Palaeontology The Palaeontological Association www.palass.org

[Palaeontology, 2023, e12634]

Testing the success of palaeontological methods in the delimitation of clam shrimp (Crustacea, Branchiopoda) on extant species

by MANJA HETHKE[1,*], KAI HARTMANN[1], MATTHIAS ALBERTI[2], THERESA KUTZNER[1] *and* MARTIN SCHWENTNER[3]

[1]Fachbereich Geowissenschaften, Institut für Geologische Wissenschaften, Fachrichtung Paläontologie, Freie Universität Berlin, D-12249 Berlin, Germany; manja.hethke@fu-berlin.de

[2]State Key Laboratory for Mineral Deposits Research, School of Earth Sciences & Engineering, Centre for Research & Education on Biological Evolution & Environment, & Frontiers Science Center for Critical Earth Material Cycling, Nanjing University, Nanjing, 210023 Jiangsu China

[3]Naturhistorisches Museum Wien, AT-1010 Wien, Austria; martin.schwentner@nhm-wien.ac.at

*Corresponding author

Typescript received 21 March 2022; accepted in revised form 20 September 2022

Abstract: Fossil spinicaudatan taxonomy heavily relies on carapace features (size, shape, ornamentation) and palaeontologists have greatly refined methods to study and describe carapace variability. Whether carapace features alone are sufficient for distinguishing between species of a single genus has remained untested. In our study, we tested common palaeontological methods on 481 individuals of the extant Australian genus *Ozestheria* that have been previously assigned to ten species based on genetic analysis. All species are morphologically distinct based on geometric morphometrics ($p \leq 0.001$), but they occupy overlapping regions in *Ozestheria* morphospace. Linear discriminant analysis of Fourier shape coefficients reaches a mean model performance of 93.8% correctly classified individuals over all possible 45 pairwise species comparisons. This can be further increased by combining the size and shape datasets. Nine of the ten examined species are clearly sexually dimorphic but male and female morphologies strongly overlap within species with little influence on model performance. Ornamentation is commonly species-diagnostic; seven ornamentation types are distinguished of which six are species-specific while one is shared by four species. A transformation of main ornamental features (e.g. from punctate to smooth) can occur among closely related species suggesting short evolutionary timescales. Our overall results support the taxonomic value of carapace features, which should also receive greater attention in the taxonomy of extant species. The extensive variation in carapace shape and ornamentation is noteworthy and several species would probably have been assigned to different genera or families if these had been fossils, bearing implications for the systematics of fossil Spinicaudata.

Key words: clam-shrimp carapace, morphometrics, species delimitation, Spinicaudata, *Ozestheria*, Australia.

SPINICAUDATA are branchiopod crustaceans whose body is enclosed in a bivalved carapace that resembles a clam shell, which led to their common name, 'clam shrimp'. With approximately 30 fossil families (Astrop & Hegna 2015), spinicaudatan fossil diversity appears to be much larger than their extant diversity, with only four currently recognized families that contain 194 species in 16 genera (Rogers 2020). The spinicaudatan carapace exhibits a high preservation potential in the fossil record, as many species build in calcium phosphate and/or calcium carbonate biominerals, and the carapace is resistant to biostratinomic processes (Stigall *et al.* 2008; Astrop *et al.* 2015; Hegna *et al.* 2020). Thus, it forms a key role in fossil species discrimination, which mainly focuses on carapace traits, such as growth-band ornamentation and ratios of linear variables (e.g. Raymond 1946; Tasch 1969; Zhang *et al.* 1976; Kozur & Seidel 1983; Li & Batten 2004*a*, 2004*b*, 2005; Gallego 2010; Scholze & Schneider 2015; Hethke *et al.* 2018; Gallego *et al.* 2020; Hegna & Astrop 2020; Li & Wu 2021). Conversely, extant species are predominately delimited by soft-part features (e.g. Daday de Deés 1915; Rogers 2020) with the carapace being considered in less detail compared to palaeontological studies (e.g. Baird 1849; Grube 1865; Daday de Deés 1915). Soft parts are, in turn, rarely preserved in the fossil record (see Orr & Briggs 1999, for an exception), rendering it difficult to compare extant and fossil species.

doi: 10.1111/pala.12634

Molecular genetic markers have shown that several extant species are in fact complexes of multiple, morphologically highly similar species (Schwentner *et al.* 2011, 2014, 2015*a*) indicating a high cryptic biodiversity. Once individuals were genetically assigned to species, minute morphological differences could often be identified (Schwentner *et al.* 2012, 2014; Tippelt & Schwentner 2018). However, many soft-part characters exhibit extensive intraspecific variability, which often overlapped with interspecific variation, a well-known phenomenon and problem in the taxonomy of extant Spinicaudata.

Carapace traits have been only superficially studied in extant species, and our understanding of intra- and interspecific variability is limited. On the one hand, a valuable character set may have been underappreciated in extant taxonomy. On the other hand, we lack quantitative and comparative studies that evaluate the performance of carapace traits to accurately discriminate species. This may have important ramifications for palaeontological studies, if commonly used carapace traits such as shape or ornamentation patterns are not species-specific or are too variable. A better understanding of carapace traits in extant species will help mediate between extant and fossil taxonomy.

The overarching goal of this study is to test the diagnostic power of carapace traits using a set of known species. As a test case, we selected species of the extant Australian cyzicid *Ozestheria*, which had been studied by Schwentner *et al.* (2015*a*). This is the ideal test dataset, as each individual has previously been studied genetically and thereby assigned to a species. So, compared to a 'classical' palaeontological study, we have a prior species identification independent of carapace traits that we can use to evaluate model performance. Also, all putatively morphologically cryptic species have already been identified and delimited (for example, *Ozestheria packardi* was split into multiple species including the herein studied species A, B, K, Q3, Q5 and S). A total of 21 *Ozestheria* species have been genetically delimited (denoted species A–U) (Schwentner *et al.* 2015*a*). Four of these species represent highly divergent genetic lineages, which probably contain additional species whose status as distinct species is debatable as their genetic differentiation is comparatively low (these were denoted with an additional number, e.g. Q1–Q5). Several of the species were represented by only a single or handful of individuals.

In our study, we focus on the ten species (A, B, C, K, M, N, O, Q3, Q5, S) that yield a sufficiently high number of individuals for statistical analysis. Of the ten species considered here, there are four pairs in sister-group relationships (Schwentner *et al.* 2020), including three sympatrically (A + B) or even syntopically (M + N; Q3 + Q5) occurring sister species pairs and a fourth pair of which the species occur in separate geographical regions (O + S). The studied *Ozestheria* species inhabit a wide range of temporary water bodies across Australia's (semi)arid region.

Based on traditional and geometric morphometrics, we examine whether 'palaeontological' diagnostic criteria allow for the discrimination of these *Ozestheria* species. Our hypotheses (H1–H3) separately address size, shape, and ornamentation. Preliminary species assignments of each individual are based on the genetic data of Schwentner *et al.* (2015*a*):

H1. Size and shape are informative for the discrimination of *Ozestheria* species.
H1a: Species are morphologically distinct (separate analyses of size and shape).
H1b: The combination of size and shape variables leads to higher accuracies in species discrimination than analysing each separately.
H2: Carapace ornamentation is species-diagnostic.
H2a: Carapace ornamentation is distinct for each species.
H2b: Carapace ornamentation observed under incident light reflects genetic groupings.
H3: It is possible to distinguish meaningful morphotypes (ideally species) based on size and shape without additional a-priori genetics-based species assignments (*explorative approach*).

For analyses relating to H1, we calculate functions based on predefined individuals and we use these functions to classify individuals of unknown species identity. The number of correctly classified individuals will be our indicator for model performance and morphological separation between species. For H2, one individual per species will be studied in detail under the scanning electron microscope, while all other individuals are studied under incident light. In addition, our long-term goal is to establish a statistical routine for species discrimination based on size, shape and ornamentation for specimen datasets with no prior information on species identity. Thus, we tested an explorative approach (H3) to distinguishing species that does not require any a-priori information on the species affiliation of individuals.

MATERIAL AND METHOD

Metadata

Specimens (n = 481) of the present study were collected by Brian V. Timms, Stefan Richter, Claire Sieves and Martin Schwentner from 108 different localities in Australia (Fig. 1) between 1999 and 2011, and are registered at the Australian Museum Sydney. Schwentner *et al.* (2015*a*) featured more than 20 *Ozestheria* species, many represented by a single or very few individuals.

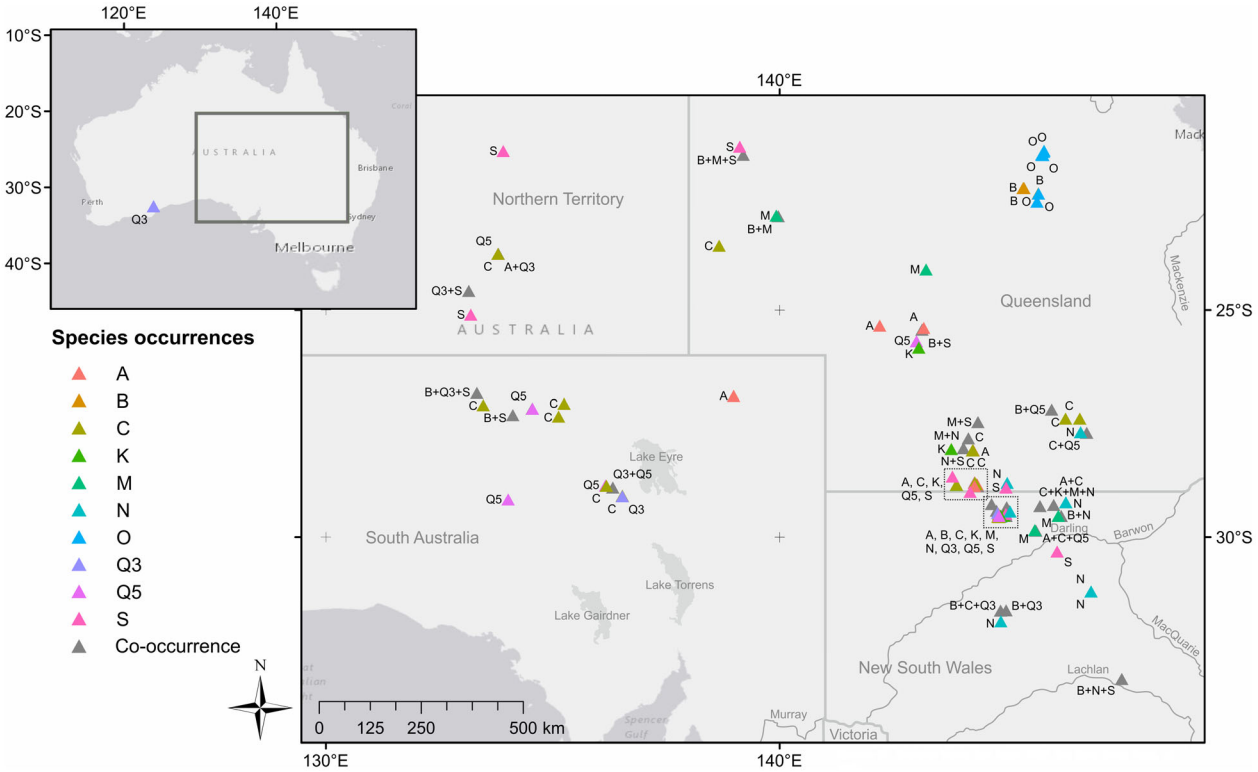

FIG. 1. Species (co-)occurrences of the 108 sampled localities in Australia (produced in ArcMap 10.7; Esri, HERE, Garmin, © OpenStreetMap contributors, and the GIS user community). Plus signs (+) between species specify co-occurrences. Dotted boxes indicate two densely sampled areas, for which occurrences are summarized. *Ozestheria* was only one of several genera that occurred in these pools. *Eulimnadia*, *Limnadia*, *Paralimnadia*, *Limnadopsis*, *Eoleptestheria* and *Eocyzicus* as well as *Triops* and Anostraca were present with a large number of species. Up to four *Ozestheria* species and ten spinicaudatan species occurred within a single pool.

Here we considered all (sub-)species containing 30+ individuals (A, B, C, K, M, N, O, Q5, S) to ensure that within-species morphological disparity is sufficiently captured. As a consequence, more than half of the species that may bias species delimitation success had to be excluded. However, we included species Q3 (with 23 individuals), to assess interspecific variation in two very closely related species (Q3 and Q5).

Metadata (Table 1; Hethke *et al.* 2022, data S1) for the studied individuals include information on accession numbers, species, sex, crowding and geographical coordinates. The 'crowding' category distinguishes between the presence or absence of growth-increment crowding in spinicaudatan carapaces, which represents a morphological feature that can also be used in palaeontological studies. It is assumed that crowding of growth increments occurs when individuals have reached maturity and reduce their growth rate. Some individuals yield multiple areas of carapace crowding, implying that environmental changes may trigger crowding as well. During data acquisition, we noticed that a few non-crowded individuals yielded eggs or fully developed claspers, indicating that crowding is only loosely related to maturity.

Data acquisition, transformation and standardization

Outlining and size measurements. All individuals were photographed under incident light at the Center of Natural History (CeNak), Universität Hamburg, Germany, and at Naturhistorische Museum Wien, Austria. Carapaces were then outlined and measured using the vector graphics software CorelDRAW X7 and Inkscape v0.92.3, producing an image set of 481 outlines and a corresponding dataset of nine linear measurements per individual (Fig. 2; variables *a*, *b*, *c*, *Arr*, *Av*, *Ch*, *Cr*, *H*, and *L* based on Defretin-Lefranc 1965 and Tasch 1987). The position of the posterior extremity of the dorsal margin (landmark E) was commonly clearly defined in *Ozestheria* individuals (except for several specimens of species K), while that of the anterior dorsal extremity (D) was often difficult to identify, leading to variable measurements of *Av*.

Size data. In a first step, we used raw length as a simple size measure. For further analyses, the linear dataset (Hethke *et al.* 2022, data S2) was scaled by dividing each measurement by the arithmetic mean of all nine measurements of a particular individual. Scaled values were then

TABLE 1. Metadata summary, including the total number of individuals, females, males and juveniles (based on the examination of soft parts), the number of individuals that lack growth-band crowding (proxy for the juvenile stage in fossils), and length range.

