BMC Ecology and Evolution

RESEARCH

Open Access

Do morphometric data improve phylogenetic reconstruction? A systematic review and assessment

Emma J. Holvast[1*], Mélina A. Celik[2], Matthew J. Phillips[2] and Laura A. B. Wilson[1,3,4]

Abstract

Background Isolating phylogenetic signal from morphological data is crucial for accurately merging fossils into the tree of life and for calibrating molecular dating. However, subjective character definition is a major limitation which can introduce biases that mislead phylogenetic inferences and divergence time estimation. The use of quantitative data, e.g., geometric morphometric (GMM; shape) data can allow for more objective integration of morphological data into phylogenetic inference. This systematic review describes the current state of the field in using continuous morphometric data (e.g., GMM data) for phylogenetic reconstruction and assesses the efficacy of these data compared to discrete characters using the PRISMA-EcoEvo v1.0. reporting guideline, and offers some pathways for approaching this task with GMM data. A comprehensive search string yielded 11,123 phylogenetic studies published in English up to Oct 2023 in the Web of Science database. Title and abstract screening removed 10,975 articles, and full-text screening was performed for 132 articles. Of these, a total of twelve articles met final inclusion criteria and were used for downstream analyses.

Results Phylogenetic performance was compared between approaches that employed continuous morphometric and discrete morphological data. Overall, the reconstructed phylogenies did not show increased resolution or accuracy (i.e., benchmarked against molecular phylogenies) as continuous data alone or combined with discrete morphological datasets.

Conclusions An exhaustive search of the literature for existing empirical continuous data resulted in a total of twelve articles for final inclusion following title/abstract, and full-text screening. Our study was performed under a rigorous framework for systematic reviews, which showed that the lack of available comparisons between discrete and continuous data hinders our understanding of the performance of continuous data. Our study demonstrates the problem surrounding the efficacy of continuous data as remaining relatively intractable despite an exhaustive search, due in part to the difficulty in obtaining relevant comparisons from the literature. Thus, we implore researchers to address this issue with studies that collect discrete and continuous data sets with directly comparable properties (i.e., describing shape, or size).

Keywords Continuous characters, Geometric morphometrics, Landmarks, Phylogeny reconstruction, Quantitative characters, Shape

*Correspondence:
Emma J. Holvast
Emma.Holvast@anu.edu.au
Full list of author information is available at the end of the article

Introduction

Discrete morphological data have been the primary focus of traditional systematic methods for phylogenetic reconstruction. Even following the success of molecular phylogeny, discrete morphological data remain crucial for merging fossils into the tree of life and for calibrating molecular dating [1]. However, a major limitation of discrete morphological data is subjectivity in the definition of characters and character states, which in turn, may potentially introduce biases and mislead phylogenetic inference and divergence time estimation. The use of quantitative data e.g., geometric morphometric (GMM; shape) landmark data can allow for more objective integration of morphology for phylogenetic inference. Two- or three-dimensional landmarks can be used to capture the geometry of a biological structure, and shape differences can then be quantified using the relative positions of landmarks for each object in the sample [2–5] by scaling and superimposing the landmark configurations, typically using Procrustes analysis [6].

However, challenges associated with the application of GMM data to the inference of phylogenetic topology include the widespread non-independence (e.g., covariation) of landmarks due to functional or developmental correlation, which violates assumptions of standard trait evolution models. Therefore, a single landmark should not be treated as an independent character, but rather the entire landmark configuration [7]. To accommodate this concern, Álvarez-Carretero et al. [8] described a Bayesian method that extended the work of Parins-Fukuchi [9] by explicitly accounting for character correlation within a dataset of 3D cranium landmarks from carnivoran mammals, showing promise for the analysis of continuous characters for phylogenetic and divergence time estimation. Moreover, landmark configurations require Procrustes superimposition to remove non-shape differences (scaling, rotation, translation) between configurations in a dataset, allowing geometric comparisons to be made in a shape space. This procedure scales the landmark configurations to centroid size, and allows for further assessment of allometry, the covariation between shape and size. Allometry can be assessed within a geometric morphometric framework through regression of Procrustes superimposed landmarks against centroid size. Allometric variation can be difficult to tease apart from, but may also contribute to, true phylogenetic signal as the two sources of variation are often confounded [10].

Subjectivity may remain with GMM through the choice i.e., number and position, and manual placement of landmarks which can result in observer and measurement error [11–13]. However, the development of automated GMM (whole bone shape) methods [14–18] that aim to reduce observer and measurement error associated with manual landmark placement, and thus increase accuracy in approximations of shape, has received recent attention in evolutionary morphology research [12, 19–21]. Although, further methodological studies are required to determine the most appropriate approach(es) for individual taxa and for the goals of individual studies [11]. For example, some automated landmarking approaches [e.g., ALPACA; 19] may be better employed within species or among closely related species; while others [e.g., MALPACA; 21] are particularly suitable in broad phylogenetic contexts for landmarking the types of morphologically diverse (multi-taxa) samples commonly encountered in evolutionary studies.

If choosing to discretize GMM data for phylogenetic analysis [22, 23], another difficulty is in characterising and coding (i.e., discretizing) landmarks into character states. Moreover, arbitrarily delimiting discrete states from variation that is inherently continuous [24] (i.e., GMM; shape) can result in information loss [25]. However, improvements in discretization methods for continuous data have recently been proposed [22].

Methods for analysing undiscretized continuous morphological data i.e., as continuous quantitative characters in phylogenetic inference have been available using cladistic approaches such as parsimony [TNT; 25–29] and model-based methods such as maximum likelihood (ML) [CONTML in PHYLIP; 30] and Bayesian methods [31]. However, most attempts to integrate GMM data into phylogenetic inference have been criticised due to methodological concerns and unreliability [7, 32–34], although few studies have performed well in reconstructing 'true' phylogenetic relationships [35–38].

Analysis of continuous data (including GMM data) has been more commonly applied to taxonomy, especially in the delineation of species [e.g., 39, 40]. Among marsupial mammals cranial and/or dental linear measurement data have been applied in several studies on bandicoots (Peramelemorphia) [41, 42], whereas 3D landmark-based GMM of the cranium has been applied in *Antechinus* (Dasyuromporphia) [43].

Examples of reconstructing phylogeny from GMM data include the use of landmark configurations as continuous characters in what is referred to in the literature as 'landmark analysis under parsimony' (LAUP; spatial parsimony) or 'Phylogenetic Morphometrics' (PM) as proposed by Catalano et al. [27], Goloboff and Catalano [28], Catalano and Goloboff [26], [e.g., 23, 44–46]. Other studies have used landmark coordinates as continuous characters under parsimony [e.g., 47], squared-change parsimony [e.g., 23], minimum evolution [48, 49] [i.e., "Euclidean parsimony"; 50] [e.g., 32], and with Brownian motion modelling of evolutionary change under maximum likelihood (ML) [e.g., 23] and under Bayesian

inference [e.g., 31]; while others have used principal component (PC) scores [e.g., 24, 47, 51] and eigenscores [eigenshape analysis descriptors; e.g., 37] as continuous characters under parsimony, and PC scores as continuous characters under maximum likelihood [e.g., 24, 35, 39, 51]. However, the use of PC scores and landmark coordinates in phylogenetic inference have been heavily criticized by some researchers, with Adams et al. [33] regarding PCs of shape data as both inappropriate and ineffective as cladistic characters [33], and Varon-Gonzalez et al. [32] suggesting the unreliability of phylogenetic estimation from shape data such as landmark coordinates [32], see also criticisms by Catalano et al. [27]; Goloboff [52].

An alternative approach to character-based methods for reconstructing phylogeny from GMM data is to create trees based on distances (i.e., Procrustes, Euclidean) between taxa using cluster analysis. For example, neighbour-joining [NJ; 53] [e.g., 23, 24, 35, 44, 54–56] and unweighted pair group method with arithmetic mean [UPGMA; 58] [e.g., 24, 35, 44, 56, 57, 59, 60] or other methods such as minimum evolution [ME; 61] [e.g., 59] and maximum likelihood using flexibly weighted least squares methods [fWLS; e.g., 62, 63].

Another procedure is to estimate the positions of fossil taxa along a scaffold (molecular) phylogeny of extant taxa using quantitative data. For example, Revell et al. [64] described a maximum likelihood method for placing individual taxa into a phylogeny of extant taxa using continuous character data. This approach performed well but was limited to the placement of only extant and recently extinct taxa, and with only a single taxon placed at a time. Extending this approach, Parins-Fukuchi [36] presented a Bayesian method that places multiple fossil taxa on a phylogeny of extant taxa using quantitative characters modelled under Brownian Motion [65]. Importantly, their model treats branch lengths in terms of morphological divergence as opposed to time, and moreover allows the placement of long extinct fossil taxa to be estimated.

