
Overcoming problems with the use of ratios as continuous characters for phylogenetic analyses

NICOLÁS MONGIARDINO KOCH IGNACIO M. SOTO & MARTÍN J. RAMÍREZ

Submitted: 2 January 2015
Accepted: 30 April 2015
doi:10.1111/zsc.12120

Mongiardino Koch, N., Soto, I. M., Ramírez, M. J. (2015). Overcoming problems with the use of ratios as continuous characters for phylogenetic analyses. —*Zoologica Scripta*, 00, 000–000.

The use of quantitative morphometric information for phylogenetic inference has been an intensely debated topic for most of the history of phylogenetic systematics. Despite several drawbacks, the most common strategy to include this sort of data into phylogenetic studies is the use of ratios, that is quotients between morphometric variables. Here, we discuss one particular problem associated with such methodology: the fact that the often arbitrary election of which variable serves as numerator and which as denominator affects the phylogenetic outcome of the analysis. We describe the cause for such an effect, and study its implications with the use of several published data matrices. Alternative coding schemes for ratio characters result in very different phylogenetic hypotheses, an effect that may even be strong enough to affect studies that combine continuous and discrete morphological information. Some of the resulting incongruence is produced by the differences in magnitude of the continuous characters involved, although different rescaling techniques are shown to decrease, but not eliminate, the confounding effect. To eliminate such problematic effect, ratios should be either log-transformed before their use or replaced by more effective ways to capture morphometric information.

Corresponding author: *Nicolás Mongiardino Koch, Intendente Güiraldes 2160, Ciudad Universitaria (C1428 EHA), Capital Federal, Buenos Aires, Argentina. E-mail: nmongiar@ege.fcen.uba.ar*

Nicolás Mongiardino Koch and Ignacio M. Soto, Departamento de Ecología, Genética y Evolución, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires and Instituto de Ecología, Genética y Evolución de Buenos Aires (IEGEB – CONICET/UBA), Intendente Güiraldes 2160, Ciudad Universitaria (C1428 EHA), Capital Federal, Buenos Aires, Argentina, E-mail: soto@ege.fcen.uba.ar

Martín J. Ramírez, Museo Argentino de Ciencias Naturales – CONICET, Ángel Gallardo 490 (C1405 DJR), Capital Federal, Buenos Aires, Argentina. E-mail: ramirez@macn.gov.ar

Introduction

Background

The disciplines of morphometrics and phylogenetic systematics have a history of complex and uneasy dialogue, which derives, at least partially, from the underlying differences in their theoretical bases (Rae 2002; Jensen 2003; Clouse *et al.* 2011). Despite the fact that both fields share a common interest in the analysis of morphology and the nature of morphological variation (MacLeod 2002), their differences in focus and approach have limited the assimilation of each other's developments (MacLeod & Forey 2002). For example, the successes of the so-called revolution of geometric morphometrics (Rohlf & Marcus 1993) have proven to be especially difficult to incorporate into phylogenetic analyses (Adams *et al.* 2004, 2013).

One of the main threads of controversy between systematists and morphometricians relates to the question of whether morphometric variables are valid characters for parsimony analyses (discussed extensively in Catalano *et al.* 2010). Much of this discussion is in fact part of the wider debate regarding the use of continuous characters in a cladistic framework (Jensen 2003; González-José *et al.* 2008; Smith & Hendricks 2013), yet certain types of morphometric data (such as partial warps, landmarks and principal components of shape, among others) have received special criticism (Naylor 1996; Rohlf 1998; Monteiro 2000; MacLeod 2001; Rae 2002; Jensen 2003; Adams *et al.* 2011).

However, the debate has not been exclusively restricted to the kind of data used in geometric morphometrics. In fact, the use of simple ratios (*i.e.* quotients between

morphometric measurements) as cladistic characters has also raised an intense debate. As has been noted, the use of ratios has become thoroughly ingrained into scientific methodology (Atchley *et al.* 1976), in part given their capability to present, in concise and easily interpretable ways, quantities that are otherwise non-comparable (Leirmann *et al.* 2004). Consequently, their use is extensive throughout the taxonomic and systematic literatures (Reyment *et al.* 1984; Winston 1999; Schuh 2000). Among these fields, the purpose behind the use of ratios is either to remove the body size component to the observed variation of an anatomical feature (Atchley *et al.* 1976) or to represent shapes or proportions (Baur & Leuenberger 2011) that are useful to identify and describe living organisms. However, as Atchley *et al.* (1976) originally demonstrated, ratios can present several undesirable statistical properties. Among these, the most disturbing for cladistic studies is that ratios may both fail to eliminate size information while introducing at the same time new relationships between previously uncorrelated characters. Both of these issues violate the fundamental principle of character independence (Doyle 1997). To limit such caveats in the use of ratios, some have proposed several techniques to produce adjusted ratios that maximize the elimination of size information (Hills 1978; Albrecht *et al.* 1993; Corruccini 1977), while others have focused on analyses of independence that eliminate co-varying continuous characters from the data matrix (de Bivort *et al.* 2010; Clouse *et al.* 2010). However, many