Lineage	Individuals	Females	Males	Juveniles	Juveniles (based on crowding)	Length range (mm)
A	35	11	18	6	6	1.75–7.28
B	48	13	18	17	15	2.12–6.50
C	74	36	37	1	12	9.03–13.75
K	62	26	34	2	12	2.50–5.42
M	30	14	16	0	1	4.46–7.79
N	58	24	25	9	18	1.63–8.24
O	30	18	8	4	0	3.21–8.05
Q3	23	10	13	0	0	3.69–6.96
Q5	43	19	23	1	3	3.14–5.50
S	78	33	34	11	9	2.18–7.33

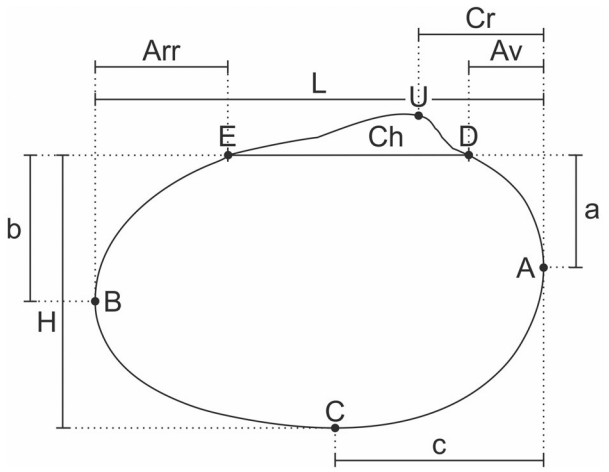

FIG. 2. Example outline of species B (male) and corresponding nine linear variables, measured following Defretin-Lefranc (1965) and Tasch (1987). *Landmarks*: A, most anterior point of the valve; B, most posterior point of the valve; C, most ventral point of the valve; D, anterior extremity of the dorsal margin; E, posterior extremity of the dorsal margin; U, midpoint of the larval valve. *Variables*: a, vertical distance of A to the dorsal margin; b, vertical distance of B to the dorsal margin; c, horizontal distance of C to the extension of the most anterior part of the valve; Arr, distance of E to the extension of the most posterior part of the valve; Av, distance of D to the extension of the most anterior part of the valve; Ch, length of the dorsal margin; Cr, distance of U to the extension of the most anterior part of the valve; H, valve height; L, valve length. Abbreviations for distances L, H, Ch, Av, and Arr derive from the French words longueur, hauteur, charnière, avant and arrière, respectively.

transformed using the natural logarithm, which brought the variables on a real space with similar algebraic structure as the shape data (below). Such ratios of linear variables are often used to describe simple shape relationships (Tasch 1987, p. 35; Scholze & Schneider 2015). This standardization of size variables also reduced ontogenetic variability in the size set (see column 'length range' in Table 1) and ensured normality as precondition for further statistical analyses. In order to guarantee full-rank covariance matrices for eigenspace approaches (linear discriminant analysis, principal component analysis), we checked the pairwise relationships of the scaled size-variables, (1) for the full dataset and (2) for each species, using the Pearson correlation coefficient of R-function *pairs.panels()* of package *psych*. Subsequently, we removed variable *Ch* (see Bivariate Scatter Plots, below).

Shape data. Since the early 1970s, many algorithms have been developed to describe shape in biology. Reviews from different points of application are provided by, Haines & Crampton (2000), Claude (2008), Pappas *et al.* (2014) and Wishkerman & Hamilton (2018), for example. For the *Ozestheria* dataset we chose Fourier shape analysis following Haines & Crampton (2000), which decomposes *xy*-coordinates into harmonically related sine and cosine curves. Two Fourier coefficients are produced per harmonic, and these coefficients constitute the shape variates for further analyses. Outlines of right valves and mirrored left valves were automatically traced using programme tpsDig2 (v2.31; Rohlf 2018). 2500 equidistant *xy*-coordinates, which describe each outline in counter clockwise direction, were then generated using the well-defined posterior dorsal extremity as starting position. Shape was explored by means of Fourier shape analysis by describing the cumulative change in angle of a tangent vector to the outline as first introduced by Zahn & Roskies (1972) using software packages HANGLE and HMATCH (Crampton & Haines 1996). HANGLE removed high frequency pixel noise resulting from automatic outline tracing (set to 10 smoothing iterations) and performed a Fast Fourier Transformation (FFT). HMATCH then normalized for starting position and orientation (Haines & Crampton 2000).

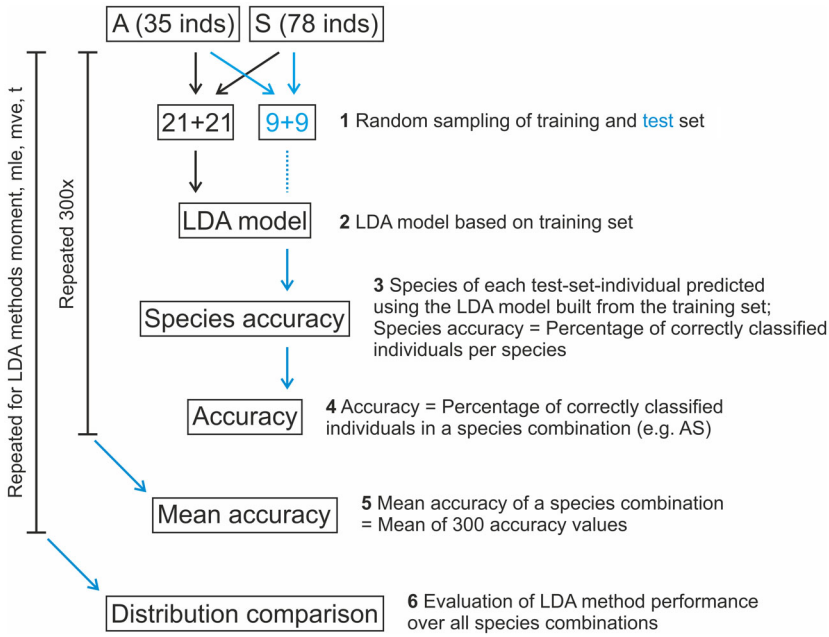

FIG. 3. Analytical steps for H1a and H1b, example for lineage pair A + S. The routine was repeated for all possible lineage combinations (pairwise = 45×). The test set contains individuals of unknown species affiliation. *Abbreviation*: LDA, linear discriminant analysis.

Fourier shape analysis method comparison

In a preliminary study, we identified the highest order positive harmonic to be kept based on species combination B + K + N + Q5. The resultant number was used in all further Fourier shape analyses of this study. Five Fourier shape analyses were performed by setting HANGLE to 4, 8, 12, 16 and 24 harmonics, respectively. For each of the five settings, we carried out the statistical routine outlined in Figure 3 (described below) and used mean accuracy over 300 iterations and the standard deviation of the resulting distribution as estimators for model performance. Combination B + K + N + Q5 was chosen because measurements and outlines of these four species were available at an early stage during data acquisition. Based on this preliminary analysis we set all further Fourier shape analyses to 8 harmonics, resulting in a total of 14 shape variates (= Fourier coefficients A2–A8 and B2–B8; Hethke *et al.* 2022, data S3, Data_analysis S4).

Intraspecific variability

The full dataset represents a non-filtered set of specimens of various ontogenetic stages, sex and ecophenotypes, similar to a dataset that one would collect during a fossil excavation. To constrain some of the resulting intraspecific variability, we assessed individuals of all ontogenetic stages ('full set') and, in a second step, individuals yielding crowded growth bands only ('reduced set'), which should reduce ontogenetic variability by excluding juveniles. Nevertheless, the reduced set still contains a wide range of ontogenetic stages, including individuals yielding just one crowded growth band as well as late stages with multiple areas of crowding on the carapace (compare with column 'length range' in Table 1).

To address sexual dimorphism, we tested whether there are significant differences between males and females: (1) in standardized and log-transformed linear variables; and (2) in shape variables using R-function *manova()* of package *stats* (for each species). Secondly, we tested whether using single-sex datasets (female-only or male-only) increases species delimitation success (part of the H1-routine described below).

Data analysis for H1 (Are species morphologically distinct?)

Length. The taxonomic value of raw length was evaluated as a simple measure for size by using the reduced set. To assess whether there are significant differences between species, we used a combination of the Kruskal–Wallis and Wilcoxon rank sum tests provided by R-functions *kruskal.test()* and *pairwise.wilcox.test()* of package *stats* and adjusted the *p*-values using the Bonferroni correction of function *p.adjust()*.

Non-parametric MANOVA. Differences in scaled and log-transformed linear variables and shape variates between individuals of the ten different species were explored by means of non-parametric MANOVA using the software PAST v4.07b (Hammer *et al.* 2001).

Linear discriminant analysis of the size and shape datasets. To address morphological relationships between species pairs, we calculated linear discriminant models that can be used to classify individuals of unknown species affiliation and then assessed model performance for all 45 possible pairwise combinations. Data analyses for H1a and H1b followed the six steps outlined in Figure 3, and they were implemented in R v.4.2.1 (R Core Team 2013; see Hethke *et al.* 2022, renv.lock, Data_analysis S4–S7 for markdown files, which feature both code and results):

1. We randomly sampled 21 (full set) and 20 (reduced set) individuals per species for the training set, and 9 individuals per species for the test set (without replacement of individuals over training and test set, so the latter contains individuals without predefined species affiliation). The number of individuals chosen for data subsetting was based on species M and O, which both contain 30 individuals in the full set (Table 1). As the full and reduced sets of Q3 both contain 23 individuals, we sampled 16 and 7 individuals of each species for the training and test sets of combinations that contained Q3, respectively. Training and test sets of female-only and male-only analyses were created by randomly sampling 70% and 30% of the population size, respectively, due to highly variable sample sizes of males and females between species (Table 1). Subsequently, randomly sampled multivariate datasets of single species were combined to reflect the desired species combination (e.g. A + S in Fig. 3). This standardization of sample size makes results of different species combinations better comparable.
2. The resultant training sets were then projected using linear discriminant analysis (LDA; R-function *lda()* of package MASS; Venables & Ripley 2002; Ripley *et al.* 2021), with species as the grouping factor and scaled and log-transformed linear variables or Fourier coefficients as the discriminators (covariates). *A priori* species allocations were based on previous molecular species delimitation (Schwentner *et al.* 2015*a*). LDA searched for the linear combination of covariates that maximized separation between the species groups (Hammer & Harper 2006).
3. Based on the resultant LD model, individuals of the test set were then allocated to a species using the posterior probabilities calculated by function *predict()* of package *stats*. Thus, in case of combination A + S (Fig. 3), the model was trained on 42 individuals and tested by classifying the 18 individuals of the test set. For each analysis, we then calculated 'species accuracy', which gives the percentage of correctly identified individuals per species.
4. The term 'accuracy' describes mean species accuracy (see step 3) in a species combination.
5. As the training and test sets were randomly sampled for each LDA, the coefficients of the respective discriminant functions differed between single runs, affecting accuracy values. Thus, steps 1–4 were iterated 300 times and accuracy was noted for each run to estimate 'mean accuracy' for a species combination, which provides a measure for model performance and, in biological terms, for morphological separation between respective species combinations. This measure becomes more robust the higher the number of individuals in the training- and test sets.
6. The performance of the four different LDA methods provided by R-function *lda()* of package MASS ('moment', 'mle', 'mve' and 't') was evaluated by using mean accuracy as the estimator. To estimate 'overall mean accuracy' of the 45 pairwise species combinations, we used the geometric mean of 300 × 45 accuracy values.

Combination of size and shape datasets for H1b. In order to combine the standardized and log-transformed linear variables and the Fourier coefficients of species A, B, K, M, N, O, Q3, Q5, and S (407 individuals full set; 343 individuals reduced set), we carried out an LDA for each dataset and saved the respective LD-scores provided by R-function *predict()* of package *stats*. Species C was excluded, as it was morphologically highly differentiated from the other nine species based on the H1a results. The two size and shape LD-datasets were then combined and weight-transformed using the 2.5% and 97.5% quantiles as weight constants (after equation (2) with $b_j = Q_{2.5\%}$ and $a_j = Q_{97.5\%}$ in Miesch 1980). Then the H1b-analysis followed steps 1–6 of Figure 3.

Data analysis for H2 (Is carapace ornamentation species-diagnostic?)

Ornamental features of the ten lineages were described in detail based on scanning electron microscopy (SEM) of one individual per species (H2a). All individuals were then examined under incident light: (1) to quantitatively assess within-species variability in ornamentation; and (2) to confirm that all carapaces of a species yielded the same ornamentation type (H2b). Descriptive terms used for main carapace and ornamental features (Table 2) follow Scholze & Schneider (2015) and Gallego *et al.* (2020). A graphical definition of the terms 'growth band', 'growth increment', 'growth line' and 'concentric ridge', in the way these terms are applied here, is illustrated in Figure 4. The term 'growth band', defined as the space between two growth lines, is well-established in the palaeontological literature (e.g. Li &

TABLE 2. Descriptive terms for main carapace and ornamental features used in this study.

Term	Description
Main carapace features	
Concentric ridge	Ridges parallel to growth lines (Raymond 1946); 'concentric ribs' of Scholze & Schneider (2015).
Growth line	Suture between two growth bands.
Growth band	Space between two growth lines; in the *Treatise on Invertebrate Paleontology*, Tasch (1969, p. R94), defined as the 'area between any two growth lines. . .; synonyms used interchangeably include intervales, growth zone, growth band.' The term 'growth band' is used by many palaeontological research groups (e.g. Li & Batten 2004*a*, 2004*b*, 2005; Astrop & Hegna 2015; Scholze & Schneider 2015; Gallego *et al.* 2020; Li 2022). To avoid confusion with overall orientation on the carapace, we use the terms 'upper' and 'lower' instead of 'dorsal' and 'ventral', respectively, for positions within single growth bands.
Secondary growth phase	Growth bands following carapace crowding, yielding abnormal ornamentation and reduced concentric ridges, corresponding to the 'disorganized ornamentation near the ventral margin' of Sun & Cheng (2022).
Ornamental features on growth bands	
Punctae	Small, roundish depressions.
Radial lirae	Radial structures, can be straight or anastomosing. Scholze & Schneider (2015) further distinguished 'radial lirae' and 'radial fringe' (the latter being anastomosing); we do not follow this distinction here as there are transitions between these features within a single carapace.
Reticulation	Polygonal ornamentation.
Serrate margins	Lower margins of growth bands with notches, which represent setal pores (compare Shen 2003).
Setal pores	Pores associated with broken-off setae, located on concentric ridges or at the distal part of concentric ridges, the latter leading to serrate margins.

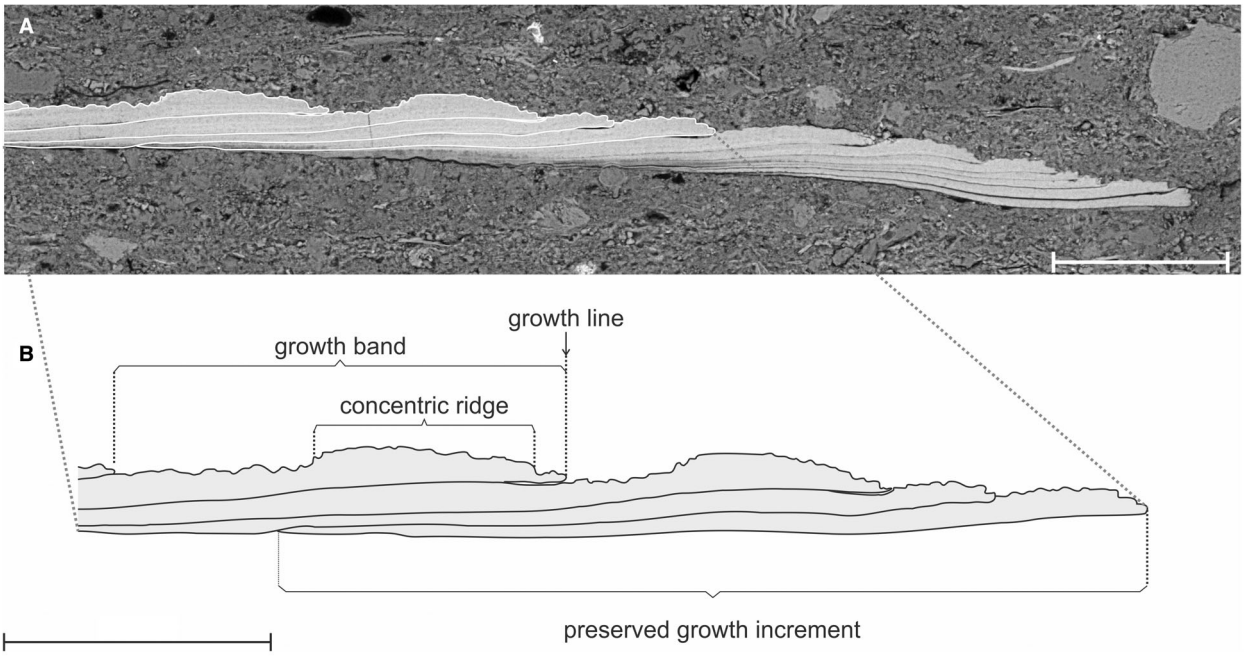

FIG. 4. Main carapace features. A, carapace margin of an exceptionally well-preserved fossil spinicaudatan in thin section (*Eosestheria*, Early Cretaceous; SEM image using back-scattered electrons); the fossil consists of fluorapatite, crystallized from calcium phosphate biominerals during fossil diagenesis; preserved growth increments are thinning towards the dorsal part of the carapace, their internal lamellar structure is well preserved (similar to cross sections of extant spinicaudatan cuticles in Rieder *et al.* 1984). B, growth lines represent the sutures between two growth increments, growth bands the space between two growth lines. Scale bars represent 100 μm.

