaspidids and the thyestids of each simulated data set, which included 9 and 12 taxa, respectively.

Finally, correlation between disparity and taxonomic diversity over time was evaluated in PASW Statistics (http://www.spss.com.hk/statistics), considering the different data sets and disparity metrics; taxonomic diversity data were obtained from Sansom *et al.* (2015). Correlations on first-difference transformed data were also checked, to eliminate the role of autocorrelation.

RESULTS

Morphospace occupation and phylomorphospaces

Plots of taxa in the multivariate space generated from the categorical and geometric morphometric data sets as well as the percentage of total variance summarized by each axis are shown in Figure 3 and Table 1 respectively. Analysis of both the complete and subsampled categorical data sets recovered similar patterns of taxon clustering within ordination space (note that the mirroring of taxa on principal co-ordinates PCo1 and PCo2 is an artefact of the

arbitrary direction of ordination). Ordinations of categorical data sets tend to disperse variance explained across a large number of axes (Lloyd 2016) and, thus, considering just the first three axes may give an incomplete view of morphological disparity. However, K-means clustering recovered identical grouping of taxa using the first three (52.7% variance explained) and the first 15 axes (84.6% variance explained) of the ordination of categorical data. Therefore, even though considering just a subset of the axes can be problematic, it is unlikely to play a large role in the major structure and clustering of the current data set. Benneviaspidids and thyestiids, which cover most of the morphospace, occupy different regions showing only a small overlapping area. A number of zenaspidids are also relatively well separated in both ordinations based on the complete and subsampled categorical data sets. Non-cornuates (e.g. *Hemicyclaspis*), basal cornuates (e.g. *Cephalaspis*) and some generalized cornuate forms with uncertain phylogenetic affinities (e.g. *Pattenaspis*, *Hildenaspis*, *Mimetaspis*, *Zychaspis*) occupy smaller areas that exhibit significant overlap with those of other groups in the ordination based on the complete categorical data set, but separate better in the ordination based on the subsampled data set (mostly on PCo3). Freshwater and marine genera are restricted to different areas of the morphospace with brackish representatives in intermediate positions. This is probably due to the existence of a strong association between taxonomic affinity, morphology, and habitat in benneviaspidids and thyestiids, which are represented mostly by freshwater and marine forms respectively. Phylomorphospaces obtained following the preordination and postordination methods show similar patterns. In both cases, two strong branching trends are recognized, one within Benneviaspidida (along PCo2) which is characterized by an anterolateral extension of the cornua and increasing length of the rostrum (e.g. *Boreaspis*, *Spatulaspis*, *Hoelaspis*), and another within Thyestiida (along PCo3), characterized by a reduction and eventual loss of the cornuae (e.g. *Oeselaspis*, *Witaaspis*). The latter is better represented in the subsampled data set. Reversals are much more common among non-cornuates, basal and generalized cornuates, and Zenaspidida. The ordination based on the continuous character data set exhibits a similar pattern of taxon clustering to ordinations based on categorical data sets, but there appears to be a greater overlap when genera are grouped by both major taxa and habitats.

FIG. 3. Morphospace occupation and phylomorphospaces of osteostracans grouped by major taxa and habitats obtained from categorical and geometric morphometric approaches. Morphospaces and phylomorphospaces are represented by PCo1 (horizontal), PCo2 (upper vertical) and PCo3 (lower vertical) axes. Osteostracan phylogenetic relationships and headshield drawings modified from Sansom (2009a). Taxa: 1, *Hemicyclaspis*; 2, *Boreaspis*; 3, '*Benneviaspis*' *lankesteri*; 4, *Tauraspis*; 5, *Ectinaspis*; 6, *Spatulaspis*; 7, '*Benneviaspis*' *longicornis*; 8, *Hoelaspis*; 9, *Camptaspis*; 10, *Oeselaspis*; 11, *Yvonaspis*; 12, *Thyestes*; 13, *Didymaspis*; 14, *Acrotomaspis*; 15, *Norselaspis*; 16, *Witaaspis*; 17, *Procephalaspis*; 18, *Waengsjoeaspis*; 19, *Ilemoraspis*; 20, *Stensiopelta*; 21, *Trewinia*; 22, *Zenaspis*; 23, *Tegaspis*; 24, *Escuminaspis*; 25, *Spangenhelmaspis*; 26, *Pattenaspis*; 27, *Hildenaspis*; 28, *Cephalaspis*; 29, *Mimetaspis*.
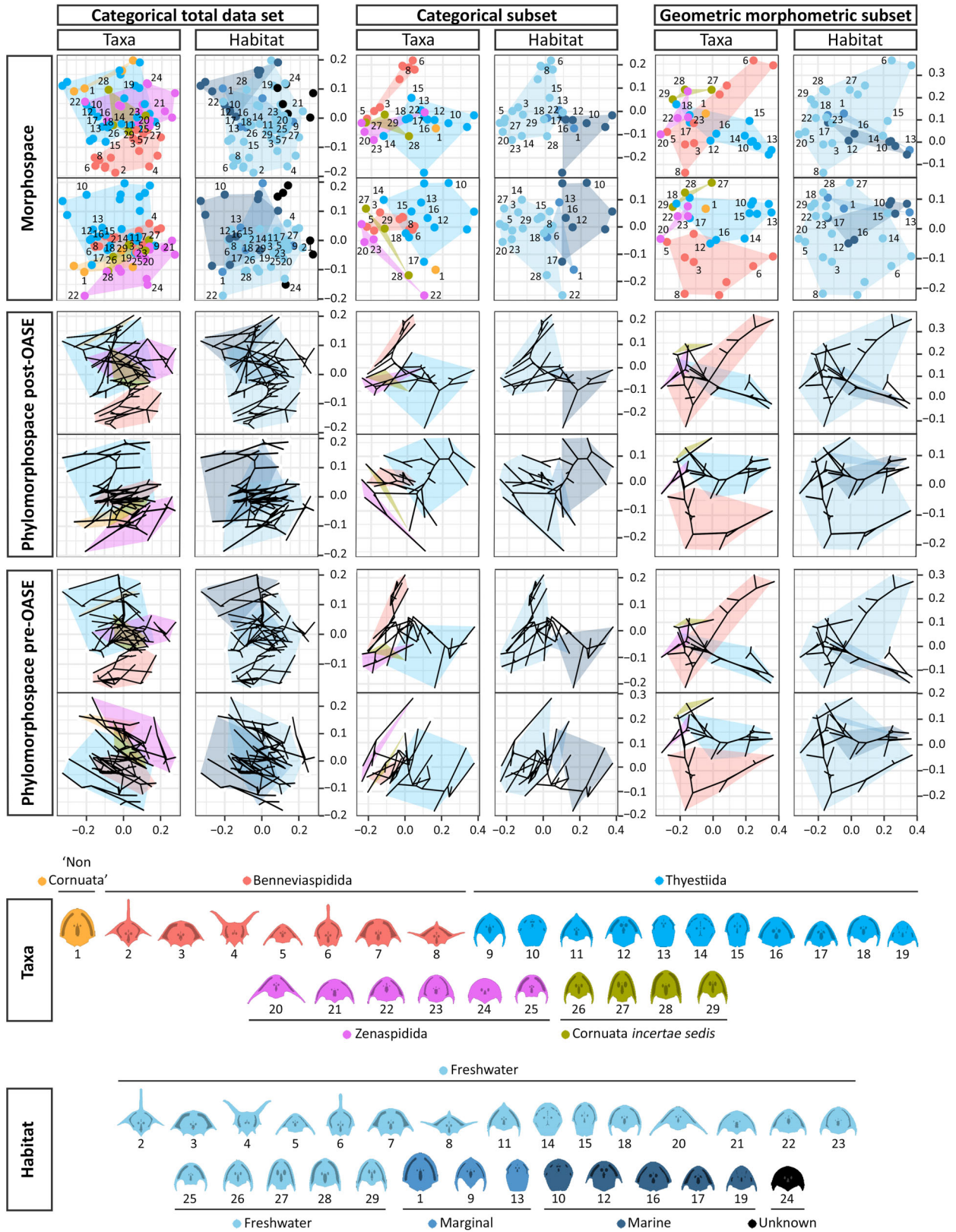
TABLE 1. Percentage variance summarized on the first three PCo axes for each of the three data sets.