Probabilistic approaches (e.g., under Bayesian inference) allow estimation of branch lengths and evolutionary rates and can improve estimates of uncertainty, thus potentially improving the accuracy of morphological phylogenetics compared to cladistic methods (i.e., parsimony analysis) [9]. Recently, Zhang et al. [31] introduced a probabilistic total-evidence method for phylogenetic inference using multiple continuous (3D GMM landmark coordinates) characters in addition to discrete morphology and molecular data from both living and subfossil taxa, and fossil ages. Their method was implemented in a flexible Bayesian framework and was found to result in a general, extendable, and fast approach for phylogenetic inference from multiple continuous characters [31], thus demonstrating promise for leveraging quantitative (i.e., GMM) data in phylogeny estimation moving forward.

Given the conflicting conclusions on the efficacy of GMM data for phylogeny reconstruction, coupled with the recent advancement in whole bone GMM approaches that enable the option to rapidly collect phenotypic data for taxonomic purposes, we review the current state of the field in using morphometric data for inferring phylogeny using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses in ecology and evolutionary biology (PRISMA-EcoEvo v1.0.) reporting guideline. In addition, we assess the efficacy of morphometric data compared to discrete characters and offer some pathways for approaching this task with GMM data (which we here consider a subset of continuous data). Our objectives were to 1) test whether morphological phylogeny congruence with a reference molecular benchmark improves when morphometric data are included, and 2) identify the challenges and benefits that could arise from the use of GMM data in phylogenetic reconstruction.

Methods

A search for relevant published literature was performed on October 11, 2023, in the Web of Science database using the PRISMA-EcoEvo version 1.0 reporting guideline of O'Dea et al. [66].

Data sources and search strategy

The search strategy was developed through an initial scoping review and subsequent rearrangement of 'phylogen*' and 'morpho*' as root terms, and with 'landmark*' as an additional term to identify all relevant articles. Asterisks were used to include search results containing variations of the root terms 'phylogeny', 'morphology', 'morphometric' and of the additional term 'landmark'. The final search string comprised the following combination of keywords and Boolean operators: phylogen* AND morpho* AND (continuous OR landmark* OR shape). The search fields selected were "All Fields". Inclusion criteria were defined as: Publication Years "2003–2023"; languages "English". Exclusion criteria were defined as: Document Types "Review Articles" and "Editorial Material" (see Table S1, Additional file 1 for inclusion and exclusion criteria). A date restriction of 2003 to 2023 was applied, allowing for articles published over the last 20 years, up to October 2023, capturing the advancement and maturity of the field of geometric morphometrics [67], which began in the late 1990s [68]. Deduplication, and both title/abstract and full-text screening of records were performed manually by one author in EndNote version 20 [69]. Articles were independently assessed for individual study quality by two authors to determine final inclusion (Table S2, Additional file 1; Fig. 1).

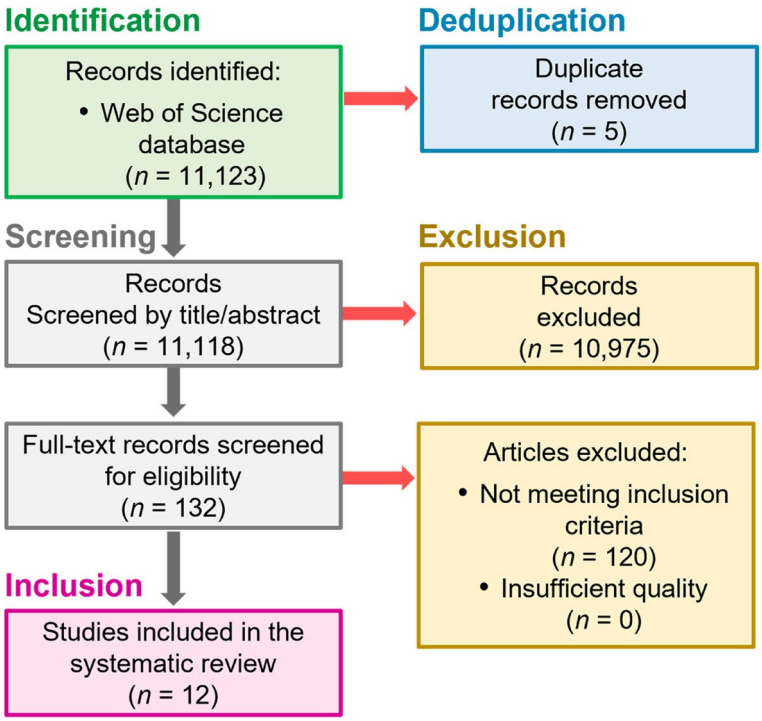

Fig. 1 PRISMA-EcoEvo flowchart adapted from O'Dea et al. [66] for Web of Science literature search and study selection in the systematic review, showing stages of the workflow

Eligibility criteria

Studies were included or excluded in this review according to predefined criteria (Table S1, Additional file 1). These criteria would enable valid comparisons between phylogenies that are based on or included continuous data with those that did not, as well as comparison to a reference phylogeny based on independent (i.e., molecular) data.

Study selection

The search strategy identified an initial 11,123 records (Fig. 1). Five duplicates were found in records identified through the search. Of the 11,118 articles retrieved, 10,975 were excluded because they did not meet the inclusion criteria. Of the 132 potentially eligible studies, a further 120 were excluded, resulting in a total of 12 articles for final inclusion in the systematic review. We note that six of the twelve studies estimate phylogeny from GMM (i.e., 2D or 3D landmark) datasets, and seven studies are based on traditional linear morphometric data (one study includes both GMM and linear morphometric data; see Fig. 2, Table 1; Tables S3-S4, Additional file 1 for full details).

Data collection, extraction

The key data sought from each study were the phylogenies that represented valid comparisons, in computer-readable tree file format, Newick or nexus (see Table 1; S2-S3, Additional file 1) for phylogenetic reconstruction methods and tree comparisons). As the tree files were not available online for any of the included studies, the author(s) of each study were contacted and requested to provide these data. For six of the 12 studies the authors either could not be reached or were not able to provide the tree files. In these cases, phylogenies were written manually in Newick format based on article figures, using the function read.tree from the ape package [version 5.7-1; 79] in the R programming language [version 4.2.2; 80]. Data collection was performed by one author.

Analyses

Molecular constraint trees

Molecular phylogenies were constructed for three studies in which a molecular reference phylogeny was not available, to provide a constraint against which the morphology-based trees of these studies could be compared (see S4, Additional file 1 for molecular analyses and GenBank accession numbers).