Batten 2004*a*, 2004*b*, 2005; Astrop & Hegna 2015; Scholze & Schneider 2015; Gallego *et al.* 2020; Li 2022), including later works of Tasch, such as his seminal work on southern hemisphere fossil conchostracans (Tasch 1987). The terms 'intervale' and 'interval' are often used in biological literature (e.g. Rogers 2020). 'Intervale' was defined by Raymond (1946, p. 224) as 'the area between two concentric ridges' (term adapted from French authors), while 'interval' may also represent the 'space between any two [radial] ribs, costae, or costellae' (Tasch 1969, p. R94). The two terms 'intervale' and 'interval' are very similar and easy to confuse, hence, we prefer 'growth band'. We classified setae on the carapace as either being short and stout, or long and thin, corresponding to wide and narrow setal pores, respectively. We did not specify the type of setae in more detail, as we are focusing on carapace features that yield a high preservation potential. Also, these were often broken-off or damaged.

Data analysis for H3 (Is it possible to distinguish species without prior genetic information?)

H3 explores a procedure for identifying a meaningful number of morphotypes in a quantitative set of outlines from a single pool or excavation (application for palaeoecological studies) and for all ten species (application for taxonomic studies) by using a combination of principal component analysis (PCA) and cluster analysis. This approach does not require any *a priori* information on the species affiliation of individuals. To our knowledge, the highest number of *Ozestheria* species co-occurring in a single pool is four (a claypan west of Engonia in New South Wales, which features lineages C, K, M and N; Fig. 1), while the common *Ozestheria* species diversity per pool ranges between one and three (Schwentner *et al.* 2015*b*). For our approach, we used the conservative number of four species in a water body and based data analysis following H3 on the real-life combination C + K + M + N (n = 181) and on our test dataset B + K + N + Q5 (n = 163). We also tested combining all ten species (reduced set). The following analytical steps were applied:

1. We performed a PCA for the size and shape datasets (separately).
2. From both PCAs, we combined the scores of the first three PCs and performed a second PCA on this six-dimensional coordinate set.
3. In order to find the best linkage method for hierarchical cluster analysis, we calculated the highest agglomerative coefficient (AC) using R-function *agnes()* from package *cluster* (Maechler *et al.* 2021), considering the linkage methods 'average', 'single', 'complete' and 'ward' (Kaufman & Rousseeuw 1990).
4. A hierarchical cluster analysis was performed by using R-function *agnes()* with the linkage method yielding the highest AC.
5. For finding the optimal number of clusters, we used the gap-statistic by Tibshirani *et al.* (2001) and the R-function *clusgap()* of package *cluster*. This algorithm compares the total intra-cluster variance for different values of k (number of clusters) with their expected values for a distribution with no clustering. The optimal number of clusters is automatically indicated by a dashed line in the plot produced by R-function *fviz_gap_stat()* of package *factoextra*.
6. We cut the hierarchical cluster tree based on the optimal cluster number k and compared the distribution of individuals among the resulting clusters with the known species assignments (based on Schwentner *et al.* 2015*a*).

RESULTS

Fourier shape analysis method comparison

The number of harmonics kept strongly influences model performance in the B + K + N + Q5 test set (Fig. 5; Hethke *et al.* 2022, Data_analysis S4). Harmonics 8 and 12 yield similar accuracy distributions with decreasing model performances towards 4 and 24 harmonics. Setting HANGLE to eight harmonics yields the highest mean accuracy (96.8%). In addition, there are no remarkable differences between methods 'moment', 'mle' and 't', which produce similar mean model accuracies and ranges (Hethke *et al.* 2022, Data_analysis S4, 3.2.3). In turn, method 'mve' yields significantly lower mean accuracy values across all examined species combinations ('Method comparison' chapters of Hethke *et al.* 2022, Data_analysis S5–S7).

Bivariate scatter plots

Bivariate relationships of both size and shape variates are affected by species C, which is well separated from all other species (Hethke *et al.* 2022, Data_analysis S5, 1.7.1, S6, 1.5.1). For example, the relationship of shape variates A2 and B2 is comparatively strong in the dataset containing ten species (n = 481, r = −0.74), but the correlation value sharply decreases when species C is excluded in a dataset containing nine species (r = −0.20). Generally, relationships of shape variates are weak when single species are examined (Hethke *et al.* 2022, Data_analysis S6, 1.5.3), indicating that they will be informative for further species discrimination.

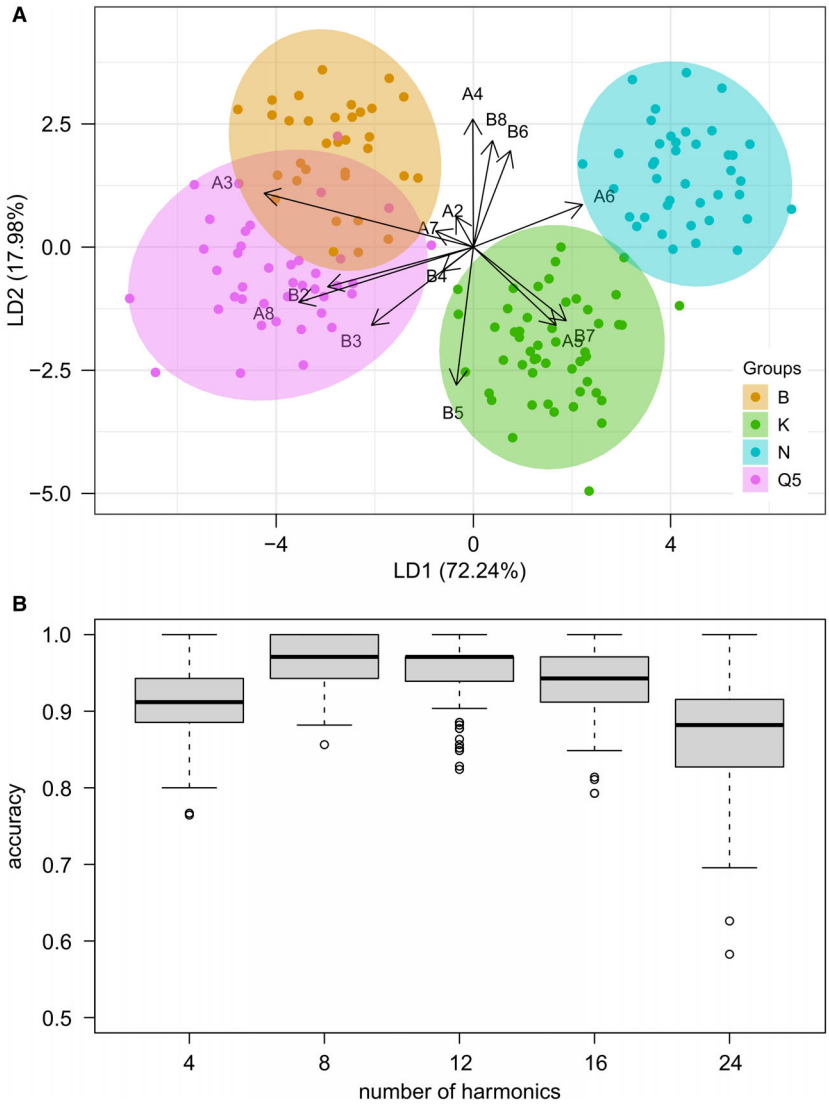

FIG. 5. B + K + N + Q5 shape analysis (H1 approach). A, LDA biplot, reduced set. B, accuracy distributions for the test datasets (reduced set, 300 iterations, method 'moment') of five different Fourier shape analyses with the highest order positive harmonic kept set to 4, 8, 12, 16 and 24, respectively; summary statistics are indicated by a horizontal black line (median), two hinges that represent the 25th and 75th percentiles, and two whiskers indicating the range 1.5 times the inter-quartile range from the lower and upper hinges, respectively; outliers are specified as points; model performance is highest when setting HANGLE to 8 harmonics.

Pearson correlation coefficients of the size dataset for nine species (n = 407; Hethke *et al.* 2022, Data_analysis S5, 1.7.2) indicate strong linear dependencies between *Arr* and *Ch* (r = −0.85) and *b* and *L* (r = −0.72). The latter is species-specific: while the carapace grows, *b* is becoming smaller relative to *L* in lineages O, Q3, and S. In contrast, almost all investigated lineages yield a strong linear relationship between *Ch* and *Arr* (Hethke *et al.* 2022, Data_analysis S5, 1.7.3). Hence, we decided to drop *Ch* in further analyses.

H1: Are species morphologically distinct?

Pairwise length. Length as a simple size measure distinguishes C from all other species (Fig. 6). Of the remaining nine species, Q5 yields the smallest and N and O the largest carapaces. Although there is plenty of overlap in length between species, each species yields a narrow length range that is especially confined in Q5, K, and M. The Kruskal–Wallis test indicates that there are significant differences in length between species ($p < 0.001$). Pairwise comparisons show that sister species A + B, M + N, Q3 + Q5 and O + S yield distinct lengths (despite some overlap), while species pairs that are more distantly related, such as B + Q5, cannot be distinguished by length.

Pairwise size. Pairwise species comparisons based on scaled and log-transformed size variables yield overall mean accuracies in the assignment of individuals to species of 92.9% (full set) and 93.6% (reduced set; Fig. 7A) over all 45 combinations (Hethke *et al.* 2022, Data_analysis S5, 4.1.1, 4.1.2). However, mean accuracies differ between lineage pairs: 37 lineage pairs (reduced set) yield

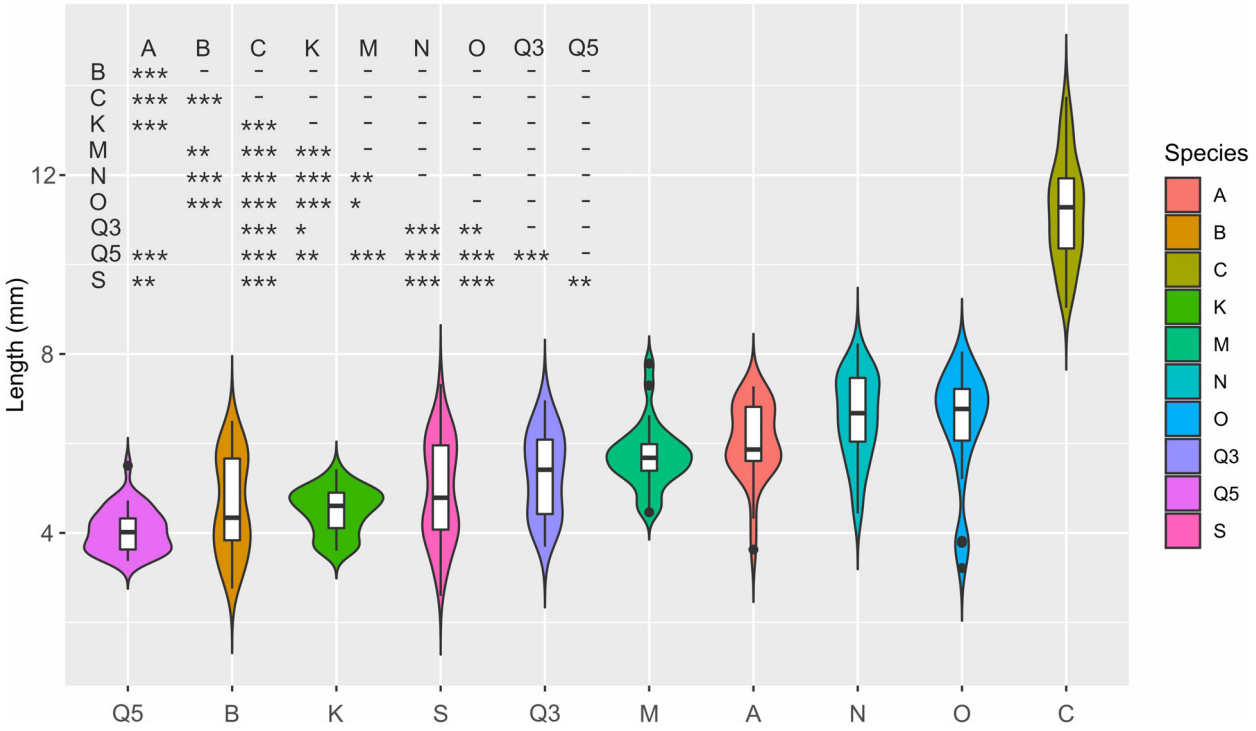

FIG. 6. Box and violin plot of raw length values, reduced set, including the interquartile range, median and the kernel density estimate. Summary statistics as for Figure 5B. Species were ordered by median. The inset table gives the results of the Wilcoxon pairwise test, which indicates that sister species pairs A + B, M + N, Q3 + Q5 and O + S yield distinct lengths. Significance codes represent: ***$0 < p \leq 0.001$; **$0.001 < p \leq 0.01$; *$0.01 < p \leq 0.05$.

mean accuracies of 100–90%, four of 90–80%, and four of <80%. The four pairs yielding the lowest mean accuracy values include two pairs of sister species: Q3 + Q5 and O + S (mean accuracies of 76.9% and 68.8%, respectively). The sympatrically occurring sister pair A + B and the syntopically occurring sister pair M + N are comparatively well separated by linear variables (mean accuracies of 89.2% and 91.6%, respectively). As the anterior dorsal extremity was often hard to identify in *Ozestheria*, we excluded variable Av in a test run but found that the inclusion of *Av* into the analysis enhances overall model performance. According to NPMANOVA, each species is morphologically distinct at $p \leq 0.001$ (Table 3).