Data set	Percentage variance summarized			
	Axis 1	Axis 2	Axis 3	Sum
Cladistic	19.11%	10.29%	7.58%	36.98%
Cladistic subset	27.45%	15.41%	9.85%	52.71%
Geometric morphometric	57.67%	18.99%	13.33%	89.99%

The phylomorphospaces obtained from the continuous data set following preordination and postordination methods are again extremely similar, revealing in both cases that there is a high degree of morphological convergence towards forms with well-developed caudally positioned cornual processes, most corresponding to freshwater genera of Benneviaspidida (e.g. *Ectinaspis*), Thyestiidae (e.g. *Waengsjoeaspis*), Zenaspidida (e.g. *Stensiopelta*) and basal and generalized cornuates (e.g. *Cephalaspis*, *Mimetaspis*) (towards negative values of PCo1). On the other hand, two similar branching trends to those recorded by categorical data can be recognized; a first one within Thyestiidae which is characterized by the reduction and loss of cornuae within Thyestiidae, as captured along PCo1 and PCo2, and a second one reflecting a notable development of cornual and rostral processes in Benneviaspidida, both captured along the PCo1 and PCo2.

A comparison of the preordination distances (Gower for categorical and Euclidean for geometric morphometric data sets) suggests the existence of strong correlation (Mantel state r = 0.3098, p = 0.001). These results contrast with K-means clustering on the two data sets, existing only around 58% agreement on cluster placement in the three-cluster solution. The three clusters found in the categorical data are composed of two groups that are environmentally and largely taxonomically consistent along with a more heterogeneous group. The morphometric data has one small group that is taxonomically and environmentally consistent and two larger mixed groups.

Temporal patterns of morphospace occupation and morphological disparity

Silurian and Devonian osteostracans occupy disparate regions within the morphospace in all three analyses showing important differences in taxonomic composition and habitats (Fig. 4). Wenlock and Ludlow age genera are mainly represented by marine thyestiids and non-cornuates. Pridolian genera comprise brackish thyestiids and non-cornuates that occupy intermediate positions between Wenlock–Ludlow and Devonian taxa. Devonian genera are mostly freshwater representatives of all major

osteostracan groups. The two measures of disparity we employ, preordination distance and generalized variance, characterize different temporal patterns within each data set (Fig. 5). In the complete cladistic data set, maximum disparity occurs at the beginning of the clade's history, although a second peak is revealed by the preordination distance metric in the Emsian (Fig. 5A–B). In the geometric morphometric and categorical subsets, the recorded temporal patterns of disparity appear congruent for each of the metrics, but differ strongly with those characterized by the complete categorical data set (compare Fig. 5A–B with Fig. 5C–D). Thus, the preordination distance metric reveals a peak in disparity in the Early Devonian (Lochkovian–Pragian) whereas generalized variance metric suggests comparatively stable values of disparity through the evolutionary history of the group. A positive correlation exists between taxonomic diversity and disparity, measured as generalized variance in the categorical subset. These results remain the same after detrending data for autocorrelation by applying first-difference transformation (Table 2).

Categorical versus continuous measurements of morphology in disparity analyses

Results derived from the subsampling routines, although displaying an important spread of the data, show positive trends (Fig. 6A). In fact, significant correlation is detected when considering results derived from both preordination distance (R = 0.186) and generalized variance disparity metrics (R = 0.065). In almost half of the sampled cases, the categorical and the geometric morphometric data sets disagree on which sample is most disparate (i.e. in 45.0% and 44.8% of the cases when considering preordination distance and generalized variance disparity metrics respectively), which is also the case for the empirical data from Benneviaspida and Thyestida (see asterisk on Fig. 6A). Positive trends, consistent with those obtained for the empirical data, are also detected in the simulated data when considering both the disparity ratios between randomly selected groups (preordination distance R = 0.233, 46.8% disagreement; generalized variance R = 0.025, 21.3% disagreement; Fig. 6B) and between Benneviaspida and Thyestida (preordination distance R = 0.299, 33.0% disagreement; generalized variance R = 0.037, 43.9% disagreement; Fig. 6B).