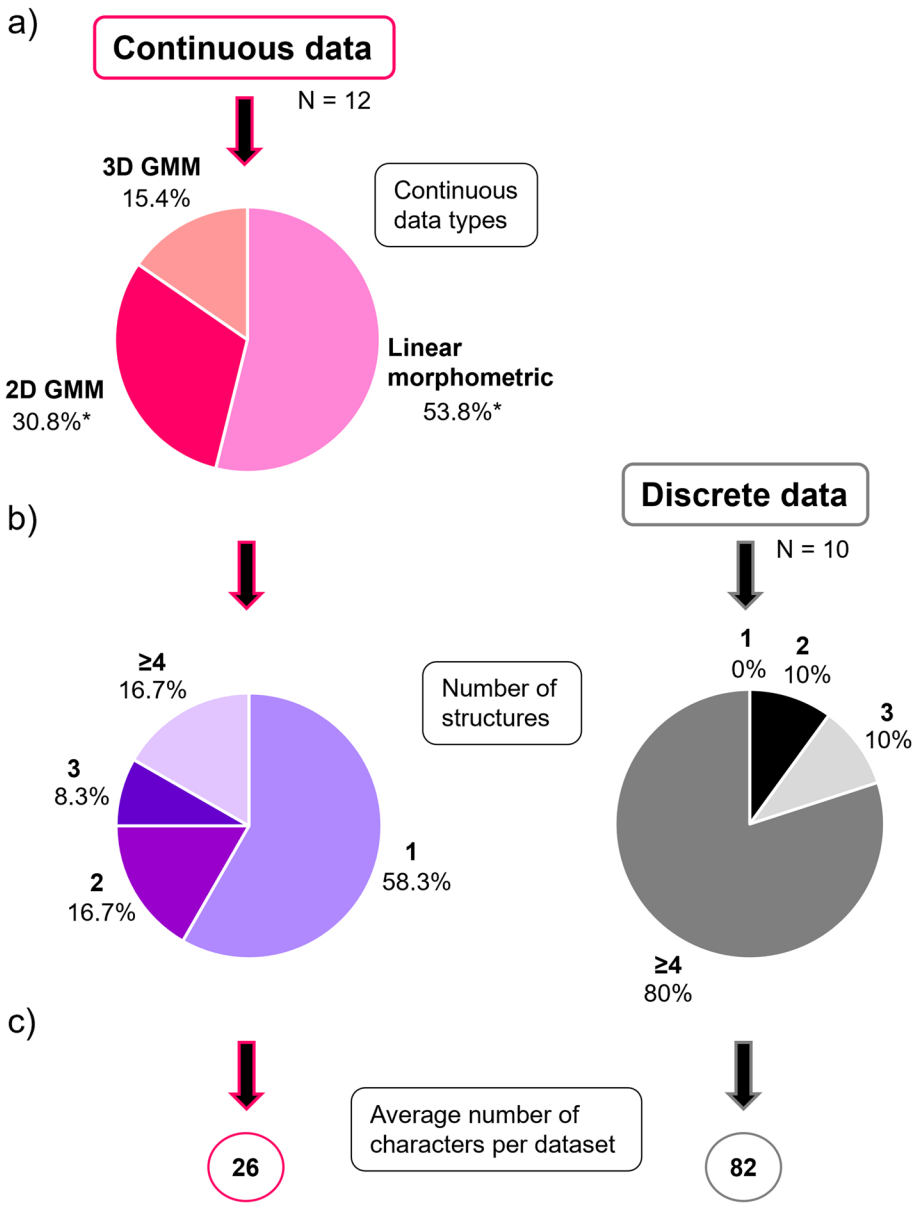

Fig. 2 Overview of morphological data obtained from included studies and analysed in the systematic review. Percentage of studies that include a) each continuous morphometric data type, *N* = 12 studies, and b) the number of morphological structures (1, 2, 3 or ≥ 4) at the 'element' level e.g., cranium, femur in continuous vs discrete (*N* = 10 studies) datasets. c) The average number of morphological characters per dataset for continuous vs discrete data. *One study was counted twice as it includes both 2D GMM and linear morphometric data

Tree comparisons

Phylogenies were read into R using either ape::read.tree or ape::read.nexus [79]. To provide equivalence for comparability between morphological and molecular phylogenies, trees based on taxa (i.e., tree tips) that did not match exactly or that differed in the number of taxa within the same study required pruning, taxon matching and/or taxon re-naming. Pruning and matching was performed using functions from the ape package in R (see S5, Additional file 1 for taxon matching). To ensure consistency, branch lengths were not included in analyses if they were not available for all tree comparisons within a single study.

First, we compared the congruence of the morphology-based trees with the molecular reference phylogeny by computing generalized Robinson-Foulds (RF) distances [81] using functions from the Treedist v2.7.0 [82] package in R (R code available at https://doi.org/10.5281/

Table 1 Overview of morphological phylogenetic studies included in the systematic review. **'**Phylogenetic morphometric method' refers to the treatment of the continuous morphometric data in the analysis with abbreviations adapted from Catalano and Torres [24]: Phylogenetic Morphometrics/ landmark analysis under parsimony (PM/LAUP); landmark coordinates as continuous characters AND (linear) parsimony (LC-P); Procrustes distances AND unweighted pair group method with arithmetic mean (PD-UPGMA); Procrustes distances AND neighbor-joining (PD-NJ); linear morphometrics AND linear parsimony (L-P); Procrustes distances AND minimum evolution (PD-ME); principal component (PC) scores as continuous characters AND parsimony (PC-P). Key: *N* number of terminal taxa in comparison trees; *SR* systematic review, *GMM* geometric morphometric (landmark-based) data, *linear MM* linear morphometric data, *discrete* discrete morphological-only dataset; *continuous* continuous-only dataset; *combined* combined discrete morphological and continuous dataset; *GW* gap-weighted, *SM GW* step-matrix gap-weighted, *7 OG* seven original continuous quantitative characters; *40* larger dataset of 40 measurement and meristic characters; *3 sig* three continuous characters carrying significant phylogenetic signal; *5 sig* five continuous characters carrying significant phylogenetic signal. Continuous and combined datasets (i.e., tree comparisons) analysed 'as such' unless otherwise stated

Study	Taxonomic group	Continuous data	Tree comparisons	Phylogenetic morphometric method	N	
					Original study	SR Analyses
Weisbecker et al. [44]	Animal (vertebrate) Marsupial mammals (Marsupalia)		1) Molecular		*N* = 61	*N* = 45
			2) Discrete		*N* = *45*	
		GMM (3D)	3) Continuous	PM/LAUP	*N* = *45*	
		GMM (3D)	4) Continuous	PD-UPGMA	*N* = 45	
		GMM (3D)	5) Continuous	PD-NJ	*N* = 45	
Gomez and Lois-Milevicich [70]	Animal (vertebrate) cowbirds (Icteridae)		1) Molecular scaffold		*N* = 20	*N* = 20
			2) Discrete		*N* = *20*	
		linear MM	3) Continuous	PM/LAUP; L-P	*N* = *20*	
		linear MM	4) Combined	PM/LAUP; L-P	*N* = *20*	
Solis-Zurita et al. [46]	Animal (vertebrate) Lizards (Squamata: Phrynosomatidae: *Sceloporus*)		1) Molecular		*N* = 20	*N* = 20
		GMM (2D); linear MM	2) Continuous	PM/LAUP; L-P	*N* = 20	
		GMM (2D); linear MM	3) Combined	PM/LAUP; L-P	*N* = *20*	
Celik et al. [59]	Animal (vertebrate) Kangaroos and wallabies (Macropodidae)		1) Molecular		*N* = 35	*N* = 13
			2) Discrete[b]		*N* = *21*	
		GMM (3D)	3) Continuous	PD-UPGMA	*N* = *33*	
		GMM (3D)	4) Continuous	PD-ME	*N* = *33*	
Perrard et al. [45]	Animal (invertebrate) Vespinae wasps (Hymenoptera: Vespidae: Vespinae)		1) Molecular		*N* = 55	*N* = 52
			2) Discrete[a]		*N* = *55*	
		GMM (2D)	3) Combined[a]	PM/LAUP	*N* = *52*	
Cichocka and Bielecki [71]	Animal (invertebrate) Hirudinid leeches (Clitellata: Hirudinida)		1) Molecular constraint		*N* = 27	*N* = 31
		linear MM	2) Continuous	L-P	*N* = *31*	
		linear MM	3) Continuous GW	L-P	*N* = *31*	
		linear MM	4) Combined	L-P	*N* = *31*	
		linear MM	5) Combined GW	L-P	*N* = *31*	
Gold et al. [47]	Animal (vertebrate) Crocodilians (Crocodylia)		1) Molecular		*N* = 18	*N* = 16
			2) Discrete[a]		*N* = *119*	
		GMM (2D)	3) Combined[a]	PC-P	*N* = *119*	
		GMM (2D)	4) Combined[a]	LC-P	*N* = *119*	
Vargas et al. [72]	Animal (invertebrate) Octocorals (Coelenterata: Octocorallia: Gorgoniidae: eastern pacific genus *Pacifigorgia*)		1) Molecular constraint		*N* = 16	*N* = 15
			2) Discrete		*N* = 25	
		linear MM	3) Continuous GW	L-P	*N* = 25	
		linear MM	4) Combined	L-P	*N* = 25	

Table 1 (continued)

Study	Taxonomic group	Continuous data	Tree comparisons	Phylogenetic morphometric method	N	
					Original study	SR Analyses
de Bivort and Giribet [73]	Animal (invertebrate) Mite harvestman (Arachnida: Opiliones: Cyphophthalmi: South African Pettalidae)		1) Molecular constraint		*N* = 38	*N* = 38
			2) Discrete		*N* = *50*	
		linear MM	3) Continuous	L-P	*N* = *50*	
		linear MM	4) Combined	L-P	*N* = *50*	
Hendrixson and Bond [74]	Animal (invertebrate) Mygalomorph spider (Araneae: Mygalomorphae: Antrodiaetidae: *Antrodiaetus*)		1) Molecular		*N* = *17*	*N* = 17
			2) Discrete		*N* = *17*	
		linear MM	3) Combined 7 OG, discretized	L-P	*N* = *17*	
		linear MM	4) Combined 7 OG, SM GW	L-P	*N* = *17*	
		linear MM	5) Combined 7 OG	L-P	*N* = *17*	
		linear MM	6) Combined 40, SM GW	L-P	*N* = *17*	
		linear MM	7) Combined 40	L-P	*N* = *17*	
		linear MM	8) Combined 3 sig, SM GW	L-P	*N* = *17*	
		linear MM	9) Combined 5 sig, SM GW	L-P	*N* = *17*	
		linear MM	10) Combined 3 sig	L-P	*N* = *17*	
		linear MM	11) Combined 5 sig	L-P	*N* = *17*	
Hardy et al. [75]	Plant Cape reeds (African Restionaceae)		1) Molecular		*N* = 297	*N* = 297
			2) Discrete[a]		*N* = *297*	
		linear MM	3) Combined[a]	L-P	*N* = *297*	
Edgar and Theriot [76]	Protist Diatoms (Bacillariophyta: *Aulacoseia*)		1) Molecular		*N* = 24	*N* = 22
			2) Discrete		*N* = *70*	
		GMM (2D)	3) Continuous SM GW[c]	L-P	*N* = *68*	
		GMM (2D)	4) Combined SM GW[c]	L-P	*N* = *70*	