Pairwise shape. Pairwise species comparisons based on shape variates yield overall mean accuracies of 93.0% (full set) and 93.8% (reduced set; Fig. 7B) over all 45 combinations (Hethke *et al.* 2022, Data_analysis S6, 3.1.1, 3.1.2). As for size, 37 lineage pairs (reduced set) yield mean accuracies of 100–90%, five of 90–80% and three of <80%. In comparison to size, shape yields higher overall mean accuracies; pair B + Q5, for example, notably increased from 86.9% to 94.7%. But there are also counterexamples where size performed better than shape: mean accuracies of sister pairs A + B and M + N both decreased to 85.4% and 88.0%, respectively. According to NPMANOVA of the shape variates of the ten examined *Ozestheria* species, all pairwise comparisons are highly significant ($p < 0.0001$; Table 3), indicating that each species is morphologically distinct.

Pairwise size and shape combined (H1b). Overall mean accuracies of the full and reduced sets of the combined analysis (H1b; 92.5% full set; 93.1% reduced set; Hethke *et al.* 2022, Data_analysis S7) are slightly higher than mean accuracies of the corresponding 36 pairwise combinations of the separate size and shape analyses (size: 91.9% full set, 92.7% reduced set; shape: 92.0% full set, 92.9% reduced set).

Intraspecific variability. Ontogenetic variability has been successfully constrained by using the reduced set over the full set, which increases overall model performance (for illustrations see Hethke *et al.* 2022, Data_analysis S5, 4.1.3 and S6, 3.1.3). For example, mean accuracy of the size dataset for pair K + N increases notably from 91.1% (full set) to 97.9% (reduced set): The datasets of both species yield numerous juvenile individuals (Table 1), hence, their exclusion from the reduced set explains the sharp increase in mean accuracy. Similar jumps in mean

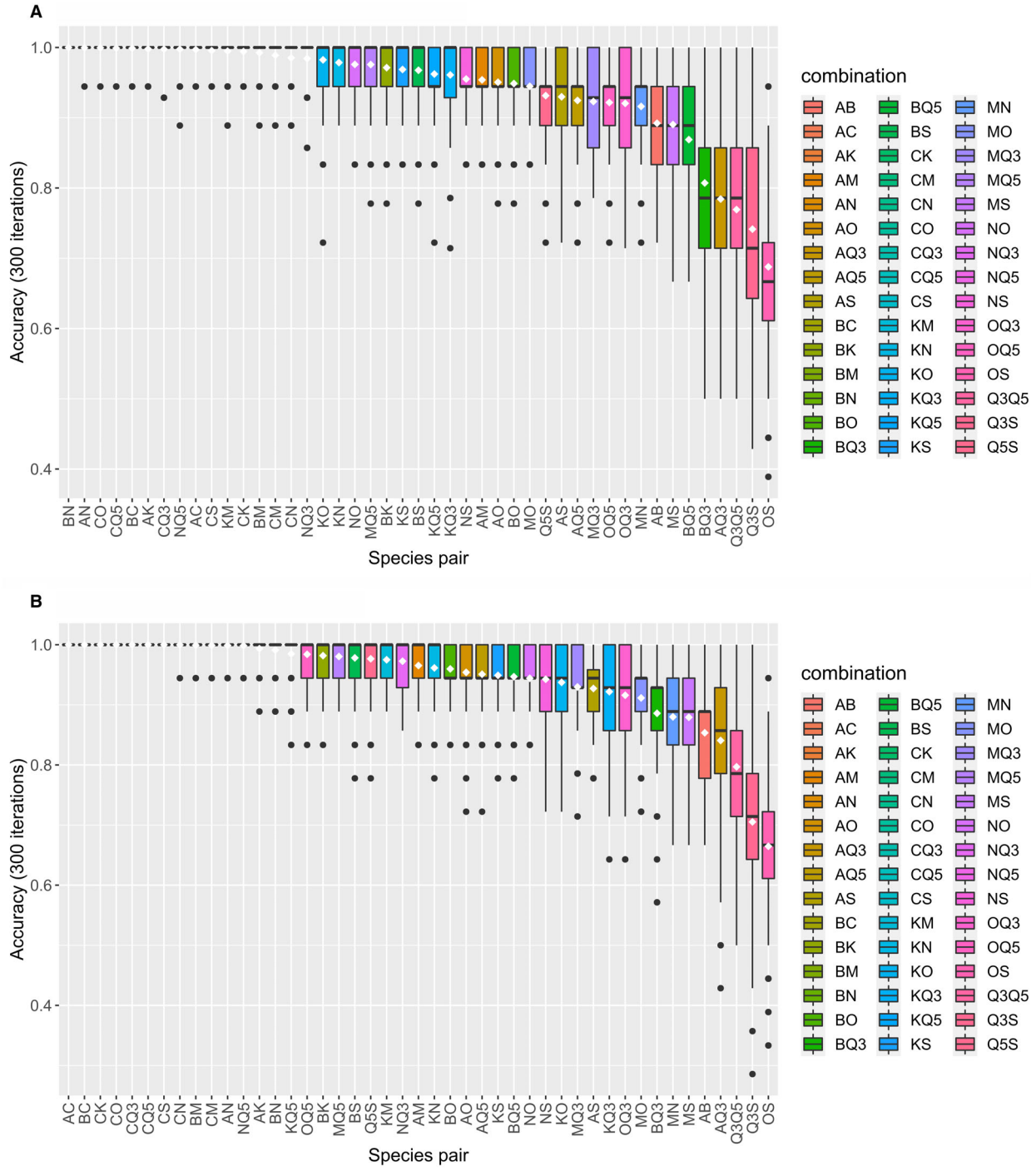

FIG. 7. Pairwise species comparisons, reduced set for: A, size variables; B, shape variates. The boxplots illustrate 45 distributions and summary statistics of 300 accuracy values each (= 13 500 LDA runs, method 'moment') sorted by mean in descending order. Summary statistics as for Figure 5B. Mean model accuracy is indicated by a white diamond for each pairwise comparison.

accuracy can be observed for combinations A + O and A + Q5. In turn, there are a few species pairs that show a (less substantial) drop in mean accuracy between the full set and the reduced sets (e.g. M + N, O + S). For the shape dataset, model performance increased most for pairs A + Q5, B + Q3, and A + Q3 when juveniles were excluded.

Nine of the ten species are clearly sexually dimorphic (Table 4). Dimorphism is especially pronounced in A (Fig. 8), K, N and S; moderate in C, O, Q3, and Q5; and

TABLE 3. Non-parametric MANOVA of scaled and log-transformed linear variables (A) and shape variates (B) of the ten examined *Ozestheria* species.

	B	C	K	M	N	O	Q3	Q5	S
Size									
A	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0011	0.0001	0.0003
B		0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
C			0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
K				0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
M					0.0002	0.0001	0.0001	0.0001	0.0001
N						0.0001	0.0001	0.0001	0.0001
O							0.0003	0.0001	0.0007
Q3								0.0001	0.0001
Q5									0.0001
Shape									
A	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
B		0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
C			0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
K				0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
M					0.0001	0.0001	0.0001	0.0001	0.0001
N						0.0001	0.0001	0.0001	0.0001
O							0.0001	0.0001	0.0001
Q3								0.0001	0.0001
Q5									0.0001

All pairwise comparisons are highly significant, indicating that each lineage is morphologically distinct.

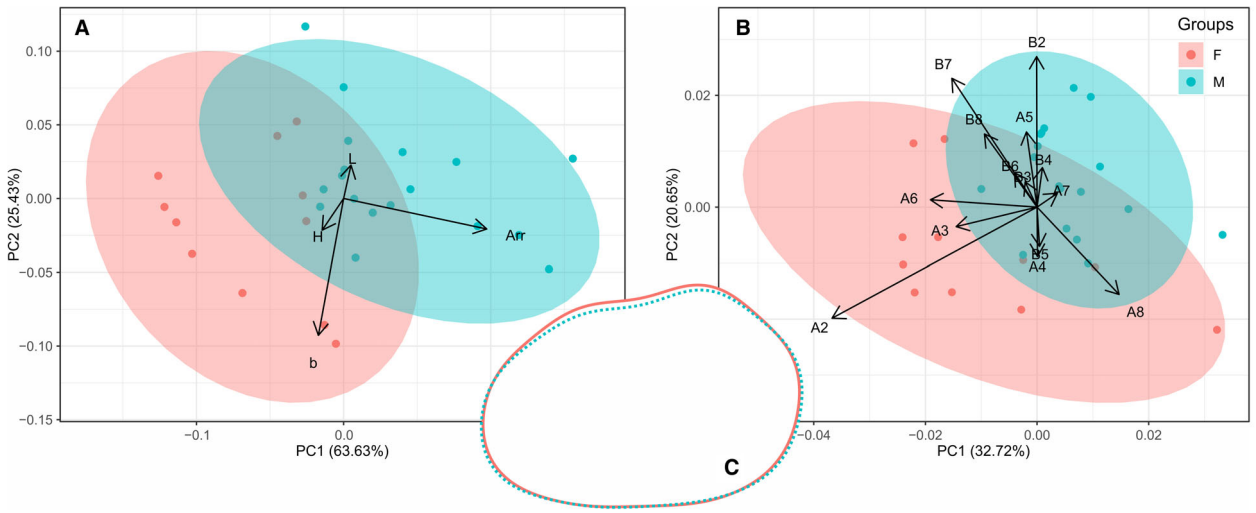

FIG. 8. A–B, example biplots of scores and loadings on PC1 and PC2 of size variables (A) and shape variates (B) in the sexually dimorphic species A. C, synthetic mean outlines of males (dotted line) and females (solid line). The female carapace is slightly higher than that of males (= higher H/L-ratio in females).

weak in B. Lineage M is weakly dimorphic based on linear variables and non-dimorphic based on shape variates, but the corresponding PCA (Hethke *et al.* 2022, Data_analysis S6, 2) indicates that morphological disparity is much larger in males than females of that species. Each dimorphic lineage yields a different set of linear variables that are significantly different among sex, of which variable *Arr* is distinct in males and females of most *Ozestheria* species. We maximized the visual separation of males and females of each lineage by restricting the respective PCA to the significant linear variables (Fig. 8A). Regardless, males and females of each examined species overlap in

TABLE 4. Sexual dimorphism in the ten examined species, based on standardized and transformed linear variables (size) and Fourier coefficients (shape).

Species	Sexual dimorphism size	Linear variables (incl. p = 0.05–0.1)	Sexual dimorphism shape
A	0.00030***	*b***, *Arr****, *H***, *L*†	0.00591**
B	0.04086*	*Arr***, *H*†, *L*†	0.06962†
C	0.00110**	*c*†, *Arr***, *Av***, *Cr*†	0.02038*
K	0.00002***	*a****, *b***, *c***, *Arr*†, *Cr*†, *H****	0.000001***
M	0.01177*	*a*†, *Arr****, *Av*†, *L*†	0.17503
N	0.00031***	*Arr***, *H*†	0.00001***
O	0.00616**	*a**, *Cr**, *H**, *L**	0.00571**
Q3	0.00159**	*Arr***	0.00064***
Q5	0.01046*	*b***, *Cr**, *L***	0.00133**
S	0.000002***	*a*†, *Arr****, *Cr*†, *H**	0.000004***

Each p-value corresponds to a one-way MANOVA that compares 8 linear response variables (*a*, *b*, *c*, *Arr*, *Av*, *Cr*, *H*, *L*) or 14 shape response variables by factor 'sex' (levels male and female). Significance values: ***$0 < p \leq 0.001$; **$0.001 < p \leq 0.01$; *$0.01 < p \leq 0.05$; †$0.05 < p \leq 0.1$.

morphospace. Hence, sexual dimorphism contributes to overall morphological disparity in *Ozestheria* species, but it is less significant than interspecific variability, exemplified by the female-only and male-only analyses that yield similar to lower overall mean accuracy values than analyses containing both sexes.

Over all 45 species combinations, the female-only (f) and male-only (m) size analyses yield mean accuracies of 93.4% and 91.5%, respectively (Hethke *et al.* 2022, Data_analysis S5, 4.1.4, 4.1.5). However, the results for sister pairs strongly diverge between females and males: M + N (f: 86.6%; m: 96.8%), O + S (f: 82.0%; m: 67.2%), A + B (f: 76.9%; m: 92.9%) and Q3 + Q5 (f: 66.8%; m: 77.9%). In comparison, mean accuracies of the reduced set, which combines males and females, of these four combinations are 91.6%, 68.8%, 89.2% and 76.9%, respectively. With only few exceptions, the performance of the reduced set is commonly better than or intermediate between the corresponding female-only and male-only analyses (Hethke *et al.* 2022, Data_analysis S5, 4.1.6).

Due to the higher number of shape variates than size variates, we eliminated Q3 (10 females) and O (8 males) from the female-only and male-only shape analyses, respectively. Again, mean accuracies of the female-only (f) and male-only (m) analyses of the sister pairs strongly diverge (Hethke *et al.* 2022, Data_analysis S6, 3.1.4, 3.1.5): M + N (f: 75.9%; m: 95.6%), O + S (f: 71.9%; m: –), A + B (f: 67.9%; m: 88.2%), and Q3 + Q5 (f: –; m: 77.9%). For female-only shape, the three analysed sister pairs yield the three lowest mean accuracy values. In comparison, the corresponding mean accuracies of the reduced set (males and females combined) are 88.0%, 66.5%, 85.4% and 79.7%, respectively (Hethke *et al.* 2022, Data_analysis S6, 3.1.2).

Ozestheria morphospace occupation. Species C is almost fully separated by LD1 in the analysis of ten lineages, explaining 76.2% and 73.0% of the between-group variance in the size and shape (Fig. 9A) reduced sets, respectively. Accordingly, C-pairwise-comparisons yield mean accuracies of 98.5–100% (size, Fig. 7A) and 99.9–100% (shape, Fig. 7B). The second highest variance in the dataset lies at an oblique angle to LD2. To capture that axis of morphological change, we re-analysed the remaining nine species in further analyses, including the combined analysis of size and shape variates (H1b).

Linear discriminants LD1, LD2, and LD3 explain 50.5%, 24.4% and 12.0% of the between-group variance in the reduced set of nine species (H1b; Fig. 9B), respectively. More than half of the between-group morphological variability is represented by a change from subrectangular carapaces with flat anterior and ventral margins at negative scores of LD1 (A, B, Q3, Q5) to oval carapaces with round anterior and ventral margins towards positive scores of LD1 (M, N, K). Sister species O and S occupy an intermediate position within the morphospace of the examined *Ozestheria* species, explaining the smaller mean-accuracy values of O and S-combinations in Figure 7. Their confined 95% confidence ellipses, however, indicate low morphological disparity in these two species. Morphological disparity within single species is overall largest in Q3 and smallest in M. Generally, all sister species overlap in morphospace but to varying degrees; of the four pairs, sister species M and N are best separated. Species K is especially well separated, reflecting the pairwise comparisons of Figure 7: Mean accuracies of all K-combinations range between 96.1% and 99.8% (size) and between 92.2% and 100% (shape).

H2: Is carapace ornamentation species-diagnostic?