DISCUSSION

Osteostracan morphological disparity in space and time

Overall, both categorical and morphometric approaches to characterizing morphology yield similar patterns of
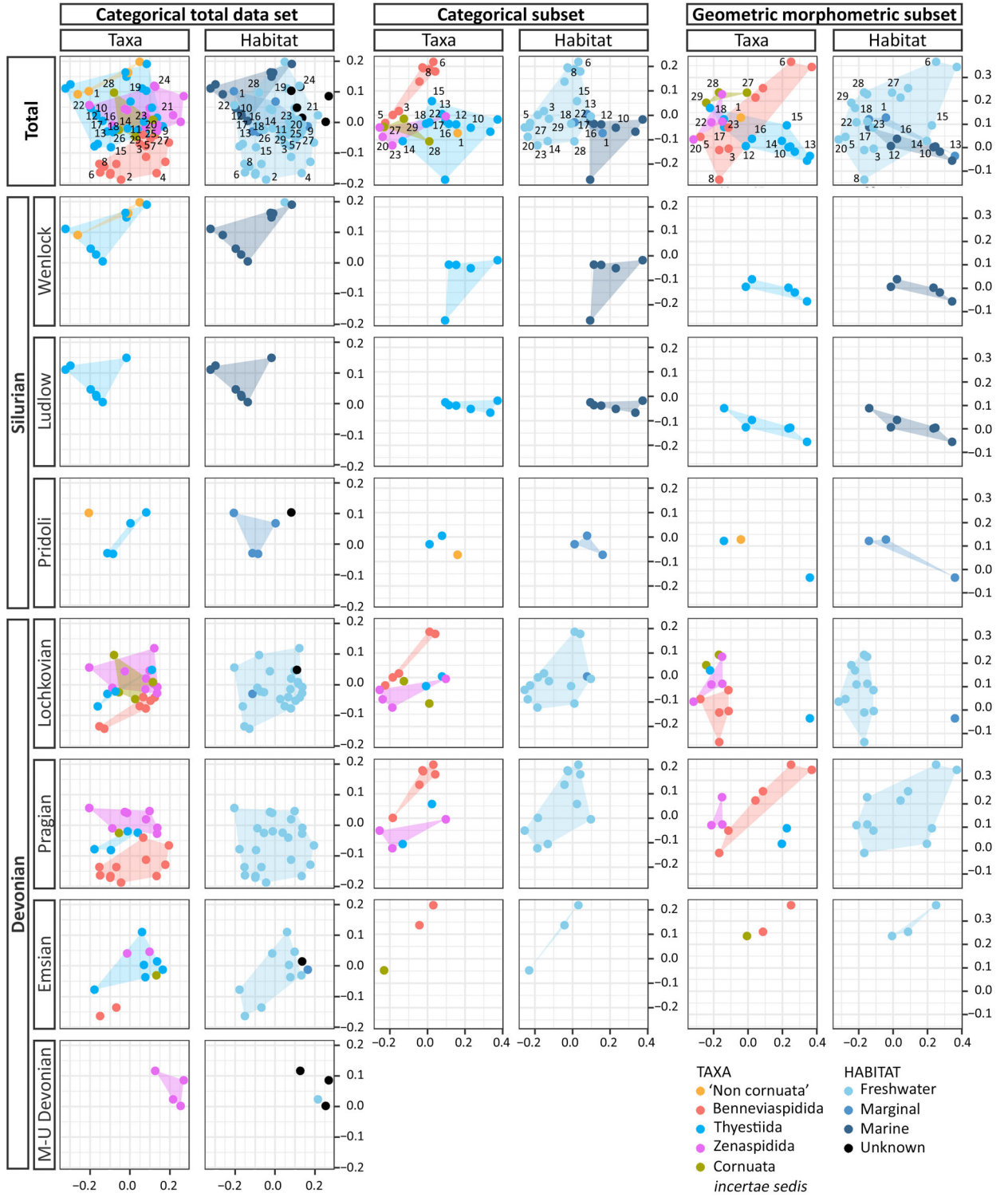
FIG. 4. Temporal patterns of morphospace occupation of osteostracans grouped by major taxa and habitats obtained from categorical and geometric morphometric approaches. Morphospaces are only represented by PCo1 (horizontal) and PCo2 (vertical) axes. Numbers in the morphospaces refer to taxa in Figure 3.
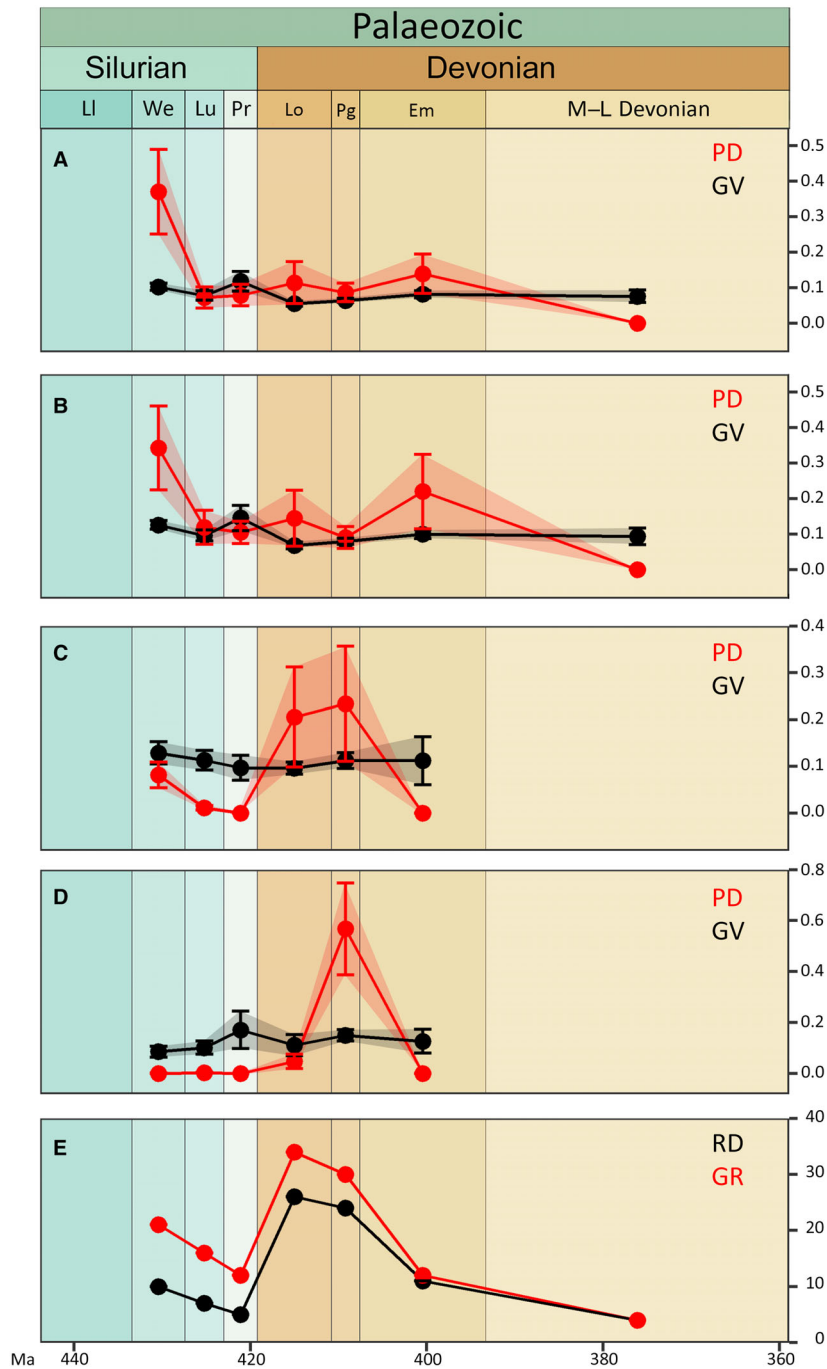
FIG. 5. A–B, categorical total data set including: A, both corporal and headshield characters; B, only headshield characters. C, categorical subset. D, geometric morphometric subset; disparity is calculated as the preordination distance (PD) and as the ratio of generalized variance (GV). E, taxonomic diversity of osteostracans through time based on raw data (RD) and including ghost ranges (GR) (data taken from Sansom *et al.* 2015). *Timescale*: Ll, Llandovery; We, Wenlock; Lu, Ludlow; Pr, Pridoli; Lo, Lochkovian; Pg, Pragian; Em, Emsian; M–L Devonian, Middle–Late Devonian. Colour online.