[a] Dataset includes DNA

[b] constructed using modern taxa only from discrete morphological matrix of Travouillon et al. [77]

[c] modification (addition) to Wiens' [78] step-matrix gap weighting method

zenodo.13357792). In contrast to standard RF metrics which are strictly based on the number of identical splits shared between two trees, generalized RF metrics consider potential similarities between non-identical splits, thus quantifying overall similarity between two trees [81]. Second, the number of resolved nodes with branch support 50 percent or higher (≥ 50%) were calculated as a percentage of the total number of nodes in the tree based on article figures or tables, or tree files. Third, support values for clades shared (i.e., equivalent splits) between alternative morphology-based trees were obtained from article figures or tree files and then averaged for each tree (see Table S8, Additional file 1 for comparable clades). Note that support values were only available for four studies; and thus, the latter two analyses were based on a subset of the twelve included studies. The overall mean value for each data category across all studies was calculated for each of the three analyses.

Results and discussion

The inclusion of continuous morphological data did not improve phylogeny reconstruction in terms of the number of resolved nodes and their statistical support or congruence with molecular benchmark phylogenies (i.e., comparing tree distances), regardless of whether the continuous data were included alone or integrated into discrete morphological datasets. We note that branch support (i.e., bootstrap) is a measure of resolution (precision), while RF distance is more a measure of accuracy, although low resolution can also manifest as reducing apparent accuracy. Continuous or combined discrete and continuous data performed similarly well to discrete data in terms of tree distances to molecular benchmarks

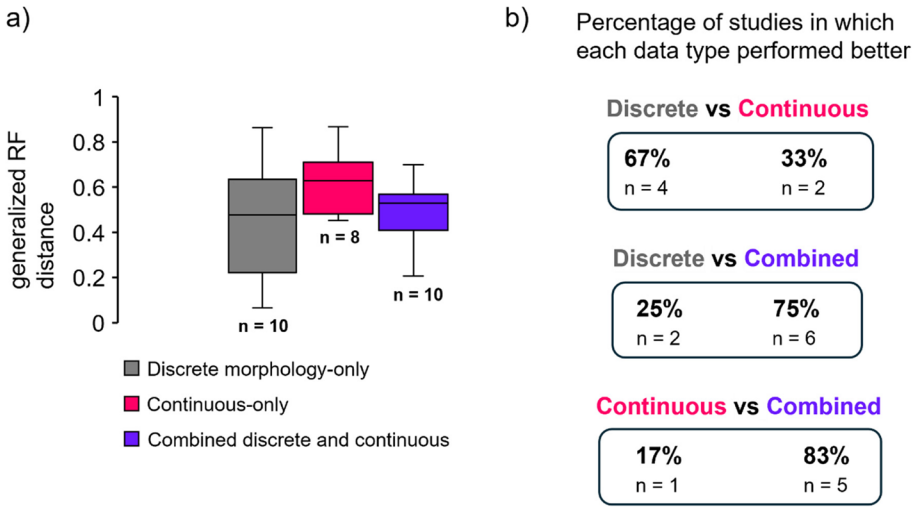

Fig. 3 a) Normalized generalized Robinson-Foulds (RF) distance from the molecular tree to the comparison trees averaged across all studies, *N* = 12. Lower values correspond to a closer match (i.e., most similar topology) with the reference tree, with zero being an equal match. b) Overall performance of different morphological data types across all studies for each pairwise data comparison where the percentage indicates the number of studies in which a given data type performed better. n indicates the number of studies included for each data type in each pairwise comparison

and support values for shared clades (Figs. 3a, 5a; see Tables 2 and 4 for full details), but are outperformed by discrete data in terms of the number of resolved nodes (≥ 50%) (Fig. 4a; see Table 3 for full details). Combined (continuous and discrete) data performed better overall than continuous data alone in two out of the three comparable analyses.

On average, the discrete- and combined data-based trees were (equally) most congruent with the molecular reference phylogeny (i.e., have the lowest generalized RF values), with the continuous data-based trees being least congruent (Fig. 3a; Table 2). When comparing discrete- vs continuous-based trees, discrete data-based trees performed better in four studies (67%), while continuous data-based trees performed better in two studies (33%) (Fig. 3b). When comparing discrete- vs combined-data based trees, discrete data-based trees performed better in two studies (25%), while combined data-based trees performed better in six studies (75%). When comparing continuous- vs combined-data based trees, continuous data-based trees performed better in one study (17%), while combined data-based trees performed better in five studies (83%). There appears to be a temporal trend in terms of the advancement of morphometric methods (i.e., from linear morphometrics to 2D and 3D GMM) implemented between 2003 and 2023 (Table 2). Approximately half of the twelve studies represent phylogenies based on datasets comprising GMM data, while the other half comprise linear morphometric data (Table 1; Table 2). We also note that two of the twelve studies do not include a discrete data comparison (see Tables 1 and 2).

On average, discrete-based trees outperformed the continuous and combined data-based trees in terms of the number of resolved nodes (≥ 50%), with discrete-based trees having the highest percentage and continuous-based trees having the lowest (Fig. 4a; Table 3). When comparing discrete- vs combined-data based trees, discrete data-based trees performed better in one study (33%), while combined data-based trees performed better in two studies (67%) (Fig. 4b). When comparing continuous- vs combined-data based trees, continuous data-based trees performed better in one study (33%), while combined data-based trees performed better in two studies (67%).

On average, discrete- and continuous-based trees performed similarly in terms of the average support value for shared clades, with continuous-based trees having the highest value and combined data-based trees having the lowest (Fig. 5a; Table 4). When comparing discrete- vs combined-data based trees, the discrete data performed better in one study (50%), while the combined data performed better in the other (50%) (Fig. 5b). When comparing continuous- vs combined-data based trees, the continuous data performed better in one study (50%), while the combined data performed better in the other (50%).

Similar results (i.e., average values) and relative performance of each data category were found when considering only those studies which include GMM data (see

Table 2 Normalized generalized Robinson-Foulds (RF) distance from the molecular tree to the comparison trees, $N = 12$ studies. Key: *discrete* discrete morphological-only dataset; *continuous* continuous-only dataset; *combined* combined discrete morphological and continuous dataset; *GW* gap-weighted; *SM GW* step-matrix gap-weighted; *7 OG* seven original continuous quantitative characters; *40* larger dataset of 40 measurement and meristic characters; *3 sig* three continuous characters carrying significant phylogenetic signal; *5 sig* five continuous characters carrying significant phylogenetic signal. Abbreviations adapted from Catalano and Torres [24]: Phylogenetic Morphometrics/ landmark analysis under parsimony (PM/LAUP); landmark coordinates as continuous characters AND (linear) parsimony (LC-P); Procrustes distances AND unweighted pair group method with arithmetic mean (PD-UPGMA); Procrustes distances AND neighbor-joining (PD-NJ); linear morphometrics AND linear parsimony (L-P); Procrustes distances AND minimum evolution (PD-ME); principal component (PC) scores as continuous characters AND parsimony (PC-P). The tree with the smallest distance from (i.e., most congruent with) the molecular tree is shown in bold. Note that some studies include more than one comparison based on a continuous morphometric or combined discrete morphological and continuous morphometric dataset (i.e., representing different treatments of continuous data and analytical approaches to the phylogenetic reconstruction)

Study	generalized RF distance				
	Discrete	Continuous		Combined	
Weisbecker et al. [44]	**0.34**	PM/LAUP PD-NJ PD-UPGMA	0.609 0.526 0.527	n/a	n/a
Gomez and Lois-Milevicich [70]	0.595	PM/LAUP, L-P	0.687	PM/LAUP, L-P	**0.570**
Solis-Zurita et al. [46]	n/a	PM/LAUP; L-P	0.773	PM/LAUP; L-P	**0.575**
Celik et al. [59]	**0.418**	PD-ME PD-UPGMA	0.719 0.664	n/a	n/a
Perrard et al. [45]	**0.066**[a]	n/a	n/a	PM/LAUP	0.11[a]
Cichocka and Bielecki [71]	n/a	GW, L-P L-P	0.457 0.453	GW, L-P L-P	0.457 **0.394**
Gold et al. [47]	0.2403[a]	n/a	n/a	PC-P LC-P	**0**[a] **0**[a]
Vargas et al. [72]	0.864	GW, L-P	0.868	L-P	**0.816**
de Bivort and Giribet [73]	0.538	L-P	**0.467**	L-P	0.477
Hendrixson and Bond [74]	0.547	n/a	n/a	7 OG, discretized, L-P 7 OG, SM GW, L-P 7 OG, L-P 40, SM GW, L-P 40, L-P 3 sig, SM GW, L-P 3 sig, L-P 5 sig, SM GW, L-P 5 sig, L-P	0.531 **0.52** **0.52** 0.566 0.699 0.527 0.537 0.55 0.537
Hardy et al. [75]	**0.164**[a]	n/a	n/a	L-P	0.207[a]
Edgar and Theriot [76]	0.755	SM GW[b], L-P	0.650	SM GW[b], L-P	**0.580**
Overall average	0.453		0.617		0.459