Ornamentation types. Seven ornamentation types (= ornamental patterns) were identified among the ten examined *Ozestheria* species (Figs 10, 11). In all examined species, ornamental features become irregular with incipient carapace crowding (example growth band arrowed in Fig. 10J). Patterns of the secondary growth phase are significantly different to mid-carapace patterns. If present, the serrate lower margins of concentric ridges are generally reduced in growth bands of the secondary growth phase (well visible in species A, M, N, Q3, Q5, S). Descriptions are mainly based on features visible under

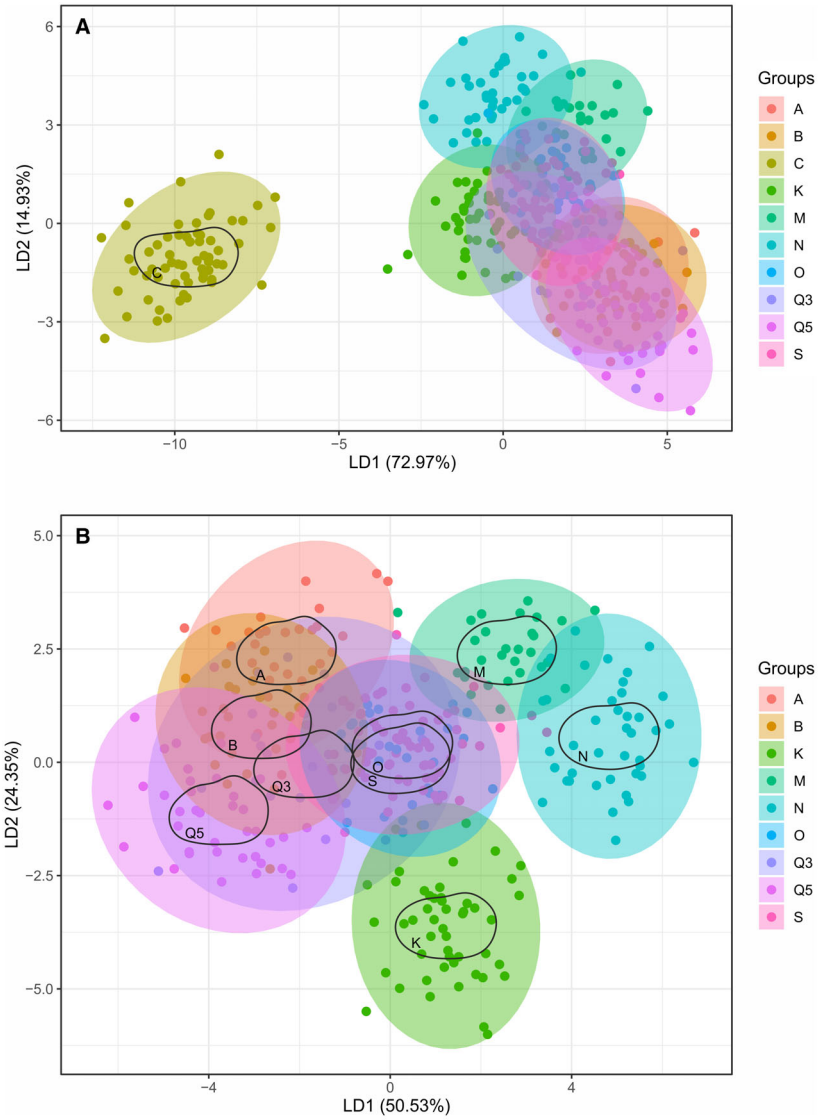

FIG. 9. LDA biplots of shape variates, ten species, reduced set (A), and of the combined size and shape variates (H1b), nine species (excluding species C), reduced set (B). Mean outlines are illustrated for each species. In A, LD1 is driven by species C; it contributes little to the discrimination of the remaining nine species. In B, more than half of the between-group morphological variability in the remaining species is represented by a change from subrectangular carapaces with flat anterior and ventral margins at negative scores of LD1 (A, B, Q3, Q5) to oval carapaces with round anterior and ventral margins towards positive scores of LD1 (M, N, K). Species O and S yield intermediate morphologies that are, however, well confined within the *Ozestheria* morphospace.

incident light, but we also considered one SEM image of a male individual per species:

1. (A; Fig. 10A–C): Early growth bands smooth with well-defined, parallel (never anastomosing or reticulating) lirae along lower margin, increasing in length in growth bands of later ontogenetic stages. Concentric ridges well-developed, raised, broad, and located at the lower end of each growth band, with nodules at its upper margin arranged in a moniliform way and a serrate lower margin (SEM observation; arrowheads in Fig. 10B). Concentric ridges narrower on growth bands of the secondary growth phase (Fig. 10C), lacking well-defined serrate margins. Crowded growth bands with well-defined lirae.
2. (K; Fig. 10G–I): Larval valve with shallow reticulations, transforming to punctae in following growth bands. Upper part of mid-carapace growth bands smooth, lower part with short, intermittent (nodular) and anastomosing (not reticulating) lirae. In later growth bands, lirae become longer and more regular (subparallel) and the smooth area becomes smaller due to growth-increment overlap; in some individuals the smooth area may be very narrow. Most notable is the very dense setation with stout setae (sometimes broken off in earlier growth bands) and corresponding wide setal pores on concentric ridges.
3. (C; Fig. 10J–L): Growth bands covered in mesh of polygonal, coarse reticulations (usually being pentagons, hexagons or heptagons); rarely earliest growth bands smooth (could be an artefact of abrasion). Reticulations become irregular on growth bands of later ontogenetic stages, short thin lirae appear on the lower part of growth bands in the posterodorsal region of the carapace that are much less conspicuous than Type 4 lirae. Coarse reticulations transition to

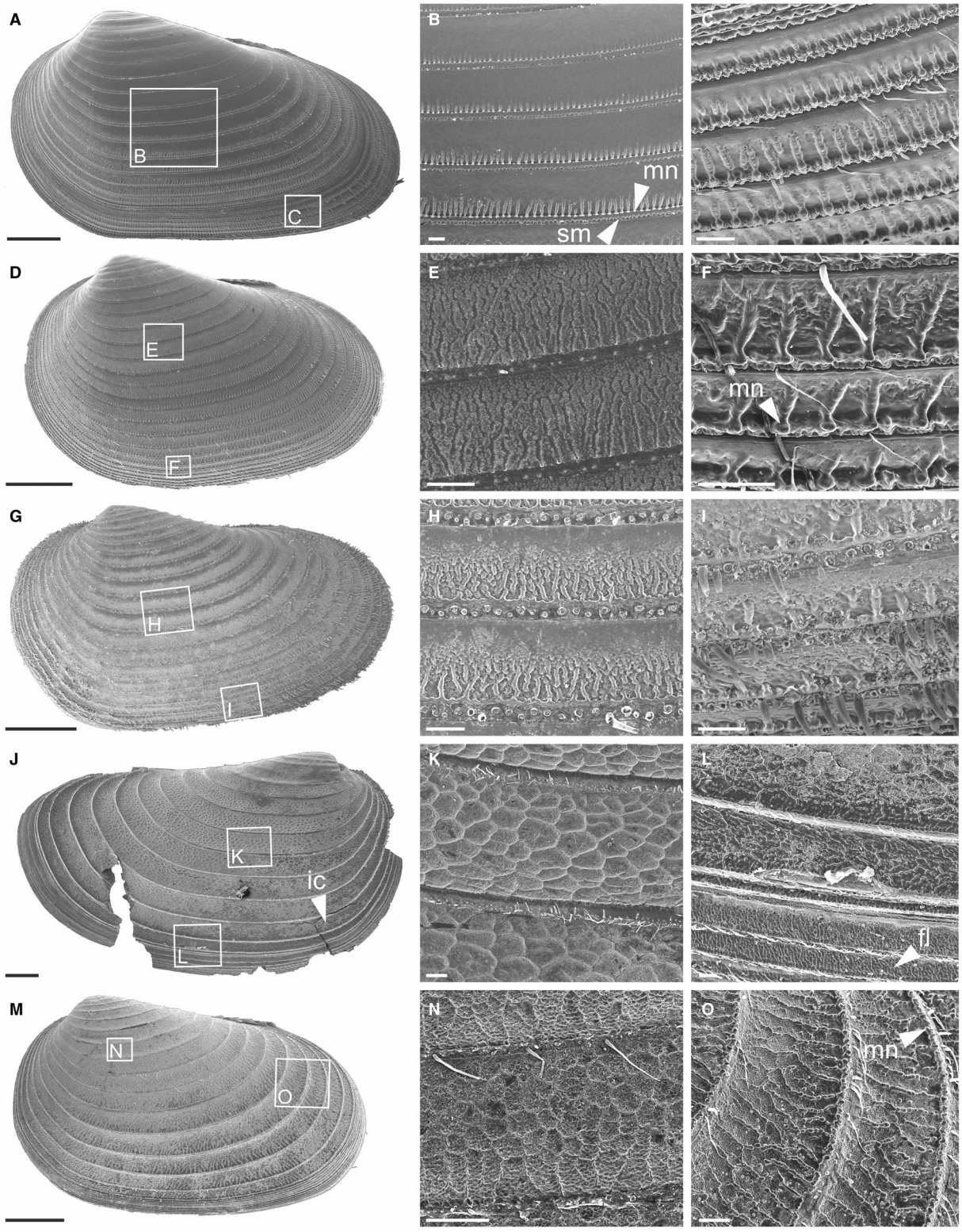

FIG. 10. SEM images of species A, B, C, K and M. A–C, species A, P.91454, male; drying of carapaces for SEM analysis led to the deformation of the ventral margin. D–F, species B, P.91496, male. G–I, species K, P.91581, male. J–L, species C, P.91285; individual not part of the morphometric analyses. M–O, species M, P.91157, male. *Abbreviations*: fl, very fine anastomosing lirae; ic, incipient growth band crowding; mn, nodules arranged in a moniliform way; sm, serrate margin. Scale bars represent: 1 mm (A, D, G, J, M); 100 μm (B, C, E, F, H, I, K, L, N, O).

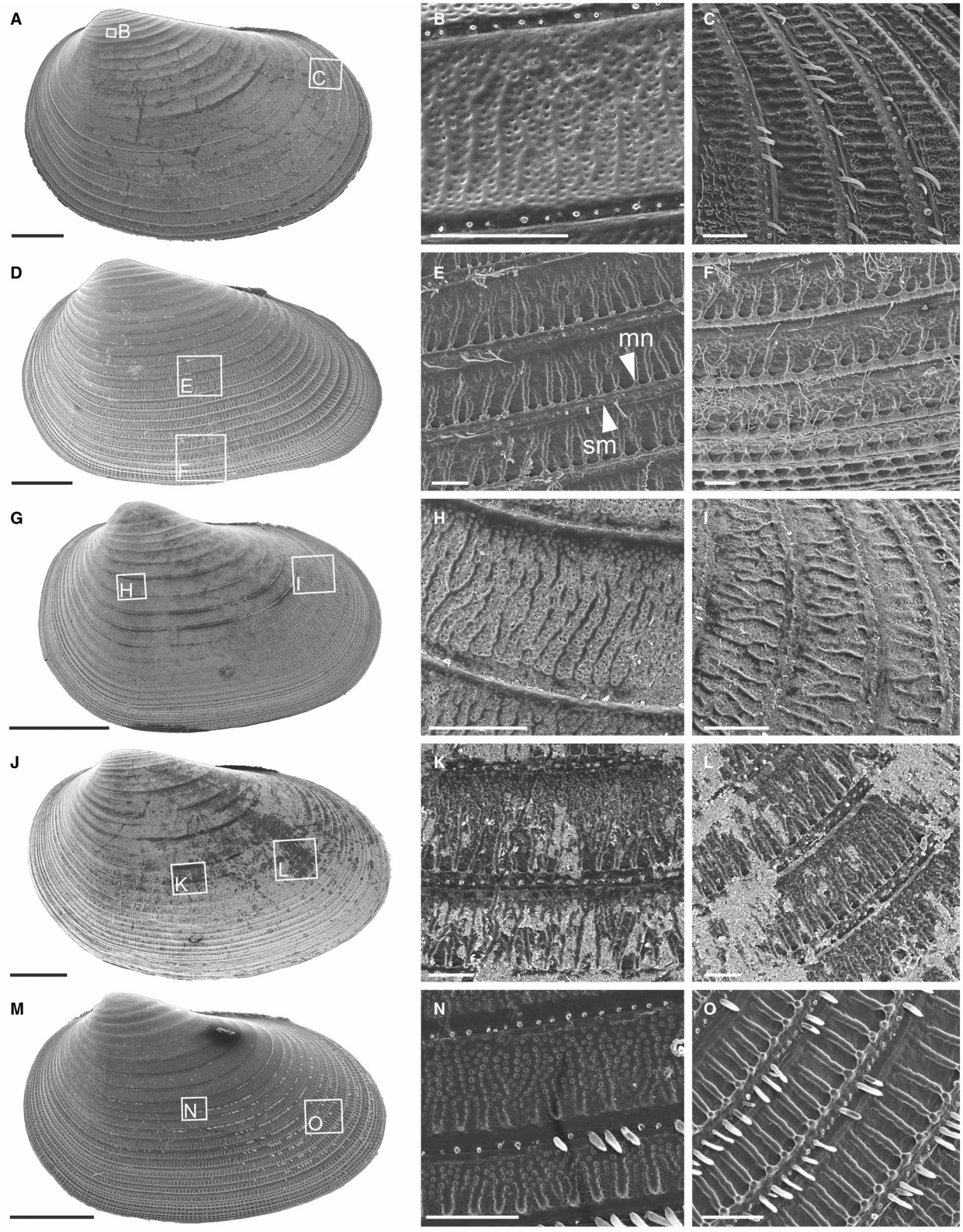

FIG. 11. SEM images of species N, Q3, Q5, O, and S. A–C, species N, P.91207, male. D–F, species Q3, P.91716, male. G–I, species Q5, P. 91739, male. J–L, species O, P.91628, male. M–O, species S, P.91797, male. *Abbreviations*: mn, nodules arranged in a moniliform way; sm, serrate margin. Scale bars represent: 1 mm (A, D, G, J, M); 100 μm (B, C, E, F, H, I, K, L, N, O).