morphospace occupation and clustering (Fig. 3). Several forms that exhibit a generalized semi-circular headshield outline, such as non-cornuate genera (e.g. *Hemicyclaspis*) or some basal cornuates, appear very close to the mean

form in all analyses, which may be considered to be the ancestral state for the group given their early branching topology in osteostracan phylogeny (Sansom 2009*a*). This morphology has been related to a benthic mode of life in

TABLE 2. Correlation results between disparity and taxonomic diversity over time in osteostracans.

		Cladistic (GV)	Cladistic (PD)	Cladistic Sub. (GV)	Cladistic Sub. (PD)	Geom. Morph. (GV)	Geom. Morph. (PD)
Diversity (raw data)	Pearson correlation	0.075	−0.694	0.946	−0.235	0.621	−0.011
	p-value	0.872	0.084	0.004	0.654	0.189	0.983
Diversity (incl. ghost ranges)	Pearson correlation	0.260	−0.502	0.968	−0.181	0.545	−0.181
	p-value	0.574	0.251	0.002	0.808	0.264	0.731
Diversity* (raw data)	Pearson correlation	0.141	−0.745	0.946	0.064	0.419	−0.483
	p-value	0.789	0.089	0.015	0.918	0.408	0.410
Diversity* (incl. ghost ranges)	Pearson correlation	0.077	−0.750	0.955	0.054	0.436	−0.452
	p-value	0.884	0.086	0.011	0.931	0.387	0.445

*Analyses performed with detrended data.

Significance threshold: p= 0.05.

GV, ratio of generalized variance; PD, preordination distance.

which the headshield is oriented flat on the substrate and is able to withstand strong currents, whilst allowing the organism to remain agile (Bunker & Machin 1991). Phylomorphospaces reveal convergence on this headshield morphology within early-branching zenaspidids, the earliest-branching benneviaspidids and thyestiids, as well as other cornuate taxa, presumably reflecting the optimality, or rather the general effectiveness of one successful form. This is also reflected in the fact that this area of the morphospace is more densely occupied than most others. Two evolutionary trends separated in time are clearly recognizable in the phylomorphospaces, involving different major groups of osteostracans occupying disparate habitats (Figs 2, 3). The first such trend is characterized by the reduction and loss of the cornual processes in thyestiids inhabiting marine environments during the Silurian. This has previously been interpreted as an adaptation to a burrowing life habit in the more derived groups including tremataspidids and kiaeraspidids (Janvier & Lawson 1985). The second trend is characterized by the development of long cornual and/or rostral processes in freshwater benneviaspidids, during the Devonian. These structures have been the focus of competing functional interpretations in other early vertebrate groups where they have been interpreted as either locomotory adaptations to enhance lift generation or reduce drag (e.g. Mark-Kurik 1992; Novitskaya 2000; Moloshnikov 2001; Botella & Fariña 2008; Fletcher *et al.* 2014), or for predator deterrence (e.g. Janvier 1977), as housing for sensory organs (e.g. Voichyshyn 2006), a substrate anchor (e.g. Dineley 1976; Janvier 1985; Wells & Dorr 1985), or an adaptation to specific feeding habits (e.g. Tarlo 1961; Dineley 1994).

We employed both preordination and postordination approaches to phylomorphospace construction, the relative merits of which were considered by Lloyd (2018). Postordination approaches are readily and therefore commonly applied, but ancestral values are forced to be within the range of sampled tip values and may lead to an underestimation of convergence. In contrast, postordination approaches are more complex, requiring prior estimation of the characteristics of the hypothetical ancestors predicted by phylogenetic hypotheses, but have some advantages including allowing increased sample size, reconstruction of missing data, and avoiding the assumption that estimated ancestors must fall within the range of tip values. However, the prior inclusion of inferred ancestors in pairwise distances and the ensuing ordination can have the effect of inflating clade disparity, perhaps artefactually. Nevertheless, in this instance, our results indicate that both preordination and postordination-based approaches to estimating phylomorphospace recover similar overall patterns (Fig. 3). Conspicuously, large areas of morphospace characterized by both the categorical and continuous character data sets remain unoccupied, which presumably reflects morphological character combinations and shapes that are hydrodynamically or functionally inefficient, unexplored as a consequence of phylogenetic, developmental or structural constraints, unpreserved or perhaps unrealized because of insufficient evolutionary time.