[a] Dataset includes DNA

[b] modification (addition) to Wiens' [78] step-matrix gap-weighting method

Additional file 1, Tables S9-11). When comparing the performance of linear morphometric vs GMM (both 2D and 3D, where possible) data (regardless of the underlying method), GMM-based studies with GMM data in combined datasets performed best, on average, in terms of congruence with the molecular reference phylogeny (i.e., had the lowest generalized RF values) (0.316) (Additional file 1, Table S12). Overall, combined datasets (0.530, 0.316) performed better than continuous-only datasets (0.617, 0.638) (Additional file 1, Table S12). Continuous-only datasets performed similarly for both GMM- (0.638) and linear morphometric-based (0.6174) studies, while combined datasets performed better for GMM-based studies (0.316) than for linear morphometric-based (0.530 studies.

When comparing the performance of linear morphometric vs GMM data, linear morphometric-based studies with linear morphometric data in combined datasets performed best, on average, in terms of the number of resolved nodes (≥ 50%) (41%) (Additional file 1, Table S13). Overall, combined datasets (41%, 36%) performed better than continuous-only datasets (26%, 11%) (Additional file 1, Table S13). Continuous-only datasets

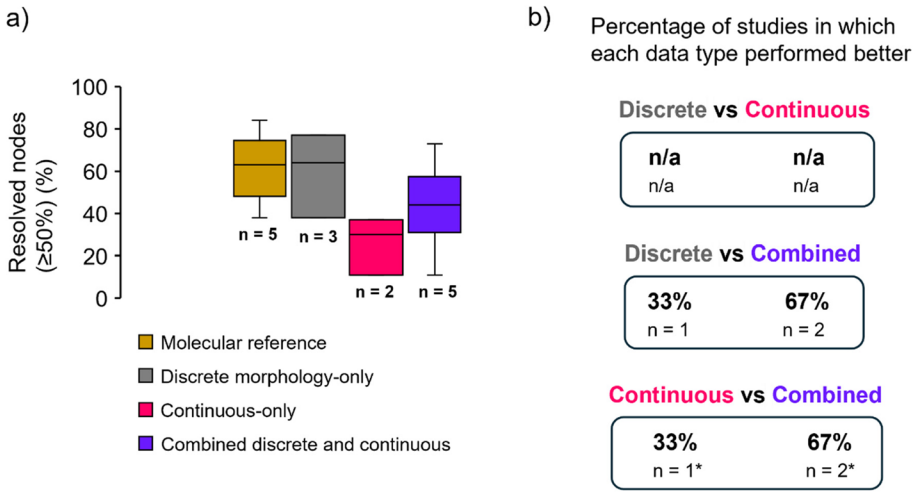

Fig. 4 a) Number of resolved nodes with branch support 50 percent or higher (≥ 50%) as a percentage of the total number of nodes in the tree averaged across all studies, $n = 5$. b) Overall performance of different morphological data types across all studies for each pairwise data comparison where the percentage indicates the number of studies in which a given data type performed better. n indicates the number of studies included for each data type in each pairwise comparison. Note that discrete vs continuous comparisons were not available (hence the n/a values for this pairwise comparison). *One study was counted twice as it performed equally well for continuous and combined data

performed better for linear morphometric-based studies (26%) than for GMM-based (11%) studies.

When comparing the performance of linear morphometric vs GMM data, GMM-based studies with GMM data in continuous-only datasets performed best, on average, in terms of the average support value for shared clades (99%) (Additional file 1, Table S14). Overall, continuous datasets performed better than combined datasets. Continuous-only datasets performed better for GMM-based studies (99%) than for linear morphometric-based studies (87.3%), while combined datasets performed similarly for both GMM- (71.1%) and linear morphometric-based (70.4%) studies.

Limitations

Limitations of our study include the small sample size (12), in part due to the eligibility criteria excluding studies that did not contain valid comparisons. For example, when the study included a continuous-only tree, but no discrete-only or combined data tree to directly compare and thus determine the effect of including continuous data. In other cases, valid comparisons were available, but a molecular reference tree was not, and it was not possible to make one. For example, if the study focused on fossil taxa then it could not be included. We are limited to existing empirical data, which at present does not allow for large sample sizes. We found great difficulty in obtaining relevant (i.e., discrete vs continuous) comparisons from the literature despite a rigorous and exhaustive search, which hinders our understanding of the performance of continuous data. Thus, we implore researchers to address this issue with more comparable analyses (i.e., targeting discrete vs continuous comparisons from the outset). We also note that a more pure "like for like" comparison i.e., between discrete data that describe shape variation and the quantitative treatment of the same data, is also interesting from a theoretical point of view and construction of such datasets should be encouraged.

The lack of support values for phylogenies within some studies meant that analyses comparing resolved nodes and average support values could only be based on a subset of the total number of studies included in the review (five and four out of 12, respectively). Furthermore, phylogenetic accuracy between studies may not be directly comparable due to different dynamics/resolution in different taxonomic groups. Different support measures might also not be directly comparable, for example, for one study [74] Bayesian posterior probability for the molecular reference was compared to bootstrap values for all morphological comparison trees. We also note that molecular trees, used here as a benchmark comparison, are not errorless, but are commonly considered the 'gold standard' in phylogenetics.

We acknowledge that there are biases associated with the definition and scoring of discrete morphological data, however this does not preclude a comparison. Discrete characters represent an alternative and adopted route used for many phylogenetic studies, especially in the field of palaeontology. Therefore, our comparison does yield practical relevance for researchers using these approaches. The GMM datasets we analysed typically cover fewer elements or structures than

Table 3 The number of resolved nodes with branch support 50 percent or higher (≥ 50%) as a percentage of the total number of nodes in the tree, $n = 5$ studies. Molecular = molecular reference tree. The morphology-based tree with the highest number of resolved nodes (≥ 50%) is shown in bold

Study	Molecular		Resolved nodes (≥ 50%)							
			Discrete		Continuous			Combined		
	Total nodes	Nodes (≥ 50%)	Total nodes	Nodes (≥ 50%)		Total nodes	Nodes (≥ 50%)		Total nodes	Nodes (≥ 50%)
Solis-Zurita et al. [46]	19	11 (58%)	n/a	n/a	PM/LAUP; L-P	19	2 (**11%**)	PM/LAUP; L-P	19	2 (**11%**)
Perrard et al. [45]	51	33 (65%)	52	40 **(77%)**[a]	n/a	n/a	n/a	PM/LAUP	47	29 (62%)[a]
Cichocka and Bielecki [71]	25	21 (84%)	n/a	n/a	GW, L-P	30	11 (37%)	GW, L-P	30	11 (37%)
					L-P	30	9 (30%)	L-P	30	16 **(53%)**
Hendrixson and Bond [74]	16	6 (38%)	16	6 (38%)	n/a	n/a	n/a	7 OG, discretized, L-P	16	7 (44%)
								7 OG, SM GW, L-P	16	8 (50%)
								7 OG, L-P	16	2 (13%)
								40, SM GW, L-P	16	6 (38%)
								40, L-P	16	9 (56%)
								3 sig, SM GW, L-P	16	7 (44%)
								3 sig, L-P	16	10 **(63%)**
								5 sig, SM GW, L-P	16	6 (38%)
								5 sig, L-P	16	2 (13%)
Hardy et al. [75]	271	171 (63%)	271	174 (64%)[a]	n/a	n/a	n/a	L-P	271	199 **(73%)**[a]
Overall average		61%		60%			26%			42%