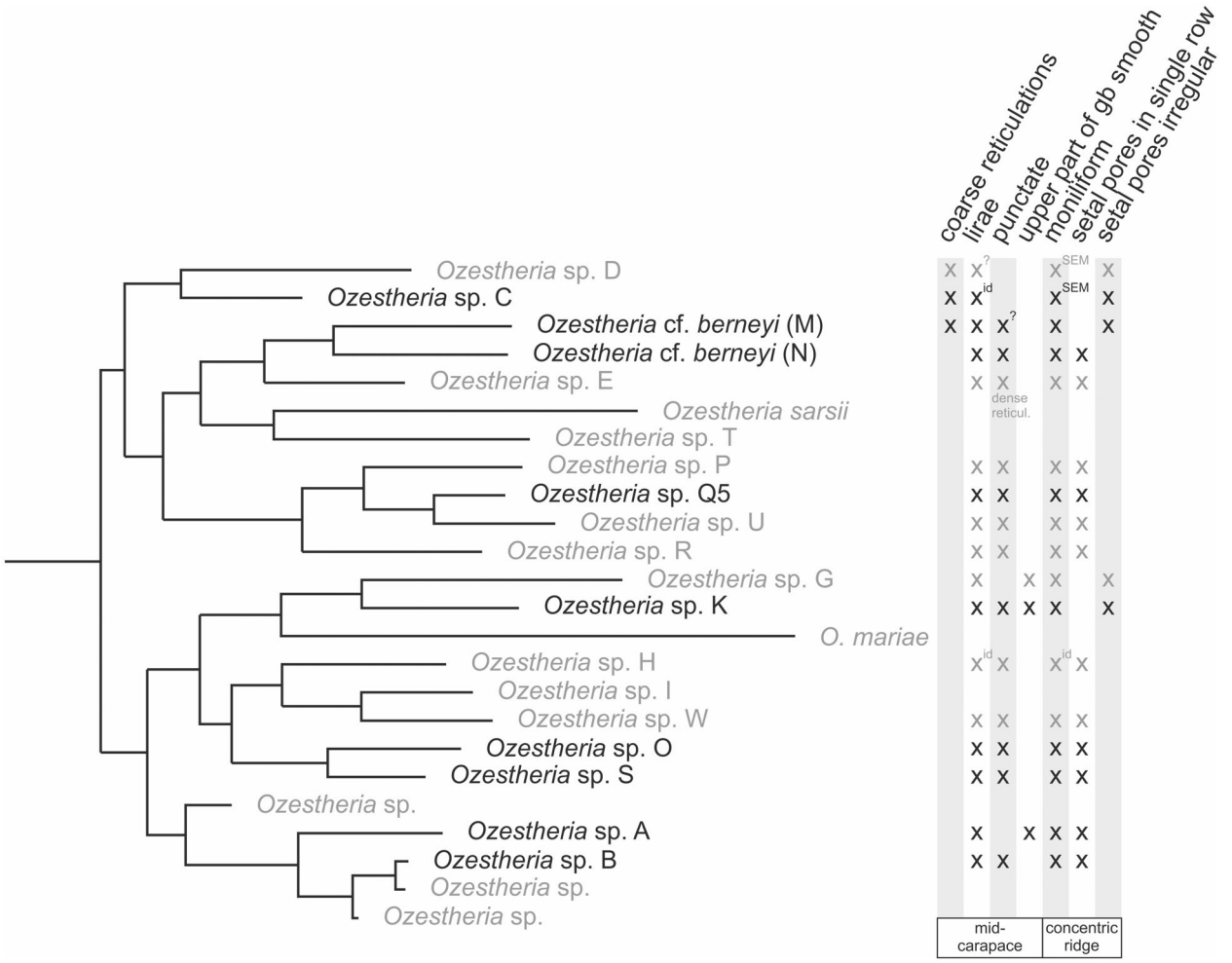

FIG. 12. Phylogenetic relationships of *Ozestheria* species (modified from Schwentner *et al.* 2020) and general ornamental features on growth bands of the mid-carapace region and on concentric ridges. Individuals of the examined species (black) were quantitatively inspected under incident light. Other species (grey) are based on the assessment of respective SEM images. The term 'moniliform' indicates the presence of nodules arranged in a moniliform way on the upper part of concentric ridges. *Ozestheria* sp. Q3 was not included by Schwentner *et al.* (2020), but is probably most closely related to Q5. *Abbreviations*: gb, growth band; id, indistinct but present; SEM, visible on scanning electron microscope images; punctate ornamentation of *Ozestheria* sp. E can also be interpreted as dense reticulation.

broken lines on growth bands of incipient carapace crowding, becoming progressively irregular on the following growth bands. Growth bands of the secondary growth phase are covered in very fine, anastomosing lirae (arrowhead in Fig. 10L).

4. (M; Fig. 10M–O): Early growth bands with shallow reticulations, followed by coarse reticulations filled with fine second-order reticulation/punctae (SEM observation). Upper part of following growth bands with reticulations which extend into nodulous, subparallel lirae ventrally, lirae become longer and dominant on progressing growth bands.
5. (N; Fig. 11A–C): Early growth bands either granular or with shallow reticulations. Following growth bands with weakly developed, anastomosing and reticulating lirae. Reticulations not as regular as in lineage M. Small punctae visible on various growth bands in some individuals
6. (B; Fig. 10D–F): Early growth bands punctate. In following growth bands subparallel, sometimes anastomosing lirae forming ventrally. In progressing growth bands lirae become longer and more dominant, from about mid-carapace only lirae and no punctae visible. In contrast to Type 7, upper part of mid-carapace growth bands nodulous (under SEM), leading to a granular appearance under incident light.
7. (O, Q3, Q5, S; Fig. 11D–O): Early growth bands dominated by punctae, forming minute reticulations on the larval valve. In following growth bands subparallel, sometimes anastomosing lirae forming on the lower part. In progressing growth bands lirae become longer and more dominant, from about mid-carapace (in

TABLE 5. Comparison of species and clusters of the realistic example C + K + M + N and the preliminary group B + K + N + Q5.

	Cluster 1	Cluster 2	Cluster 3
C	0	0	62
K	2	48	0
M	29	0	0
N	40	0	0
B	33	0	0
K	2	3	45
N	0	40	0
Q5	37	1	2

some individuals already at *c.* ¼, depending on ontogenetic stage) only lirae and no punctae visible. Lirae in most individuals weakly developed, but strongly in others (e.g. S individual in Fig. 11).

Ornamental patterns and phylogenetic relationships in Ozestheria. Under incident light, ornamentation is distinct for A, C, K, M and N, while B, O, Q3, Q5 and S are difficult to distinguish. Of the latter five species, the SEM image of B is separable from O, Q3, Q5 and S in yielding a nodulous upper part on mid-carapace growth bands, while crowded growth bands of B are very similar to species of Type 7, which represents a common ornamentation type in *Ozestheria* species that occurs across different phylogenetic groups (Fig. 12).

The ten examined *Ozestheria* species feature: (1) punctae; (2) reticulations; (3) smooth areas on growth bands; (4) the appearance of liral ornamentation during ontogeny; (5) nodules arranged in a moniliform way on the upper part of concentric ridges (= extensions of lirae that become narrower or obliterate towards concentric ridges and end in an enlarged nodule); and (6) setal pores arranged in a single row along the lower margin of concentric ridges or irregularly on concentric ridges, corresponding setae may be of different length and width.

Punctae are present in types 4(?)–7 (species B, M(?), N, O, Q3, Q5 and S), hence, across different phylogenetic groups (Fig. 12). SEM observations show that early ontogenetic stages are smooth in A, smooth to granular followed by punctate ventrally on the larval valve in B, reticulated in C, K, M, N, and transitional between minute reticulation and coarse punctae in Q3, Q5 and S (outer carapace layer of O chipped off). The arrangement of setal pores is generally regular in *Ozestheria* species, forming the serrate lower margins of concentric ridges. But there are also forms where setal pores are irregularly arranged on concentric ridges (species C, M, K).

At first glance, Type 3 (C) lacks several of these traits, however, the SEM image shows a transition from reticulation to thin lirae in the posterodorsal region of the carapace, a row of small nodules arranged in a moniliform way along the upper part of the concentric ridges, and irregularly arranged thin setae along the concentric ridges of later ontogenetic stages. Type 4 (M) is intermediate between Type 3 and all other ornamentation types, with coarse reticulations and short nodulous liral extensions in mid-carapace that each end in a nodule on the concentric margin. Long and thin setae of species M are aligned along the ventral part of the concentric ridge but less regularly than in its sister species N (Type 5), which yields narrow lirae and equidistant, setal pores.

Ornamental patterns are highly variable in *Ozestheria*, including special cases (species D, Fig. 12) that lack the described transition to liral ornamentation, showing a predominance of reticulations.

H3: Is it possible to distinguish species without prior information?

Four species. The gap statistic identified three meaningful clusters in the realistic example C + K + M + N (Table 5; Fig. 13; Hethke *et al.* 2022, Data_analysis S8, 2.2), of which species C (cluster 3) and K (cluster 2) are well separated, while the sister species M + N form a morphogroup (cluster 1).

Likewise, the gap statistic for the B + K + N + Q5 cluster dendrogram indicates three main clusters (Table 5), of which species N (cluster 2) and K (cluster 3) are well separated, while species B and Q5 (cluster 1) form a morphogroup.

Ten species. The gap statistic indicates five meaningful clusters in the dendrogram containing individuals of ten species (Hethke *et al.* 2022, Data_analysis S8, 2.3). Species C is well separated by cluster 5 (Table 6), and the sister group M + N by cluster 2. Species B, K, O and S are largely confined to single clusters, however these also contain a larger number of individuals of additional species. Sister pair O + S defines cluster 3, which contains individuals of additional species, owing to the central position of O + S in the *Ozestheria* morphospace. A, Q3 and Q5 are poorly separated. 67.5% of the Q5 individuals are clustered with K, which is surprising considering the excellent separation of the two species in previous analyses of this study (98.5% mean accuracy in H1b, reduced set), but it also indicates a morphological separation of Q5 from Q3. In summary, only two of the five clusters are meaningful: cluster 2 (sister group M + N) and cluster 5 (species C). There is no obvious pattern explaining the assignment of species to multiple clusters, for example, males and females are not assigned to separate clusters.

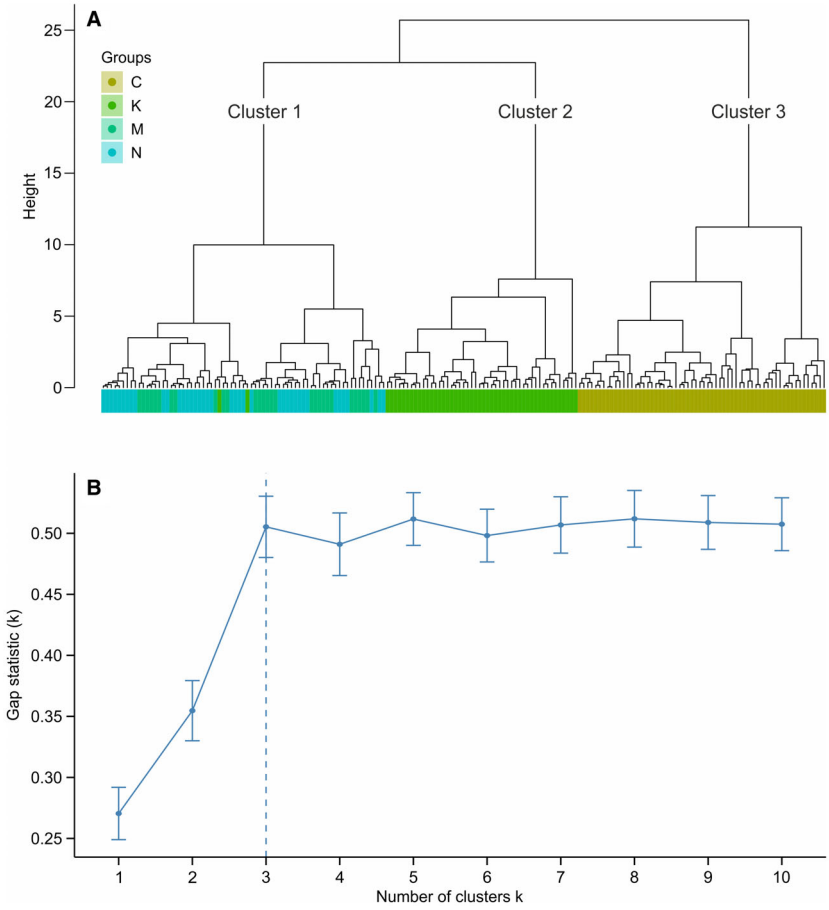

FIG. 13. Cluster dendrogram (A) and gap statistic (B) for C + K + M + N, reduced set. Clusters numbered as in Table 5. The blue dashed line indicates the optimal number of clusters.

TABLE 6. Comparison of species and clusters of the ten examined species.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
A	9	0	19	1	0
B	30	0	2	1	0
C	0	0	0	0	62
K	5	2	0	43	0
M	0	25	4	0	0
N	0	40	0	0	0
O	3	0	27	0	0
Q3	11	0	10	2	0
Q5	10	0	3	27	0
S	6	3	60	0	0

DISCUSSION

Methodological implications

Method comparisons of our study allow for general recommendations regarding: (1) the number of harmonics in Fourier shape analyses of Spinicaudata; (2) the choice of LDA-method; and (3) whether size and shape should be combined into a single dataset.

1. An increase in the number of harmonics does not increase model performance, as one would intuitively assume. Instead, highest accuracy values are reached when setting HANGLE to 8 harmonics.
2. Three of the four different LDA-methods provided by R-function *lda*() are suitable for species discrimination ('moment', 'mle' and 't'). They yield similar overall mean accuracies for the pairwise results of the size and shape analyses. But we recommend rejecting 'mve' for such morphometric analyses, as it performed least well in all LDA of this study.
3. Model performance of the shape analysis is generally higher than that of the size analysis, but there are exceptions such as pair M + N, which is better separated by linear variables (mean accuracy, reduced set: size = 91.6%, shape = 88.0%, combined = 89.9%). The combined analysis of size and shape (H1b) increases model performance to some degree (overall by 0.9% and 0.7% compared to the size and shape-reduced sets, respectively). However, the higher

number of variables needed for this combined analysis requires higher sample sizes. Also, as the results are comparable to the performance of the shape analysis, it is worth considering the time effort needed for assembling the size dataset. If both size and shape datasets are available, our H1b procedure will lead to higher mean accuracy values.

Are Ozestheria *species morphologically distinct? (H1)*

Intraspecific variability. Our *Ozestheria* dataset represents a time-averaged set of specimens of various ecophenotypes, ontogenetic stages and sex. Generally, environmental parameters greatly affect carapace size and shape (e.g. Rogers *et al.* 2012; Huang & Chou 2015, 2017; Hethke & Weeks 2020; Hethke *et al.* 2021). For our dataset, individuals were collected from 108 different localities (Fig. 1), but each pond is represented by only few individuals, so we are unable to test the contribution of single settings to model performance. Hence, in our study, mean accuracy values reflect a range of ecophenotypes that give a good representation of morphological disparity in *Ozestheria* species. Ontogenetic variability also strongly affects carapace morphology (Brown *et al.* 2014). As expected, the exclusion of juveniles ('reduced set') notably increased mean accuracy values, indicating allometric growth in *Ozestheria* species. Also, nine out of ten *Ozestheria* species are clearly sexually dimorphic but male and female morphologies strongly overlap (Fig. 8). Thus, sexual dimorphism contributes to overall morphological disparity in *Ozestheria* species, but remains low compared to interspecific variation within the genus *Ozestheria* or to dimorphism in, for example, limnadiid species (Brown *et al.* 2014). Pairwise comparisons of female-only or male-only analyses are sensitive to small and unequal sample sizes between males and females of a species. A combined-sex analysis of males and females yields mean accuracies that are mostly intermediate between or higher than the corresponding female-only and male-only values (Hethke *et al.* 2022, Data_analysis S5, 4.1.6). In the absence of a large single-sex sample size, a combined analysis will yield the most representative results, which is good news for palaeontological studies where sex is unknown.

Species distinction based on size and shape. All species are morphologically distinct, despite the intraspecific variability in the dataset, supporting the species status of each of these previously only genetically differentiated species, including the closely related species Q3 and Q5. However, nine of the ten species occupy similar positions in the *Ozestheria* morphospace (Fig. 9), requiring a more detailed evaluation of the size and shape relationships between them. The four pairs in a sister-group relationship (A + B, M + N, Q3 + Q5, O + S) overlap in morphospace and yield lower mean accuracies compared to more distantly related species pairs, so it can be assumed that closely related species are generally similar in outline. But there are also convergent morphologies in the dataset (e.g. Q3 + S).