Categorical and continuous character data sets convey different patterns of variation in the range of morphospace occupation, as do the different measures of disparity. Characterization of disparity with categorical data suggests that the maximum was established early in osteostracan evolution (Fig. 5A, B), consistent with the derived nature of some Silurian thyestiids (Figs 1, 3). This pattern is more significant when measuring disparity from preordination distances. However, capturing disparity with geometric morphometric data suggests that maximal variation was achieved later in osteostracan phylogeny, in the Pridolian or early Devonian (depending on the metric used). Overall, both data types suggest a post-Pragian decline in the morphological disparity of the Osteostraci before their Late Devonian extinction. This pattern could
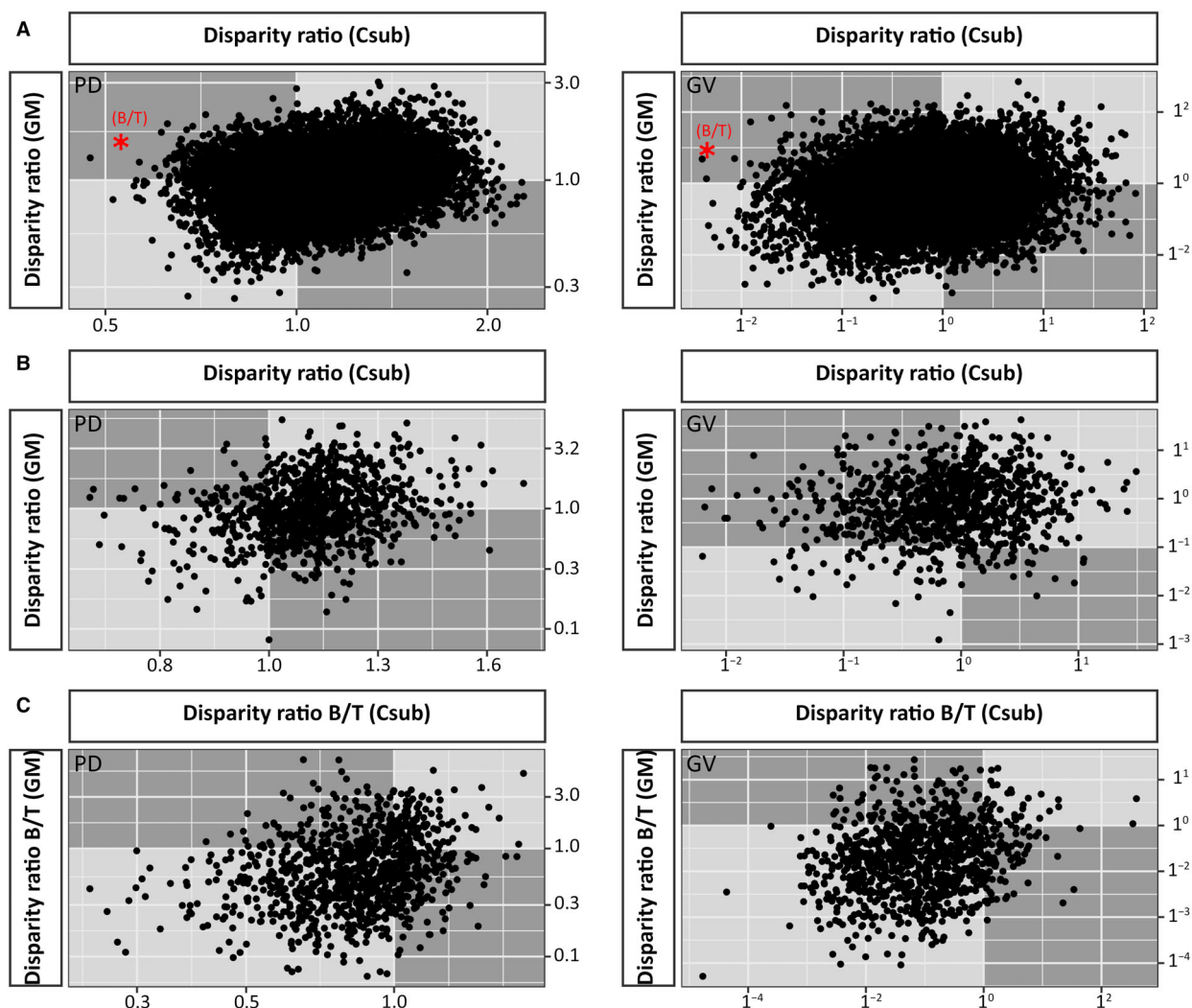
FIG. 6. Plots showing the disparity ratios derived from the categorical (Csub) and the geometric morphometric (GM) data sets in: A, empirical; B–C, modelled approaches. Disparity is calculated as both the preordination distance (PD) and as the ratio of generalized variance (GV). The disparity ratios are calculated on two randomly selected groups of 9 and 12 taxa in the empirical (A) and modelled (B) data, as well as on Benneviaspida and Thyestida (C) in the modelled data. Actual ratio of Benneviaspida and Thyestida (B/T) is denoted by an asterisk on the empirical data plots. Points on dark grey areas correspond to sampled cases where the categorical (Csub) and the geometric morphometric (GM) data sets disagree on which group is more disparate. Colour online.

be interpreted literally, congruent with Janvier & Newman's (2005) hypothesis on the decline of ostracoderm groups, reflecting an increase of ecological restriction during the Middle and Late Devonian period imposed by changes on marginal marine and freshwater environments and/or food resources, combined with limited dispersal capability (Sansom 2009b). However, flux in the diversity of osteostracans throughout this interval has been interpreted to reflect facies shifts in the rock record (Sansom et al. 2015). Indeed, a Pearson correlation test finds evidence of a significant correlation between diversity (raw standing diversity or corrected for ghost lineages) and some metrics of morphological disparity based on the

categorical data (Table 2). It is most likely that parallel changes in diversity and disparity reflect the environmentally non-uniform fossil record (Sansom et al. 2015) and the adaptation of osteostracans to the environments in which they lived.

Categorical versus continuous measurements of morphology in disparity analyses

Categorical observations and continuous measurements are non-mutually exclusive approaches to the characterization of morphology in analyses of the evolution of

morphological disparity. In attempting to derive nomothetic insights in this sense, to address hypotheses such as the universality of maximal initial disparity (Hughes *et al.* 2012), or the relationship between disparity and diversification (Foote 1993), it is important that the aspects of morphological variation summarized in disparity analyses are in some sense equivalent (Hetherington *et al.* 2015). Given the nature and scope of the phenotypic features that these approaches can characterize, it could be expected that categorical characters, which usually sample from across the breadth of phenotype, capture different aspects of morphological disparity to geometric morphometrics, which is usually focused on a subsample of overall phenotype as a proxy for the whole. Furthermore, categorical data sets which, as here, are mostly repurposed cladistic data sets, have a tendency to eschew invariant, convergent and autapomorphic characteristics. Continuous character data sets should not suffer these same ascertainment biases but capturing morphological variation through continuous characters has its own limitations. Principally, these are related to the need to limit landmarks to homologous structures present in all taxa and, therefore, the challenge of capturing neomorphic structures and losses.