[a] Dataset includes DNA

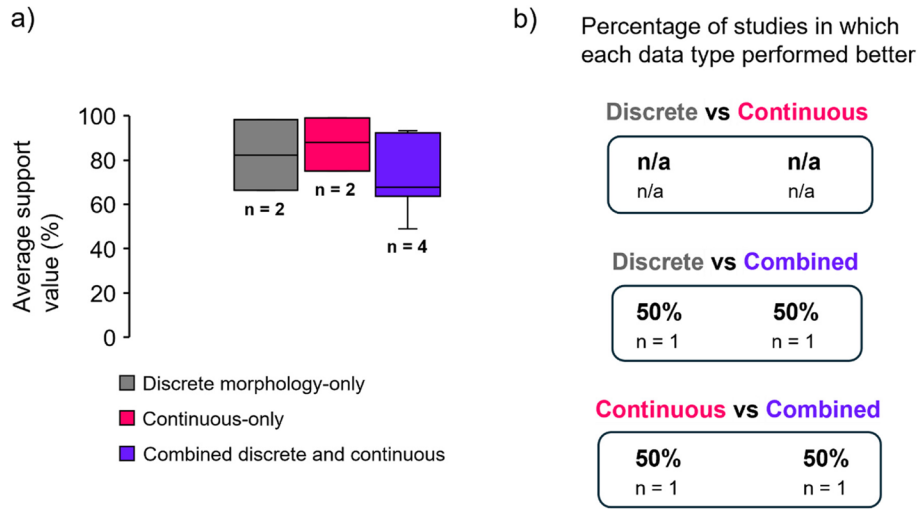

Fig. 5 a) Average support value for shared clades (i.e., equivalent splits) across alternative morphology-based comparison trees averaged across all studies, $n = 4$ studies. b) Overall performance of different morphological data types across all studies for each pairwise data comparison where the percentage indicates the number of studies in which a given data type performed better. n indicates the number of studies included for each data type in each pairwise comparison. Note that discrete vs continuous comparisons were not available (hence the n/a values for this pairwise comparison)

Table 4 Average support value for shared clades (i.e., equivalent splits) across morphology-based comparison trees, $n = 4$ studies. The tree with the highest average support value is shown in bold

	Average support value				
Study	**Discrete**	**Continuous**		**Combined**	
Solis-Zurita et al. [46]	n/a	PM/LAUP; L-P	**99**	PM/LAUP; L-P	49
Perrard et al. [45]	**98.2**[a]	n/a	n/a	PM/LAUP	93.2[a]
Cichocka and Bielecki [71]	n/a	GW, L-P L-P	88 75	GW, L-P L-P	92 **92.5**
Hendrixson and Bond [74]	66.3	n/a	n/a	7 OG, discretized, L-P 7 OG, SM GW, L-P 3 sig, SM GW, L-P 3 sig, L-P 5 sig, SM GW, L-P	67.7 66.7 64.7 **68.3** 62.7
Overall average	82.3		87.3		73

[a] Dataset includes DNA

do the discrete datasets. Thus, our comparisons more closely reflect the current state of the field for researchers deciding to use continuous vs discrete data, rather than the full potential of GMM data.

We note that only six of the twelve studies included in the systematic review estimate phylogeny from GMM (i.e., landmark) datasets. Seven studies are based on traditional linear morphometric data collected from morphological structures which are often utilised as taxonomically informative characters, which may be due to the time and labour required to collect GMM data. Furthermore, one study [76] did not include an 'as such' comparison of continuous morphometric data. Here, 2D GMM data were not analysed as quantitative variables but were discretized using a modification (addition) to Wiens' [78] step-matrix gap-weighting method in both a continuous-only dataset and in a dataset combined with discrete morphological data, and in both cases were treated as continuous data in the analyses.

Future directions

Overall, our results suggest that continuous morphometric data do not perform well when applied to the inference of phylogenetic topology. Since none of the included studies used probabilistic methods, it would be valuable

to explore how these methods might perform compared to parsimony or distance-based methods. Our findings suggest a temporal trend in terms of the advancement of morphometric methods (i.e., from linear morphometrics to 2D and 3D GMM) implemented between 2003 and 2023, however, no trend is shown in the same timeframe for the improvement in continuous data compared with discrete data. In the case of GMM this may be due to the high dimensionality of the data and variation within these types of datasets. Thus, 'noise reduction' techniques [83, 84] will be critical to the use of GMM data for phylogeny to enable the isolation of phylogenetic signal.

In addition, the full potential of GMM data e.g., whole body or whole osteology landmarking will be important to explore in future. Despite the challenges associated with estimating phylogenetic topology from GMM data, the advancement of morphological phylogenetics depends on improved methods for extracting and modelling quantitative morphological data. Hence, the potential benefits of leveraging quantitative data for this purpose warrant further investigations into how to best model the complexity and correlations inherent in GMM character data.

One promising approach is to objectively discretize the shape variation in landmark data, subdividing the landmarks into "characters" and discretizing each of these into states using clustering algorithms [22]. Additionally, some continuous GMM characters do not appear to contain phylogenetic signal at all which may be due to signal erosion over time, with rapidly evolving traits or phylogenetic signal being overwhelmed by functional covariation and phenotypic plasticity [74, 85]. Phylogenetic signal erosion/retention for molecular phylogenies have been quantified using stemminess analyses, saturation analyses and various other metrics [86–91]. Phylogenetic signal retention in morphological landmark data and across different taxonomic groups will be another important area for further research.

Conclusions

Our study provides a comprehensive and rigorous assessment of the performance of existing empirical continuous data using the PRISMA-EcoEvo version 1.0 reporting guideline, and identifies important challenges to overcome, as well as benefits that could arise from the use of GMM data in phylogenetic reconstruction. We find that the inclusion of continuous morphological data does not improve phylogeny reconstruction, and our analyses show that overall, continuous morphometric data do not perform well when applied to the inference of phylogenetic topology. However, despite the challenges associated with estimating phylogenetic topology from GMM data, the advancement of morphological phylogenetics depends on improved methods for extracting and modelling quantitative morphological data. Hence, the potential benefits of leveraging quantitative data for this purpose warrant further investigations into how to best model the complexity and correlations inherent in GMM character data.

Since morphological data is often used in conjunction with molecular data, it would also be valuable to explore how discrete and continuous morphological data respectively interact with molecular data in producing emergent phylogenetic signals. Our study demonstrates the problem surrounding the efficacy of continuous data as remaining relatively intractable despite an exhaustive search, due in part to the difficulty in obtaining relevant comparisons from the literature. Our study was performed under a rigorous framework for systematic reviews, which showed that the lack of available comparisons between discrete and continuous data hinders our understanding of the performance of continuous data.. Thus, we implore researchers to address this issue with studies that collect discrete and continuous data sets with directly comparable properties (i.e., describing shape, or size).

Abbreviations

GMM	Geometric morphometric
ML	Maximum likelihood
PCMs	Phylogenetic comparative methods
PRISMA-EcoEvo v1.0.	Preferred Reporting Items for Systematic Reviews and Meta-Analyses in ecology and evolutionary biology version 1.0
LAUP	Landmark analysis under parsimony
PM	Phylogenetic Morphometrics
PC	Principal component
NJ	Neighbour-joining
UPGMA	Unweighted pair group method with arithmetic mean
fWLS	Flexibly weighted least squares
RF	Robinson-Foulds

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12862-024-02313-3.

Additional file 1: Supporting information for this publication: inclusion and exclusion, quality assessment criteria, phylogenetic reconstruction methods, tree comparisons, molecular analyses, taxon matching, comparable clades, results, GMM data-only and results, linear morphometric vs GMM data.

Acknowledgements

We thank the authors of studies who responded and provided data that were used in this study.

Authors' contributions

L.A.B.W. conceived the study. E.J.H. analysed the data and wrote the original draft. M.J.P. and M.A.C. were involved in planning and performing the analyses. L.A.B.W., M.J.P. and M.A.C. supervised the study. E.J.H. prepared all figures and tables. All authors reviewed and approved the final manuscript.

Funding
This work was supported by an Australian Government Research Training Program Scholarship (to E.J.H.), and the Australian Research Council Discovery Program (FT200100822) (to L.A.B.W.).

Data availability
The datasets analysed during the current study and the R code are available in the Zenodo repository, https://doi.org/10.5281/zenodo.13357792.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]School of Archaeology and Anthropology, The Australian National University, Canberra, Australia. [2]School of Biology and Environmental Science, Queensland University of Technology, Brisbane, QLD, Australia. [3]School of Biological, Earth and Environmental Sciences, University of New South Wales, Kensington, NSW 2052, Australia. [4]ARC Training Centre for Multiscale 3D Imaging, Modelling and Manufacturing, Research School of Physics, The Australian National University, Acton, ACT 2601, Australia.