Generally, a drop in model performance may reflect sister-group relationships (A + B, Q3 + Q5 and O + S), an intermediate occupation of the *Ozestheria* shape space (species O, Q3 and S), and, in the case of Q3, also sample size bias (LD models for combinations comprising Q3 were built on less individuals in the training set). Hence, insufficient sampling will affect model performance. We also tested the potential of length ranges in species discrimination, which distinguish between species pairs in 32 of 45 cases, rendering length an important auxiliary diagnostic feature.

Is carapace ornamentation diagnostic? (H2)

Intra and interspecific variability. Ornamental patterns are diagnostic for A, B, C, K, M, N, and group O + Q3 + Q5 + S (seven types), supporting the concept that ornamentation is species-diagnostic in many cases (e.g. Zhang *et al.* 1976; Gallego 2010; Morton *et al.* 2017; Hethke *et al.* 2018; Li & Wu 2021; Li 2022; Sun & Cheng 2022). Although ornamental patterns are stable within single *Ozestheria* species, features may be weakly or strongly pronounced, which may reflect ecophenotypic variation or sexual dimorphism. In limnadiids, ornamental features strongly vary within species and are susceptible to environmental change. For example, algal growth on carapaces may trigger punctate surfaces (Rogers *et al.* 2012). According to Sun & Cheng (2022), adult-stage carapace ornamentation may differ between sexes, while it is similar in juvenile males and females. Sun & Cheng (2022) further distinguished between seven basic and ten combinational ornamental types. The examined *Ozestheria* species from Australia correspond to two of the ten proposed combinational types and to two new combinational types (terms 'upper' and 'lower' in Sun & Cheng 2022, changed to 'dorsal' and 'ventral', respectively): *Ozestheria* types 3 and 4 = 'large reticulations (dorsal), radial lirae (ventral)'; *Ozestheria* types 5–7 = 'punctae (dorsal), radial lirae (ventral)'. Some individuals of *Ozestheria* sp. D (Fig. 12) correspond only to the 'large reticulation' basic type.

Ozestheria Type 1 represents a new combinational type translating to 'smooth (dorsal), radial lirae (ventral)' and

a new basic type that translates to 'smooth-liral transitional ornamentation'. *Ozestheria* Type 2 is a new combinational type of 'small reticulations and punctae dorsal, smooth-liral transitional ornamentation ventral'.

Higher taxonomic levels. The observed diversity of ornamental features in *Ozestheria* species (combinations of punctae, reticulations, smooth areas and lirae) is especially interesting for the question of whether ornamental features are meaningful for higher taxonomy (Gallego 2010). Fossils of the suborder Spinicaudata consist of six superfamilies, of which the assignment of the diplostracan Leaioidea to Spinicaudata is uncertain (Astrop & Hegna 2015). Superfamily diagnoses (summarized by Astrop & Hegna 2015, and publications therein) are based on general shape, simplified to five stylized carapace shapes (in case of *Ozestheria* 'cyziciform', referring to the general shape of modern cyzicids), and on the presence or absence of additional major characters such as a recurvature of growth lines at the posterodorsal margin (e.g. in Vertexioidea) or the presence of radial ribs (e.g. in Afrograptioidea and Estherielloidea). Though growth-band ornamentation is mentioned in superfamily diagnoses (in Eosestherioidea and Estheriteoidea), it remains unspecific due to the diversity of ornamental features found in the various families of these two superfamilies.

The origin of the modern family Cyzicidae, which *Ozestheria* belongs to, is disputed. Zhang *et al.* (1976, p. 86) highlighted two possible candidates: Afrograptidae of the Afrograptioidea (with which the cyzicid genus *Caenestheriella* was associated, now divided between *Cyzicus* and *Ozestheria*) and Euestheriidae (Eosestherioidea). Astrop & Hegna (2015, fig. 2) is a reproduction of this analysis, but associates Cyzicidae with Aquilonoglyptidae of the Eosestherioidea, instead of Eustheriidae. In their own analysis, Astrop & Hegna (2015, fig. 6) grouped the cyzicids with the Eosestheriidae of the Eosestherioidea. Even though several *Ozestheria* species were formerly assigned to *Caenestheriella* (A, B, K, Q3, Q5, O, S; Schwentner *et al.* 2015*a*), Afrograptidae can be excluded as a candidate for *Ozestheria*, as the herein examined species lack the stout radiating costae diagnostic of Afrograptioidea.

In *Ozestheria* sp. C, D and M we find 'strong defined polygonal reticulation' (Astrop & Hegna 2015), supporting the placement of the Cyzicidae in Eosestherioidea. But the proposed superfamily diagnosis requires the carapace to be small containing 'no recurved growth lines or saw-toothed dorsal margins' (Zhang *et al.* 1976, in Astrop & Hegna 2015). These two specifications would exclude *Ozestheria*, and therefore Cyzicidae, from this superfamily. Growth bands of several *Ozestheria* species (e.g. B) are recurved at the dorsal posterior margin creating a 'saw-toothed' appearance that is otherwise diagnostic of the superfamily Vertexioidea and the therein included Palaeolimnadiopseidae (although less pronounced in *Ozestheria*), supporting the proposal of Astrop & Hegna (2015) that this feature cannot sustain a family.

Ozestheria species are similar in ornamentation to species of various families among the superfamily Eosestherioidea (*sensu* Astrop & Hegna 2015). The ornamentation pattern of *Ozestheria* sp. H (larval ornamentation not preserved) is reminiscent of the described family diagnoses of the Aquilonoglyptidae and Triglyptidae, in yielding punctae in mid-carapace that transition to radial lirae with punctate intercalations on growth bands of later ontogenetic stages (Li 2022), supporting a possible close relationship of aquilonoglyptids/triglyptids and cyzicids as proposed by Zhang *et al.* (1976, p. 86). In another case (*Ozestheria* sp. M), the presence of reticulated ornamentation on growth bands of early ontogenetic stages that transition into lirae on the lower part of growth bands is reminiscent of the Eosestheridae, which may support a possible close relationship of cyzicids and eosestheriids (Astrop & Hegna 2015). The ornamental spectrum in *Ozestheria* also includes exclusively coarse reticulations in several individuals of D, similar to the family Loxomegaglyptidae as described from Oligocene deposits by Gallego & Mesquita (2000). The thin lirae in the posterior region of species C, whose ornamentation is otherwise dominated by coarse reticulations, indicate that lirae were progressively reduced during the evolution of the lineage containing species C and D.

Generally, we find that a transformation of main ornamental features may occur along short evolutionary timescales in *Ozestheria*. Such a fast evolutionary transition in ornamentation is, for example, well documented by sister group A + B (transition from punctate to smooth). The resulting ornamental diversity between *Ozestheria* species can be higher than what is currently interpreted as between-family ornamental diversity in fossils (compare family diagnoses of Triglyptidae and Aquilonoglyptidae; Wang 2014; Li 2022). In *Ozestheria*, the presence or absence of punctae, for example, does not represent a strong diagnostic character for generic discrimination. In limnadiids, punctae may even represent an entirely ecophenotypic feature (Rogers *et al.* 2012). Punctae are present in various phylogenetic lineages of *Ozestheria* (Fig. 12); using this character for the *Ozestheria* genus diagnosis would lead to a polyphyletic group. Of course, it is unclear whether our findings within *Ozestheria* can be easily transferred to other spinicaudatan genera and families, however, the usefulness of specific ornamental features for higher-level classifications may be questionable while others are better suited, and this should be considered in future revisions of fossil families.

Is it possible to distinguish species without prior information? (H3)

Morphotype recognition performs well for pools containing up to four species, which represents the maximum number of *Ozestheria* species known to co-occur, while the common *Ozestheria* diversity ranges between one and three species per pool (Schwentner *et al.* 2015*b*). We find that the gap-statistic (Tibshirani *et al.* 2001) is a good tool for identifying the optimal number of clusters: This approach captures a species diversity of three in the B + K + N + Q5 dendrogram with two clusters representing species and one cluster representing a polyphyletic morphospecies made up of B and Q5 individuals. In a four-cluster solution, individuals of these two species are distributed among two clusters, rendering the three-cluster solution proposed by the R-function *clusgap()* the better choice. Note that species B and Q5 do not commonly co-occur in reality. They have been collected from 15 and 14 different localities, respectively, with only one reported co-occurrence. In the realistic example (C + K + M + N), our H3 procedure also captures a species diversity of three comprising two species and one sister group. Hence, we are confident that the explorative approach will yield useful results for ecosystem and palaeocommunity analyses of single water bodies or confined geographical areas.

However, its value for species-level taxonomic studies comprising ten or more species is limited. For instance, cluster 4 in Hethke *et al.* (2022, Data_analysis S8, 2.3.4.1) groups species K and Q5, which is surprising as they are otherwise excellently separated by LDA. Differences in ornamentation between these two poorly separated species (by H3) indicate the potential of a statistical routine that combines morphometric and ornamental features, which would isolate species Q5 in cluster 4. In general, the approach tends to slightly underestimate species diversities in four-species assemblages. In the full dataset, only half of the overall diversity was captured. Not only were morphologically similar species clustered together (e.g. M and N) but some species were assigned to multiple clusters. Thus, relying solely on size or shape analyses may underestimate species diversities in fossil assemblages. Also, carapace distortion from sediment compaction will affect the cluster result to some degree. Most individuals are preserved in two dimensions, but distinct radial 'wrinkles' form on the carapace during fossil diagenesis which accommodate some of the compactional distortion that would otherwise affect the carapace outline.

Diagnostic power of carapace features in Ozestheria

The identification of fossils primarily relies on carapace size, shape, and ornamental patterns. Ornamentation is diagnostic for A, B, C, K, M, N, and group O + Q3 + Q5 + S (seven types), but the unresolved O + Q3 + Q5 + S group highlights the need for a combined analysis of size, shape and ornamental features. Curiously, this morphogroup yields similar ornamentation patterns and similar outlines. So, both ornamentation and shape converged within these two distantly related sister species pairs (O + S and Q3 + Q5), of which Q5 is best separated by length and shape.

Species C, K, M and N can easily be discriminated based on carapace features. In this species-set, carapace length represents an additional discriminator for C. Sister species Q3 and Q5 are better separated by size and shape variables than by ornamentation, while sister species A and B are best separated by ornamental patterns. The sister species pair O + S occupies a narrow but intermediate position in the *Ozestheria* size and shape morphospace that is blanketed by Q3, which yields a generally high morphological disparity and similar ornamentation (Type 7). However, the Q3 individual studied under the SEM contains very distinct punctae on concentric ridges of later ontogenetic stages, which would separate this species from O and S, but we do not know whether this feature is present in all Q3 individuals. Hence, Q3 + O + S forms a morphospecies based on ornamentation and morphometric analyses of the carapace.

Applications of the presented methods

The study of outline characteristics is especially interesting for the classification of larger sets of individuals, as the number of individuals that can be studied under the SEM is often limited, and as fossils are often preserved with the smooth, internal side up. The fully explorative approach (H3, PCA and cluster analysis) performs well for single water bodies or a confined geographic setting but it fails to separate species sets with higher species numbers. This is in strong contrast to the overall highly successful assignment of unclassified individuals to predefined species (see H1 approach). Thus, the availability of a training dataset containing individuals that could be pre-assigned to species will greatly improve the performance.

In a real-life scenario, such a training dataset might be obtained by studying a subset of extant species genetically (similar to our approach) or by assessing ornamental features of a subset of individuals. The latter can be applied to extant as well as fossil species and may require SEM observations on several individuals. With these, a model can be trained and then used to classify the remaining individuals that cannot be studied under the SEM. This would be similar to our H1 approach based on LDA. This way, the computer learns outline characteristics of specific

groups and applies these to classify further individuals (see Hethke *et al.* 2022, Data_analysis S5–S7 for technical guidance). In the case of *Ozestheria*, this would lead to a higher number of correctly classified individuals, but also to a lumping of O, Q3, Q5 and S (if these species were grouped by ornamentation). Likewise, our H1-approach can be used by biologists to assign outlines of previously described type material to the most likely cryptic species.

The described overall good agreement between molecular and palaeontological methods indicates that we can be confident about cyzicid diversity described from the fossil record at species level, with the widely distributed genus *Ozestheria* as the model group (reported occurrences in Australia, China, India, Madagascar, Mongolia, Niger, South Africa and Thailand; Schwentner *et al.* 2015*a*, 2020; Meyer-Milne *et al.* 2020; Padhye *et al.* 2020; Sun & Cheng 2022). Our results further underscore the high taxonomic value of ornamental and morphometric carapace features in *Ozestheria* and, therefore, highlight the value of the herein employed techniques for taxonomic studies on extant spinicaudatan species. Nevertheless, morphologically similar species may be erroneously lumped into a single species, and these may not even be sister species. The evolutionary time required for the loss or appearance of main ornamental features such as punctae or reticulations can be surprisingly low and the ornamental diversity among relatively closely related species can be much larger than anticipated, demonstrating that descriptions of higher taxonomic ranks (genus and family-level) need to be revisited.

CONCLUSIONS

1. *Ozestheria* is a widely distributed extant genus of clam shrimp with a biodiversity hotspot in Australia. We used this genus as a model group for cyzicid species delimitation by testing whether carapace traits, important for the identification of fossil clam shrimp, are species-diagnostic. The ten species considered here include four pairs in sister-group relationship, three of which occur sympatrically or syntopically.
2. Our dataset includes 481 clam shrimp from 108 different localities, so studied individuals cover a wide spectrum of ecophenotypes, age and sex. Nevertheless, geometric morphometrics of the carapace indicate that all ten examined *Ozestheria* species are morphologically distinct, but they occupy overlapping regions in the *Ozestheria* morphospace.
3. 93.8% individuals of unknown species identity ('test set') can be successfully classified when an LD model is trained with individuals of known species identity, using species as the grouping factor and morphometric variables as the discriminators (e.g. Fourier coefficients). For this machine learning routine, we recommend a training set of at least 30 identified individuals per species. Thus, the routine requires molecular or, in case of fossils, ornamental data.
4. Six of the ten *Ozestheria* species yield species-specific ornamentation types, one type is shared by four species. Species delimitation based solely on ornamental patterns would recognize seven out of ten species.
5. Single pools commonly contain one to three *Ozestheria* species, and there is one pool with four co-occurring species, which we used to test if it is possible to distinguish species without prior species assignments by applying hierarchical cluster analysis to PC scores of the size and shape datasets. We successfully distinguished a diversity of three of four species by using the so-called gap-statistic; the two species grouped to represent one morphotype are in sister-group relationship.
6. Our overall results underscore the taxonomic value of carapace traits, if length, size, shape and ornamentation are combined. We find that carapace features vary extensively among species of a single genus, and several *Ozestheria* species would have been assigned to different genera or families if these had been fossils.

Acknowledgements. We thank Kathrin Philipps-Bussau, CeNak, Universität Hamburg (Germany) for technical assistance during the photographic documentation of the *Ozestheria* collection and Stephen Keable of the Australian Museum (Sydney) for kindly providing the collection material and his enduring patience and support. Brian Timms and Stefan Richter (University of Rostock) laid the foundation for this work and helped collect the studied material, for which we are most thankful. We are grateful for comments and suggestions of D. Christopher Rogers, Tim Astrop, and an anonymous reviewer, and we thank the editors Imran Rahman and Sally Thomas for their support. This work was supported by the German Research Foundation (DFG; HE 7531/1-1). Open Access funding enabled and organized by Projekt DEAL.