Despite this, the majority of benchmarking studies have found that categorical and continuous characters capture similar patterns of morphological disparity (Villier & Eble 2004; Anderson & Friedman 2012; Foth *et al.* 2012; Hetherington *et al.* 2015; Hopkins 2017; Romano *et al.* 2017; Schaeffer *et al.* 2019). However, few of these studies (Romano *et al.* 2017; Schaeffer *et al.* 2019) compare categorical and continuous characters based on the same anatomical structures; most characterize disparity based on different body regions. For example, Villier & Eble (2004) incorporated data sets for echinoids in which 80% of the categorical characters had no equivalence in their morphometric characters; Hetherington *et al.* (2015) compared geometric morphometric data on the skull surface and categorical data on internal cranial structures in caecilian amphibians; Foth *et al.* (2012) compared pterosaur disparity based on cranial morphometrics and categorical characters from the entire skeleton (Prentice *et al.* 2011; Butler *et al.* 2012), and limb measurements and ratios (Dyke *et al.* 2009). Different anatomical divisions may well co-evolve, exhibiting equivalent patterns of morphological disparity through time, supporting the practice of using proxy data sets as representative of the whole (Hopkins 2017). However, this should not be a null expectation not least since anatomical partitions of categorical data commonly generate different phylogenetic hypotheses (Mounce *et al.* 2016; Brinkworth *et al.* 2019; Li *et al.* 2019) and, therefore, they should be expected to generate equally different characterization of disparity.

Our results suggest that the patterns of osteostracan morphological disparity captured by categorical and continuous characters are correlated when cladistic geometric morphometric data sets are based on equivalent taxa and anatomical components (Fig. 6). This result is not surprising since it is possible to express continuous measurements as categorical states (e.g. Thiele 1993) and, therefore, it is possible to compile categorical data sets that are equivalent to continuous character data sets, though they have less information content since they objectively and explicitly summarize continuous variation. The correlation between the empirical and simulated data implies a strong phylogenetic signal within the morphological data. However, the empirical data exhibit greater disagreement on the relative disparity within the groups. Despite this apparent equivalence in the results derived from categorical and geometric morphometric subsets, the complete categorical data set (which samples trunk as well as cranial characters) provides a very different perspective on the evolution of morphological diversity within osteostracans both in terms of the timing of peak disparity within the clade and the relative disparity between subclades. Therefore, while both data types appear to capture similar patterns of disparity, the nature of morphometric data (highly dependent on preservation and limited to areas with recognizable homologous structures) could lead in practice to very partial results in some cases and, ultimately, to conclusions more strongly biased by the vagaries of preservation.

In a very real sense, it does not matter that these two approaches to summarizing morphology result in different perceptions of morphological variation. They provide different perspectives on the same phenomenon and that, based on the nature of the data, enriches understanding of the evolution of morphology within species and clades. Further, these alternative approaches may be better suited to different questions. For example, characterization of disparity in terms of shape variation may be of greater importance when exploring the constraining role of the aquatic environment in which osteotracans lived. Meanwhile, categorical characters may be better suited to capturing the overall disparity of osteostracan phenotype. In this sense, neither approach is necessarily superior. Nevertheless, if the objective is to capture the evolution of phenotypic disparity in general, both within and between lineages, it appears clear that the while a dense sampling of phenotype is advantageous, diverse sampling of phenotype is paramount. In attempting to integrate over the vagaries of variable fossil preservation, this may be achieved most effectively by summarizing phenotype using categorical characters.

DATA ARCHIVING STATEMENT

Data for this study (including specimen list, R scripts and images) are available in the Dryad Digital Repository: https://doi.org/10.5061/dryad.rbnzs7h7f

Editor. Lionel Cavin

REFERENCES

ADAMS, D. C. and OTÁROLA-CASTILLO, E. 2013. Geomorph: an R package for the collection and analysis of geometric morphometric shape data. *Methods in Ecology & Evolution*, **4**, 393–399.

—— COLLYER, M. L. and KALIONTZOPOULOU, A. 2019. geomorph: geometric morphometric analyses of 2D/4D landmark data. Version 3.1.0. https://cran.r-project.org/web/packages/geomorph/

AFANASSIEVA, O. B. 1992. Some peculiarities of osteostracan ecology. 61–70. *In* MARK-KURIK, E. (ed.) *Fossil fishes as living animals*. Academy of Sciences of Estonia, Tallin, 299 pp.

ANDERSON, P. S. and FRIEDMAN, M. 2012. Using cladistic characters to predict functional variety: experiments using early gnathostomes. *Journal of Vertebrate Paleontology*, **32**, 1254–1270.

ANGIELCZYK, K. D. and SHEETS, H. D. 2007. Investigation of simulated tectonic deformation in fossils using geometric morphometrics. *Paleobiology*, **33**, 125–148.

BAPST, D. W. 2012. paleotree: an R package for paleontological and phylogenetic analyses of evolution. *Methods in Ecology & Evolution*, **3**, 803–807.

BARIŞ, M. V. G. 2017. The morphological diversity of coccolithophores across environments, geographic space, and geologic time. Unpublished PhD thesis, The University of Chicago, 140 pp.

BELLES-ISLES, M. 1987. La nage et l'hydrodynamique de deux Agnathes du Paléozoïque: *Alaspis macrotuberculata* et *Pteraspis rostrata*. *Neues Jahrbuch für Geologie und Paläontologie, Ahbhandlungen*, **175**, 347–376.

BOOKSTEIN, F. L. 1997. *Morphometric tools for landmark data: geometry and biology*. Cambridge University Press, 435 pp.

BOTELLA, H. and FARIÑA, R. A. 2008. Flow pattern around the rigid cephalic shield of the Devonian agnathan *Errivaspis waynensis* (Pteraspidiformes: Heterostraci). *Palaeontology*, **51**, 1141–1150.

BRIGGS, D. E., FORTEY, R. A. and WILLS, M. A. 1992. Morphological disparity in the Cambrian. *Science*, **256**, 1670–1673.

BRINKWORTH, A. R., SANSOM, R. and WILLS, M. A. 2019. Phylogenetic incongruence and homoplasy in the appendages and bodies of arthropods: why broad character sampling is best. *Zoological Journal of the Linnean Society*, **187**, 100–116.