Received: 28 May 2024 Accepted: 2 October 2024
Published online: 18 October 2024

References

***Studies included in the systematic review**
1. Phillips MJ, Celik MA, Beck RM. The evolutionary relationships of Diprotodontia and improving the accuracy of phylogenetic inference from morphological data. Alcheringa. 2023;16:1–13.
2. Webster M, Sheets HD. A practical introduction to landmark-based geometric morphometrics. Paleontol Soc Pap. 2010;16:163–88.
3. Rohlf FJ. Geometric morphometrics and phylogeny. Syst AssocSpec. 2002;64:175–93.
4. Bookstein FL. Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. Med Image Anal. 1997;1:225–43.
5. Bookstein FL. Morphometric tools for landmark data. New York: Cambridge University Press; 1997.
6. Zelditch M, Swiderski D, Sheets HD. Introduction. In: Zelditch ML, Swiderski DL, Sheets HD, editors. Geometric morphometrics for biologists. 2nd ed. San Diego, CA: Academic Press; 2012. p. 1–20.
7. Cardini A, Marco VA. Procrustes shape cannot be analyzed, interpreted or visualized one landmark at a time. Evol Biol. 2022;49:239–54.
8. Álvarez-Carretero S, Goswami A, Yang Z, Dos Reis M. Bayesian Estimation of Species Divergence Times Using Correlated Quantitative Characters. Syst Biol. 2019;68:967–86.
9. Parins-Fukuchi C. Use of Continuous Traits Can Improve Morphological Phylogenetics. Syst Biol. 2018;67:328–39.
10. Collyer ML, Adams DC. Phylogenetically aligned component analysis. Methods Ecol Evol. 2021;12:359–72.
11. Takács P, Vitál Z, Ferincz Á, Staszny Á. Repeatability, reproducibility, separative power and subjectivity of different fish morphometric analysis methods. PLoS ONE. 2016;11:e0157890.
12. Devine J, Aponte JD, Katz DC, Liu W, Vercio LDL, Forkert ND, et al. A Registration and Deep Learning Approach to Automated Landmark Detection for Geometric Morphometrics. Evol Biol. 2020;47:246–59.
13. Ridel AF, Demeter F, Galland M, L'abbé EN, Vandermeulen D, Oettlé AC. Automatic landmarking as a convenient prerequisite for geometric morphometrics. Validation on cone beam computed tomography (CBCT)- based shape analysis of the nasal complex. Forensic Sci Int. 2020;306:110095.
14. Boyer DM, Puente J, Gladman JT, Glynn C, Mukherjee S, Yapuncich GS, et al. A New Fully Automated Approach for Aligning and Comparing Shapes. Anat Rec. 2015;298:249–76.
15. Pomidor BJ, Makedonska J, Slice DE. A landmark-free method for three-dimensional shape analysis. PLoS ONE. 2016;11:e0150368.
16. Koehl P, Hass J. Landmark-free geometric methods in biological shape analysis. J R Soc Interface. 2015;12:20150795.
17. White JD, Ortega-Castrillón A, Matthews H, Zaidi AA, Ekrami O, Snyders J, et al. MeshMonk: Open-source large-scale intensive 3D phenotyping. Sci Rep. 2019;9:6085.
18. Gao T, Yapuncich GS, Daubechies I, Mukherjee S, Boyer DM. Development and Assessment of Fully Automated and Globally Transitive Geometric Morphometric Methods, With Application to a Biological Comparative Dataset With High Interspecific Variation. Anat Rec. 2018;301:636–58.
19. Porto A, Rolfe S, Maga AM. ALPACA: A fast and accurate computer vision approach for automated landmarking of three-dimensional biological structures. Methods Ecol Evol. 2021;12:2129–44.
20. Rolfe SM, Maga AM. DeCA: a dense correspondence analysis toolkit for shape analysis. In: Wachinger, C, Paniagua, B, Elhabian, S, Li, J, Egger, J, editors. Shape in Medical Imaging. ShapeMI 2023. Lecture Notes in Computer Science, Vol 14350. Cham, Switzerland: Springer; 2023. https://doi.org/10.1007/978-3-031-46914-5_21.
21. Zhang C, Porto A, Rolfe S, Kocatulum A, Maga AM. Automated landmarking via multiple templates. PLoS ONE. 2022;17:e0278035.
22. Celik MA. Tracing the evolution of Australasian mammals: Integrating morphological, palaeontological and molecular data: PhD Thesis, Queensland University of Technology; 2020. https://eprints.qut.edu.au/135716/.
23. Ascarrunz E, Claude J, Joyce WG. Estimating the phylogeny of geoemydid turtles (Cryptodira) from landmark data: an assessment of different methods. PeerJ. 2019;7:e7476.
24. Catalano SA, Torres A. Phylogenetic inference based on landmark data in 41 empirical data sets. Zool Scr. 2017;46:1–11.
25. Goloboff PA, Mattoni CI, Quinteros AS. Continuous characters analyzed as such. Cladistics. 2006;22:589–601.
26. Catalano SA, Goloboff PA. Simultaneously Mapping and Superimposing Landmark Configurations with Parsimony as Optimality Criterion. Syst Biol. 2012;61:392–400.
27. Catalano SA, Goloboff PA, Giannini NP. Phylogenetic morphometrics (I): the use of landmark data in a phylogenetic framework. Cladistics. 2010;26:539–49.
28. Goloboff PA, Catalano SA. Phylogenetic morphometrics (II): algorithms for landmark optimization. Cladistics. 2011;27:42–51.
29. Goloboff PA, Farris JS, Nixon KC. TNT, a free program for phylogenetic analysis. Cladistics. 2008;24:774–86.
30. Felsenstein J. PHYLIP Phylogeny inference package. Department of Genetics, University of Washington, Seattle; 1993.
31. Zhang R, Drummond AJ, Mendes FK. Fast Bayesian inference of phylogenies from multiple continuous characters. Syst Biol. 2023;73:102–24.
32. Varon-Gonzalez C, Whelan S, Klingenberg CP. Estimating Phylogenies from Shape and Similar Multidimensional Data: Why It Is Not Reliable. Syst Biol. 2020;69:863–83.
33. Adams DC, Cardini A, Monteiro LR, O'Higgins P, Rohlf FJ. Morphometrics and phylogenetics: Principal components of shape from cranial modules are neither appropriate nor effective cladistic characters. J Hum Evol. 2011;60:240–3.
34. Monteiro LR. Why morphometrics is special: the problem with using partial warps as characters for phylogenetic inference. Syst Biol. 2000;49:796–800.
35. Cardini A, Elton S. Does the skull carry a phylogenetic signal? Evolution and modularity in the guenons. Biol J Linn Soc. 2008;93:813–34.
36. Parins-Fukuchi C. Bayesian placement of fossils on phylogenies using quantitative morphometric data. Evolution. 2018;72:1801–14.
37. Smith UE, Hendricks JR. Geometric Morphometric Character Suites as Phylogenetic Data: Extracting Phylogenetic Signal from Gastropod Shells. Syst Biol. 2013;62:366–85.