Author contributions. **Conceptualization** M Hethke (MH), K Hartmann (KH), M Schwentner (MS); **Data Curation** MH, MS; **Formal Analysis** MH, KH, MS; **Funding Acquisition** MH; **Investigation** MH, M Alberti (MA), T Kutzner (TK), MS; **Methodology** MH, KH, MS; **Project Administration** MH, MS; **Resources** B Timms, S Richter, C Sieves, MS; **Validation** MH, KH, MS; **Visualization** MH, MS; **Writing – Original Draft Preparation** MH, MS; **Writing – Review & Editing** MH, KH, MA, TK, MS.

DATA ARCHIVING STATEMENT

Data and R-markdown files for this study are available in the Dryad Digital Repository: https://doi.org/10.5061/dryad.fbg79cnxs

Editor. Imran Rahman

REFERENCES

ASTROP, T. I. and HEGNA, T. A. 2015. Phylogenetic relationships between living and fossil spinicaudatan taxa (Branchiopoda Spinicaudata): reconsidering the evidence. *Journal of Crustacean Biology*, **35**, 339–354.

ASTROP, T. I., SAHNI, V., BLACKLEDGE, T. A. and STARK, A. Y. 2015. Mechanical properties of the chitin-calcium-phosphate 'clam shrimp' carapace (Branchiopoda: Spinicaudata): implications for taphonomy and fossilization. *Journal of Crustacean Biology*, **35**, 123–131.

BAIRD, W. 1849. Monograph of the family Limnadiadae, a family of entomostracous Crustacea. *Proceedings of the Zoological Society of London*, **17**, 84–90.

BROWN, B. P., ASTROP, T. I. and WEEKS, S. C. 2014. Post-larval developmental dynamics of the spinicaudatan (Branchiopoda: Diplostraca) carapace. *Journal of Crustacean Biology*, **34**, 611–617.

CLAUDE, J. 2008. *Morphometrics with R*. Springer.

CRAMPTON, J. S. and HAINES, A. J. 1996. Users' manual for programs HANGLE, HMATCH, and HCURVE for the Fourier shape analysis of two-dimensional outlines. Institute of Geological & Nuclear Sciences, Science Report.

DADAY DE DEÉS, E. 1915. Monographie systématique des Phyllopodes Conchostracés. *Annales des Sciences Naturelles, Zoologie, série*, **9** (20), 39–330. [in Latin and French]

DEFRETIN-LEFRANC, S. 1965. Etude et révision de Phyllopodes Conchostracés en provenance d'U.R.S.S. *Annales de la Société Géologique du Nord*, **85**, 15–48.

GALLEGO, O. F. 2010. A new crustacean clam shrimp (Spinicaudata: Eosestheriidae) from the Upper Triassic of Argentina and its importance for 'conchostracan' taxonomy. *Alcheringa*, **34**, 179–195.

GALLEGO, O. F. and MESQUITA, M. V. 2000. First record of oligocene conchostracans (Tremembé formation, Taubaté Basin) from São Paulo, Brazil. *Journal of South American Earth Sciences*, **13**, 685–692.

GALLEGO, O. F., MONFERRAN, M. D., ZACARÍAS, I. A., JIMÉNEZ, V. C., BUSCALIONI, A. D. and LIAO, H. 2020. Clam shrimp fauna (Diplostraca-Spinicaudata and Estheriellina) from the Lower Cretaceous of Las Hoyas, Cuenca (Spain). *Cretaceous Research*, **110**, 104389.

GRUBE, E. 1865. *Ueber die Gattungen* Estheria *und* Limnadia *und einen neuen* Apus. Nicolaische Verlagsbuchhandlung, Berlin.

HAINES, A. J. and CRAMPTON, J. S. 2000. Improvements to the method of Fourier shape analysis as applied in morphometric studies. *Palaeontology*, **43**, 765–783.

HAMMER, Ø. and HARPER, D. 2006. *Paleontological data analysis*. Blackwell Publishing.

HAMMER, Ø., HARPER, D. A. T. and RYAN, P. D. 2001. PAST: paleontological statistics software package for education and data analysis. *Palaeontologia Electronica*, **4**, art. 4.

HEGNA, T. A. and ASTROP, T. I. 2020. The fossil record of the clam shrimp (Crustacea; Branchiopoda). *Zoological Studies*, **59**, 43.

HEGNA, T. A., CZAJA, A. D. and ROGERS, D. C. 2020. Raman spectroscopic analysis of the composition of the clam-shrimp carapace (Branchiopoda: Laevicaudata, Spinicaudata, Cyclestherida): a dual calcium phosphate-calcium carbonate composition. *Journal of Crustacean Biology*, **40**, 756–760.

HETHKE, M. and WEEKS, S. C. 2020. Population density effects on carapace growth in clam shrimp: implications for palaeontological studies. *Zoological Studies*, **59**, 33.

HETHKE, M., FÜRSICH, F. T., MORTON, J. D. and JIANG, B. 2018. Analysis of morphological variability in the clam shrimp *Eosestheria middendorfii* (Crustacea, Spinicaudata) from the Lower Cretaceous of China, and its implications for spinicaudatan taxonomy. *Papers in Palaeontology*, **4**, 21–53.

HETHKE, M., WEEKS, S. C., SCHÖTTLE, V. and ROGERS, D. C. 2021. Preliminary study of temperature effects on size and shape in the modern spinicaudatan *Eulimnadia texana* (Crustacea: Branchiopoda). *Zoological Studies*, **60**, 2.

HETHKE, M., HARTMANN, K., ALBERTI, M., KUTZNER, T. and SCHWENTNER, M. 2022. Data from: Testing the success of palaeontological methods in the delimitation of clam shrimp (Crustacea, Branchiopoda) on extant species. *Dryad Digital Repository*. https://doi.org/10.5061/dryad.fbg79cnxs

HUANG, W. and CHOU, L. 2015. Temperature effect on development and reproduction of the androdioecious clam shrimp, *Eulimnadia braueriana* (Branchiopoda: Spinicaudata). *Journal of Crustacean Biology*, **35**, 330–338.

HUANG, W. and CHOU, L. 2017. Temperature effects on life history traits of two sympatric branchiopods from an ephemeral wetland. *PLoS One*, **12**, e0179449.

KAUFMAN, L. and ROUSSEEUW, P. J. 1990. *Finding groups in data: An introduction to cluster analysis*. Wiley.

KOZUR, H. and SEIDEL, G. 1983. Revision der Conchostracen-Faunen des unteren und mittleren Buntsandsteins. Teil I. *Zeitschrift für Geologische Wissenschaften*, **11**, 295–423.

LI, G. 2022. New taxonomic features of fossil clam shrimp *Aquilonoglypta* (Spinicaudata) and its implication for origin of triglyptids. *Paleontological Research*, **26**, 8–17.

LI, G. and BATTEN, D. J. 2004*a*. *Cratostracus*? *cheni*, a new conchostracan species from the Yixian Formation in western Liaoning, north-east China, and its age implications. *Cretaceous Research*, **25**, 577–584.

LI, G. and BATTEN, D. J. 2004*b*. Revision of the conchostracan genera *Cratostracus* and *Porostracus* from Cretaceous deposits in north-east China. *Cretaceous Research*, **25**, 919–926.

LI, G. and BATTEN, D. J. 2005. Revision of the conchostracan genus *Estherites* from the Upper Cretaceous Nenjiang Formation of the Songliao Basin and its biogeographic significance in China. *Cretaceous Research*, **26**, 920–929.

LI, G. and WU, Z. 2021. Early Cretaceous clam shrimp *Yanjiestheria* (Spinicaudata, Crustacea) from eastern Liaoning, northeastern China. *Cretaceous Research*, **128**, 104962.

MAECHLER, M., ROUSSEEUW, P., STRUYF, A., HUBERT, M., HORNIK, K., STUDER, M., ROUDIER, P., GONZALEZ, J., KOZLOWSKI, K., SCHUBERT, E. and MURPHY, K. 2021. cluster: Cluster Analysis Basics and Extensions. R package v2.1.2. https://CRAN.R-project.org/package=cluster

MEYER-MILNE, E., MLAMBO, M. C. and ROGERS, D. C. 2020. Distribution of clam shrimps (Crustacea:

Laevicaudata and Spinicaudata) in South Africa, with new records from the Northern Cape Province. *Zoological Studies*, **59**, 39.

MIESCH, A. T. 1980. Scaling variables and interpretation of eigenvalues in principle component analysis of geologic data. *Mathematical Geology*, **12**, 523–538.

MORTON, J. D., WHITESIDE, D. I., HETHKE, M. and BENTON, M. J. 2017. Biostratigraphy and geometric morphometrics of conchostracans (Crustacea, Branchiopoda) from the Late Triassic fissure deposits of Cromhall Quarry, UK. *Palaeontology*, **60**, 349–374.

ORR, P. J. and BRIGGS, D. E. G. 1999. Exceptionally preserved conchostracans and other crustaceans from the Upper Carboniferous of Ireland. *Special Papers in Palaeontology*, **62**, 1–68.

PADHYE, S. M., INGALE, P. A. and VANJARE, A. I. 2020. Taxonomical account of the Indian cyzicid *Ozestheria annandalei* (Daday, 1913) (Branchiopoda: Spinicaudata) from peninsular India. *Zootaxa*, **4852**, 3.

PAPPAS, J. L., KOCIOLEK, J. P. and STOERMER, E. F. 2014. Quantitative morphometric methods in diatom research. *Nova Hedwigian, Beiheft*, **143**, 281–306.

RAYMOND, P. E. 1946. The genera of fossil Conchostraca – an order of bivalved Crustacea. *Bulletin of the Museum of Comparative Zoology at Harvard College*, **96**, 217–307.

R CORE TEAM. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

RIEDER, N., ABAFFY, P., HAUF, A., LINDEL, M. and WEISHÄUPL, H. 1984. Funktionsmorphologische Untersuchungen an den Conchostracen *Leptestheria dahalacensis* und *Limnadia lenticularis* (Crustacea, Phyllopoda, Conchostraca). *Zoologische Beiträge Neue Folge*, **28**, 417–444.

RIPLEY, B., VENABLES, B., BATES, D. M., HORNIK, K., GEBHARDT, A. and FIRTH, D. 2021. MASS: Support Functions and Datasets for Venables and Ripley's MASS. R package v7.3-54. https://cran.r-project.org/web/packages/MASS/index.html

ROGERS, D. C. 2020. Spinicaudata catalogus (Crustacea: Branchiopoda). *Zoological Studies*, **59**, 45.

ROGERS, D. C., RABET, N. and WEEKS, S. C. 2012. Revision of the extant genera of Limnadiidae (Branchiopoda: Spinicaudata). *Journal of Crustacean Biology*, **32**, 827–842.

ROHLF, F. J. 2018. tpsDig: digitize coordinates of landmarks and capture outlines. v2.31. Department of Ecology & Evolution, State University of New York at Stony Brook. http://www.sbmorphometrics.org/soft-dataacq.html

SCHOLZE, F. and SCHNEIDER, J. W. 2015. Improved methodology of 'conchostracan' (Crustacea: Branchiopoda) classification for biostratigraphy. *Newsletters on Stratigraphy*, **48**, 287–298.

SCHWENTNER, M., TIMMS, B. V. and RICHTER, S. 2011. An integrative approach to species delineation incorporating different species concepts: a case study of *Limnadopsis* (Branchiopoda: Spinicaudata). *Biological Journal of the Linnean Society*, **104**, 575–599.

SCHWENTNER, M., TIMMS, B. V. and RICHTER, S. 2012. Description of four new species of *Limnadopsis* from Australia (Crustacea: Branchiopoda: Spinicaudata). *Zootaxa*, **3315**, 42–64.

SCHWENTNER, M., TIMMS, B. V. and RICHTER, S. 2014. Evolutionary systematics of the Australian *Eocyzicus* fauna (Crustacea: Branchiopoda: Spinicaudata) reveals hidden diversity and phylogeographic structure. *Journal of Zoological Systematics & Evolutionary Research*, **52**, 15–31.

SCHWENTNER, M., JUST, F. and RICHTER, S. 2015*a*. Evolutionary systematics of the Australian Cyzicidae (Crustacea, Branchiopoda, Spinicaudata) with the description of a new genus. *Zoological Journal of the Linnean Society*, **173**, 271–295.

SCHWENTNER, M., TIMMS, B. V. and RICHTER, S. 2015*b*. Spinicaudata (Branchiopoda: Diplostraca) in Australia's arid zone: unparalleled diversity at regional scales and within water bodies. *Journal of Crustacean Biology*, **35**, 366–378.

SCHWENTNER, M., RABET, N., RICHTER, S., GIRIBET, G., PADHYE, S., CART, J.-F., BONILLO, C. and ROGERS, D. C. 2020. Phylogeny and biogeography of Spinicaudata (Crustacea: Branchiopoda). *Zoological Studies*, **59**, 44.

SHEN, Y. 2003. Review of the classification of the family Afrograptidae (Crustacea: Conchostraca). *Acta Palaeontologica Sinica*, **42**, 590–597.

STIGALL, A. L., BABCOCK, L. E., BRIGGS, D. E. G. and LESLIE, S. A. 2008. Taphonomy of lacustrine interbeds in the Kirkpatrick Basalt (Jurassic), Antarctica. *PALAIOS*, **23**, 344–355.

SUN, X. and CHENG, J. 2022. Sexually dimorphic ornamentation in modern spinicaudatans and the taxonomic implications for fossil clam shrimps. *Acta Palaeontologica Polonica*, **67**, 475–492.

TASCH, P. 1969. Branchiopoda. R128–R191. *In* MOORE, R. C. (ed.) *Treatise on invertebrate paleontology. Part R. Arthropoda 4 (1)*. University of Kansas Press & The Geological Society of America.

TASCH, P. 1987. Fossil Conchostraca of the southern hemisphere and continental drift, paleontology, biostratigraphy, and dispersal. *The Geological Society of America Memoirs*, **165**, 1–290.

TIBSHIRANI, R., WALTHER, G. and HASTIE, T. 2001. Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B*, **63**, 411–423.

TIPPELT, L. and SCHWENTNER, M. 2018. Taxonomic assessment of Australian *Eocyzicus* species (Crustacea: Branchiopoda: Spinicaudata). *Zootaxa*, **4410**, 401–452.

VENABLES, W. N. and RIPLEY, B. D. 2002. *Modern applied statistics with S*, Fourth edition. Springer. ISBN 0-387-95457-0

WANG, S. E. 2014. Triglyptidae fam. nov. and its significance in evolution and biostratigraphy. *Acta Palaeontologica Sinica*, **53**, 486–496. [in Chinese]

WISHKERMAN, A. and HAMILTON, P. B. 2018. Shape outline extraction software (DiaOutline) for elliptic Fourier analysis application in morphometric studies. *Applications in Plant Sciences*, **6**, e1204.

ZAHN, C. and ROSKIES, R. 1972. Fourier descriptors for plane closed curves. *IEEE Transactions on Computers*, **C-21**, 269–281.

ZHANG, W., CHEN, P. and SHEN, Y. 1976. *Fossil Conchostraca of China*. Science Press, Beijing. [in Chinese]