BRUSATTE, S. L., BENTON, M. J., RUTA, M. and LLOYD, G. T. 2008. The first 50 Myr of dinosaur evolution: macroevolutionary pattern and morphological disparity. *Biology Letters*, **4**, 733–736.

BUNKER, S. J. and MACHIN, K. E. 1991. The hydrodynamics of cephalaspids. *Society of Experimental Biology Seminar Series*, **36**, 113–129.

BUTLER, R. J., BRUSATTE, S. L., ANDRES, B. and BENSON, R. B. J. 2012. How do geological sampling biases affect studies of morphological evolution in deep time? A case study of pterosaur (Reptilia: Archosauria) disparity. *Evolution*, **66**, 147–162.

CABRERA, J. M. and GIRI, F. 2016. Evomorph: evolutionary morphometric simulation. https://cran.r-project.org/web/packages/Evomorph/index.html

CIAMPAGLIO, C. N., KEMP, M. and McSHEA, D. W. 2001. Detecting changes in morphospace occupation patterns in the fossil record: characterization and analysis of measures of disparity. *Paleobiology*, **27**, 695–715.

DAVIES, B. E. 2009. An experimental morphological investigation into the hydrodynamics and locomotion of the Palaeozoic jawless vertebrates *Poraspis*, *Errivaspis* and *Ateleaspis*. Unpublished Masters thesis, University of Leicester, 123 pp.

DELINE, B. 2009. The effects of rarity and abundance distributions on measurements of local morphological disparity. *Paleobiology*, **35**, 175–189.

—— and AUSICH, W. I. 2017. Character selection and the quantification of morphological disparity. *Paleobiology*, **43**, 68–84.

—— —— and BRETT C. E. 2012. Comparing taxonomic and geographic scales in the morphologic disparity of Ordovician through Early Silurian Laurentian crinoids. *Paleobiology*, **38**, 538–553.

DENISON, R. H. 1951. Evolution and classification of the Osteostraci. *Fieldiana Geology*, **11**, 157–196.

DINELEY, D. L. 1976. New species of *Ctenaspis* (Ostracodermi) from the Devonian of arctic Canada. 26–43. *In* CHURCHER, C. S. (ed.) *Essays on Palaeontology in Honour of Loris Shano Russell*. Royal Ontario Museum, Life Sciences, Miscellaneous Publications.

—— 1994. Cephalaspids from the Lower Devonian of Prince of Wales Island, Canada. *Palaeontology*, **37**, 61–70.

DONOGHUE, P. C. J. and KEATING, J. N. 2014. Early vertebrate evolution. *Palaeontology*, **57**, 879–893.

DYKE, G. J., McGOWAN, A. J., NUDDS, R. L. and SMITH, D. 2009. The shape of pterosaur evolution: evidence from the fossil record. *Journal of evolutionary biology*, **22**, 890–898.

FERRÓN, H. G., GREENWOOD, J. M., DELINE, B., MARTÍNEZ-PÉREZ, C., BOTELLA, H., SANSOM, R.

S., RUTA, M. and DONOGHUE, P. C. J. 2020. Data from: Categorical versus geometric morphometric approaches to characterizing the evolution of morphological disparity in Osteostraci (Vertebrata, stem Gnathostomata). *Dryad Digital Repository*. https://doi.org/10.5061/dryad.rbnzs7h7f

FLETCHER, T., ALTRINGHAM, J., PEAKALL, J., WIGNALL, P. and DORRELL, R. 2014. Hydrodynamics of fossil fishes. *Proceedings of the Royal Society B*, **281**, 20140703.

FOOTE, M. 1993. Contributions of individual taxa to overall morphological disparity. *Paleobiology*, **19**, 403–419.

—— 1997. The evolution of morphological diversity. *Annual Review of Ecology & Systematics*, **28**, 129–152.

FOTH, C., BRUSATTE, S. L. and 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**, 904–915.

GUILLERME, T. 2018. dispRity: a modular R package for measuring disparity. *Methods in Ecology & Evolution*, **9**, 1755–1763.

HETHERINGTON, A. J., SHERRATT, E., RUTA, M., WILKINSON, M., DELINE, B. and DONOGHUE, P. C. J. 2015. Do cladistic and morphometric data capture common patterns of morphological disparity? *Palaeontology*, **58**, 393–399.

HOPKINS, M. J. 2014. The environmental structure of trilobite morphological disparity. *Paleobiology*, **40**, 352–373.

—— 2017. How well does a part represent the whole? A comparison of cranidial shape evolution with exoskeletal character evolution in the trilobite family Pterocephaliidae. *Palaeontology*, **60**, 309–318.

HUELSENBECK, J. P., NIELSEN, R., BOLLBACK, J. P. and SCHULTZ, T. 2003. Stochastic mapping of morphological characters. *Systematic Biology*, **52**, 131–158.

HUGHES, M., GERBER, S. and WILLS, M. A. 2012. Clades reach highest morphological disparity early in their evolution. *Proceedings of the National Academy of Sciences*, **110**, 13875–13879.

JANVIER, P. 1977. Contribution à la connaissance de la systématique du genre *Boreaspis*, Stensiö (Agnatha, Cephalaspidomorphi, Osteostraci) du Dévonien inférieur du Spitsberg. *Annales de Paléontologie*, **63**, 1–32.

—— 1985. *Les Céphalaspides du Spitsberg (Anatomie - phylogénie et systématique des ostéostracés siluro-devoniens. Révisions des ostéostracés de la Formation de Wood Bay - Dévonien inférieur du Spitsberg)*. Cahiers de Paléontologie, **7**, Muséum national d'Histoire naturelle, Paris, 176 pp.

—— 1996. *Early vertebrates*. Oxford University Press, 408 pp.

—— and LAWSON, J. D. 1985. Environmental framework of the diversification of the Osteostraci during the Silurian and Devonian. *Philosophical Transactions of the Royal Society B*, **309**, 259–272.

—— and NEWMAN, M. J. 2005. On *Cephalaspis magnifica* Traquair, 1893, from the Middle Devonian of Scotland, and the relationships of the last osteostracans. *Transactions of the Royal Society of Edinburgh: Earth Sciences*, **95**, 511–525.

LAURIN, M. 2004. The evolution of body size, Cope's rule and the origin of amniotes. *Systematic Biology*, **53**, 594–622.

LI, Y., RUTA, M., WILLS, M. A. and FOSTER, P. 2019. Craniodental and postcranial characters of non-avian Dinosauria often imply different trees. *Systematic Biology*, syz077.