38. Hetherington AJ, Sherratt E, Ruta M, Wilkinson M, Deline B, Donoghue PC. Do cladistic and morphometric data capture common patterns of morphological disparity? Palaeontology. 2015;58:393–9.
39. Caumul R, Polly PD. Phylogenetic and environmental components of morphological variation: Skull, mandible, and molar shape in marmots (*Marmota*, Rodentia). Evolution. 2005;59:2460–72.
40. Viacava P, Blomberg SP, Sansalone G, Phillips MJ, Guillerme T, Cameron SF, et al. Skull shape of a widely distributed, endangered marsupial reveals little evidence of local adaptation between fragmented populations. Ecol Evol. 2020;10:9707–20.
41. Travouillon KJ, Gurovich Y, Beck RMD, Muirhead J. An exceptionally well-preserved short-snouted bandicoot (Marsupialia; Peramelemorphia) from Riversleigh's Oligo-Miocene deposits, northwestern Queensland. Australia J Vertebr Paleontol. 2010;30:1528–46.
42. Travouillon K. Notes on a new method to identify Golden Bandicoot and Northern Brown Bandicoot in the Kimberley region. Rec W Aust Mus. 2022;57:37.
43. Viacava P, Baker AM, Blomberg SP, Phillips MJ, Weisbecker V. Using 3D geometric morphometrics to aid taxonomic and ecological understanding of a recent speciation event within a small Australian marsupial (*Antechinus*: Dasyuridae). Zool J Linn Soc. 2022;196:963–78.
44. * Weisbecker V, Beck RMD, Guillerme T, Harrington AR, Lange-Hodgson L, Lee MSY, et al. Multiple modes of inference reveal less phylogenetic signal in marsupial basicranial shape compared with the rest of the cranium. Philos Trans R Soc Biol Sci. 2023;378:20220085.
45. * Perrard A, Lopez-Osorio F, Carpenter JM. Phylogeny, landmark analysis and the use of wing venation to study the evolution of social wasps (Hymenoptera: Vespidae: Vespinae). Cladistics. 2016;32:406–25.
46. * Solis-Zurita C, De Luna E, Gonzalez D. Phylogenetic relationships in the *Sceloporus variabilis* (Squamata: Phrynosomatidae) complex based on three molecular markers, continuous characters and geometric morphometric data. Zool Scr. 2019;48:419–39.
47. * Gold MEL, Brochu CA, Norell MA. An Expanded Combined Evidence Approach to the *Gavialis* Problem Using Geometric Morphometric Data from Crocodylian Braincases and Eustachian Systems. PLoS ONE. 2014;9:1932–6203.
48. Cavalli-Sforza L, Edwards A. Phylogenetic analysis; models and estimation procedures. Am J Hum Genet. 1967;19:233–57.
49. Thompson E. The method of minimum evolution. Ann Hum Genet. 1973;36:333–40.
50. Klingenberg CP, Gidaszewski NA. Testing and quantifying phylogenetic signals and homoplasy in morphometric data. Syst Biol. 2010;59:245–61.
51. Gonzalez-Jose R, Escapa I, Neves WA, Cuneo R, Pucciarelli HM. Cladistic analysis of continuous modularized traits provides phylogenetic signals in *Homo* evolution. Nature. 2008;453:775–U4.
52. Goloboff PA. Refining phylogenetic analyses: phylogenetic analysis of morphological data: volume 2. Boca Raton: CRC Press; 2022.
53. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 1987;4:406–25.
54. Couette S, Escarguel G, Montuire S. Constructing, bootstrapping, and comparing morphometric and phylogenetic trees: a case study of New World monkeys (Platyrrhini, Primates). J Mammal. 2005;86:773–81.
55. Lockwood CA, Kimbel WH, Lynch JM. Morphometrics and hominoid phylogeny: Support for a chimpanzee–human clade and differentiation among great ape subspecies. Proc Natl Acad Sci. 2004;101:4356–60.
56. Scalici M, Panchetti F. Morphological cranial diversity contributes to phylogeny in soft-furred sengis (Afrotheria, Macroscelidea). Zoology. 2011;114:85–94.
57. Watanabe A, Slice DE. The utility of cranial ontogeny for phylogenetic inference: a case study in crocodylians using geometric morphometrics. J Evol Biol. 2014;27:1078–92.
58. Sneath PH, Sokal RR. Numerical Taxonomy. San Francisco: W. H Freeman; 1973.
59. * Celik M, Cascini M, Haouchar D, Van der Burg C, Dodt W, Evans AR, et al. A molecular and morphometric assessment of the systematics of the *Macropus* complex clarifies the tempo and mode of kangaroo evolution. Zool J Linn Soc. 2019;186:793–812.
60. Frédérich B, Pilet A, Parmentier E, Vandewalle P. Comparative trophic morphology in eight species of damselfishes (Pomacentridae). J Morphol. 2008;269:175–88.
61. Rzhetsky A, Nei M. A simple method for estimating and testing minimum-evolution trees. Mol Biol Evol. 1992;9:945–67.
62. Waddell PJ. Extended distance-based phylogenetic analyses applied to 3D *Homo* fossil skull evolution. arXiv preprint. 2014. https://doi.org/10.48550/arXiv.1501.00019.
63. Waddell PJ. Expanded distance-based phylogenetic analyses of fossil *Homo* skull shape evolution. arXiv preprint. 2015. https://doi.org/10.48550/arXiv.1512.09115.
64. Revell LJ, Mahler DL, Reynolds RG, Slater GJ. Placing cryptic, recently extinct, or hypothesized taxa into an ultrametric phylogeny using continuous character data: A case study with the lizard *Anolis roosevelti*. Evolution. 2015;69:1027–35.
65. Felsenstein J. Maximum-likelihood estimation of evolutionary trees from continuous characters. Am J Hum Genet. 1973;25:471.
66. O'Dea RE, Lagisz M, Jennions MD, Koricheva J, Noble DWA, Parker TH, et al. Preferred reporting items for systematic reviews and meta-analyses in ecology and evolutionary biology: a PRISMA extension. Biol Rev. 2021;96:1695–722.
67. Adams DC, Rohlf FJ, Slice DE. A field comes of age: geometric morphometrics in the 21st century. Hystrix. 2013;24:7–14.
68. Adams DC, Rohlf FJ, Slice DE. Geometric morphometrics: ten years of progress following the "revolution." Ital J Zool. 2004;71:5–16.
69. The EndNote Team. EndNote. Clarivate, Philadelphia, PA. https://endnote.com/. 2013.
70. * Gomez RO, Lois-Milevicich J. Phylogenetic signal in the skull of cowbirds (Icteridae) assessed by multivariate and cladistic approaches. Zool Anz. 2020;286:52–7.
71. * Cichocka JM, Bielecki A. Phylogenetic utility of the geometric model of the body form in leeches (Clitellata: Hirudinida). Biologia. 2015;70:1078–92.
72. * Vargas S, Breedy O, Guzman HM. The phylogeny of *Pacifigorgia* (Coelenterata, Octocorallia, Gorgoniidae): a case study of the use of continuous characters in the systematics of the Octocorallia. Zoosystema. 2010;32:5–18.
73. * de Bivort BL, Giribet G. A systematic revision of the South African Pettalidae (Arachnida : Opiliones : Cyphophthalmi) based on a combined analysis of discrete and continuous morphological characters with the description of seven new species. Invertebr Syst. 2010;24:371–406.
74. * Hendrixson BE, Bond JE. Evaluating the efficacy of continuous quantitative characters for reconstructing the phylogeny of a morphologically homogeneous spider taxon (Araneae, Mygalomorphae, Antrodiaetidae, *Antrodiaetus*). Mol Phylogenet Evol. 2009;53:300–13.
75. * Hardy CR, Moline P, Linder HP. A phylogeny for the African Restionaceae and new perspectives on morphology's role in generating complete species phylogenies for large clades. Int J Plant Sci. 2008;169:377–90.
76. * Edgar SM, Theriot EC. Phylogeny of *Aulacoseira* (Bacillariophyta) based on molecules and morphology. J Phycol. 2004;40:772–88.
77. Travouillon KJ, Butler K, Archer M, Hand SJ. Two new species of the genus *Gumardee* (Marsupialia, Macropodiformes) reveal the repeated evolution of bilophodonty in kangaroos. Alcheringa. 2022;46:105–28.
78. Wiens JJ. Character analysis in morphological phylogenetics: Problems and solutions. Syst Biol. 2001;50:689–99.
79. Paradis E, Claude J, Strimmer K. APE: analyses of phylogenetics and evolution in R language. Bioinformatics. 2004;20:289–90.
80. R Core Team. R: A Language and Environment for Statistical Computing (Version 4.2.2). Vienna, Austria: R Foundation for Statistical Computing; 2023. https://www.R-project.org/.
81. Smith MR. Information theoretic generalized Robinson-Foulds metrics for comparing phylogenetic trees. Bioinformatics. 2020;36:5007–13.
82. Smith MR & actions-user. TreeDist: distances between phylogenetic trees. R package. ms609/TreeDist: v2.7.0. Zenodo. 2023. https://doi.org/10.5281/zenodo.10043369.
83. Motani R, Schmitz L. Phylogenetic versus functional signals in the evolution of form–function relationships in terrestrial vision. Evolution. 2011;65:2245–57.
84. Pratt RC, Gibb GC, Morgan-Richards M, Phillips MJ, Hendy MD, Penny D. Toward Resolving Deep Neoaves Phylogeny: Data, Signal Enhancement, and Priors. Mol Biol Evol. 2009;26:313–26.
85. Díaz-Cruz JA, Alvarado-Ortega J, Ramírez-Sánchez MM, Bernard EL, Allington-Jones L, Graham M. Phylogenetic morphometrics, geometric

morphometrics and the Mexican fossils to understand evolutionary trends of enchodontid fishes. J S Am Earth Sci. 2021;111:103492.
86. Rohlf FJ, Chang W, Sokal R, Kim J. Accuracy of estimated phylogenies: effects of tree topology and evolutionary model. Evolution. 1990;44:1671–84.
87. Fiala KL, Sokal RR. Factors determining the accuracy of cladogram estimation: evaluation using computer simulation. Evolution. 1985;39:609–22.
88. Longhorn SJ, Pohl HW, Vogler AP. Ribosomal protein genes of holometabolan insects reject the Halteria, instead revealing a close affinity of Strepsiptera with Coleoptera. Mol Phylogenet Evol. 2010;55:846–59.
89. Tong KJ, Duchêne DA, Duchêne S, Geoghegan JL, Ho SYW. A comparison of methods for estimating substitution rates from ancient DNA sequence data. BMC Evol Biol. 2018;18:70.
90. Xiang C-Y, Gao F, Jakovlić I, Lei H-P, Hu Y, Zhang H, et al. Using PhyloSuite for molecular phylogeny and tree-based analyses. Meta. 2023;2:e87.
91. White WT, Hills SF, Gaddam R, Holland BR, Penny D. Treeness triangles: Visualizing the loss of phylogenetic signal. Mol Biol Evol. 2007;24:2029–39.

Publisher's Note