LLOYD, G. T. 2016. Estimating morphological diversity and tempo with discrete carácter-taxon matrices: implementation, challenges, progress, and future directions. *Biological Journal of the Linnean Society*, **118**, 131–151.

—— 2018. Journeys through discrete-character morphospace: synthesizing phylogeny, tempo, and disparity. *Palaeontology*, **61**, 637–645.

MAECHLER, M., ROUSSEEUW, P., STRUYF, A., HUBERT, M. and HORNIK, K. 2019. cluster: "finding groups in data": cluster analysis extended Rousseeuw et al. Version 2.1.0. https://cran.r-project.org/web/packages/cluster/

MARK-KURIK, E. 1992. Functional aspects of the armour in the early vertebrates. 107–115. *In* MARK-KURIK, E. (ed.) *Fossil fishes as living animals*. Academy of Sciences of Estonia, 299 pp.

MOLOSHNIKOV, S. V. 2001. New data on *Pycnosteus palaeformis* Preobrazhensky (Heterostraci, Psammosteiformes) from the Aruküla Deposits. *Paleontological Journal*, **35**, 410–414.

MONGIARDINO KOCH, N., CECCARELLI, F. S., OJANGUREN-AFFILASTRO, A. A. and RAMIREZ, M. J. 2017. Discrete and morphometric traits reveal contrasting patterns and processes in the macroevolutionary history of a clade of scorpions. *Journal of Evolutionary Biology*, **30**, 814–825.

MORRISSEY, L. B., BRADDY, S. J., BENNETT, J. P., MARRIOTT, S. B. and TARRANT, P. R. 2004. Fish trails from the lower Old Red Sandstone of Tredomen Quarry, Powys, southeast Wales. *Geological Journal*, **39**, 337–358.

MOUNCE, R. C., SANSOM, R. and WILLS, M. A. 2016. Sampling diverse characters improves phylogenies: Craniodental and postcranial characters of vertebrates often imply different trees. *Evolution*, **70**, 666–686.

NOVITSKAYA, L. I. 2000. Adaptation for swimming in the external morphology and skeleton of early vertebrates (Agnatha: Heterostraci). *Paleontological Journal*, **34**, 583–592.

OKSANEN, J., BLANCHET, F. G., KINDT, R., LEGENDRE, P., MINCHIN, P. R., O'HARA, R. B., SIMPSON, G. L., SOLYMOS, P., STEVENS, M. H. H. and WAGNER, H. 2013. vegan: community ecology package. Version 2.0-7. https://cran.r-project.org/web/packages/vegan/

PARADIS, E. and SCHLIEP, K. 2018. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, **35**, 526–528.

PRENTICE, K. C., RUTA, M. and BENTON, M. J. 2011. Evolution of morphological disparity in pterosaurs. *Journal of Systematic Palaeontology*, **9**, 337–353.

R CORE TEAM. 2019. R: a language and environment for statistical computing. Version 3.6.2. R Foundation for Statistical Computing.

REVELL, L. J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology & Evolution*, **3**, 217–223.

ROHLF, F. J. 2016a. *TpsDig v. 2.26*. Department of Ecology & Evolution, State University of New York at Stony Brook. http://life.bio.sunysb.edu/morph

—— 2016*b*. *TpsRelw v. 1.16*. Department of Ecology & Evolution, State University of New York at Stony Brook. http://life.bio.sunysb.edu/morph

ROMANO, M., BROCKLEHURST, N. and FRÖBISCH, J. 2017. Discrete and continuous character-based disparity analyses converge to the same macroevolutionary signal: a case study from captorhinids. *Scientific reports*, **7**, 1–9.

SANSOM, R. S. 2008. The origin and early evolution of the Osteostraci (Vertebrata): a phylogeny for the Thyestiida. *Journal of Systematic Palaeontology*, **6**, 317–332.

—— 2009*a*. Phylogeny, classification and character polarity of the Osteostraci (Vertebrata). *Journal of Systematic Palaeontology*, **7**, 95–115.

—— 2009*b*. Endemicity and palaeobiogeography of Osteostraci and Galeaspida: a test of scenarios of gnathostome evolution. *Palaeontology*, **52**, 1257–1273.

—— RANDLE, E. and DONOGHUE, P. C. 2015. Discriminating signal from noise in the fossil record of early vertebrates reveals cryptic evolutionary history. *Proceedings of the Royal Society B*, **282**, 20142245.

SCHAEFFER, J., BENTON, M. J., RAYFIELD, E. J. and STUBBS, T. J. 2019. Morphological disparity in theropod jaws and the application of discrete characters and geometric morphometrics. *Palaeontology*, **63**, 283–299.

STAYTON, C. T. and RUTA, M. 2006. Geometric morphometrics of the skull roof of stereospondyls (Amphibia: Temnospondyli). *Palaeontology*, **49**, 307–337.

TARLO, L. 1961. *Rhinopteraspis cornubica* (McCoy), with notes on the classification and evolution of the pteraspids. *Acta Palaeontologica Polonica*, **6**, 367–400.

THIELE, K. 1993. The holy grail of the perfect character: the cladistic treatment of morphometric data. *Cladistics*, **9**, 275–304.

VILLIER, L. and EBLE, G. J. 2004. Assessing the robustness of disparity estimates: the impact of morphometric scheme, temporal scale, and taxonomic level in spatangoid echinoids. *Paleobiology*, **30**, 652–665.

VOICHYSHYN, V. 2006. New osteostracans from the Lower Devonian terrigenous deposits of Podolia, Ukraine. *Acta Palaeontologica Polonica*, **51**, 131–142.

WEBSTER, M. and HUGHES, N. C. 1999. Compaction-related deformation in Cambrian Olenelloid trilobites and its implications for fossil morphometry. *Journal of Paleontology*, **73**, 355–371.

WELLS, N. A. and DORR, J. A. 1985. Form and function of the fish *Bothriolepis* (Devonian: Placodermi, Antiarchi): The first terrestrial animal. *Michigan Academician*, **17**, 167–173.

WICKHAM, H. 2016. *ggplot2: elegant graphics for data analysis*. 2nd edn. Springer, 213 pp.

WILLS, M. A. 1998. Cambrian and recent disparity: the picture from priapulids. *Paleobiology*, **24**, 177–199.

—— BRIGGS, D. E. and FORTEY, R. A. 1994. Disparity as an evolutionary index: a comparison of Cambrian and Recent arthropods. *Paleobiology*, **20**, 93–130.