authors continue to use ratios as continuous characters without any sort of modification (see e.g. Mannion *et al.* 2013; Mongiardino Koch *et al.* 2015) probably because further treatment of the data (such as the use of PCA or independence analysis) produces less intuitive results. Likewise, a consensus on the appropriate methodology to apply is still lacking.

We discuss here a subsequent caveat towards the use of ratios in cladistic analyses: the fact that the often arbitrary choice between which measurement is used as numerator and which as denominator affects the phylogenetic outcome of the analysis. We illustrate this problem with a simple hypothetical example, and then demonstrate its importance with the analysis of several published data matrices. Finally, we propose the use of log-transformed ratios as a suitable replacement, given the fact that they allow the inclusion of classic morphometric variables that do not depend on the choice of numerator and denominator.

Example

The drawings in Fig. 1 represent five theoretical species of arthropods, four of which belong to the same genus (A–D) and the fifth to a different one (E, used as outgroup). The monophyly of the ingroup is considered well supported due to the presence of a central furrow (absent in E and other close outgroups), but the relationships within the genus are yet to be resolved. Morphological variation at

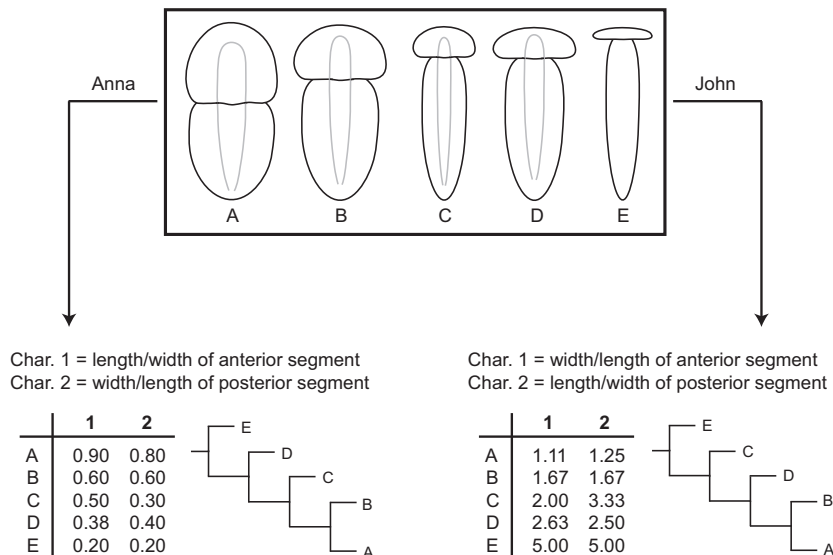


Fig. 1 Example showing the effect of using alternative codifications of ratios as characters in a phylogenetic analysis. Taxa A–E are hypothetical arthropods with a two-segment body plan, and with scarce morphological differences except for variation in body proportions. Two different systematists (Anna and John) end up building a two-character matrix with the same information, although the resulting characters are reciprocals. The respective phylogenetic analyses produce different hypotheses of relationships even after eliminating differences in scale (see text for explanation).

such low taxonomic levels often take the form of quantitative differences in general morphology, with few (or none) discrete morphological attributes differing between species (Rae 1998; MacLeod & Forey 2002). In an effort to understand the phylogenetic relationships between them, two different systematists (referred to as Anna and John) take four different measurements, and decide to combine them into two continuous characters that express, in the form of ratios, the shape of two different body regions. However, given that the choice between numerator and denominator is essentially arbitrary (e.g. length/width is as much a representation of the shape of an object as width/length is) each one ends up building a two character matrix in which all characters are the reciprocal of the ones built by the other (e.g. character 2 in John's matrix = $1/\text{character 2}$ in Anna's matrix, see Fig. 1). A tree search using implicit enumeration results in a single most parsimonious tree for each matrix, yet the trees are different (Fig. 1); none of them can be considered better in any sense, as both are built using the same information, and there seems to be no logical way in which to choose which of the original measurements should become the numerator and which the denominator.

At least two different phenomena are contributing to this effect. The first one is that inverting a character results in a change in the magnitude in which such character is varying. In general, characters that vary in higher orders of magnitude will dominate the analysis (Thiele & Ladiges 1988; Wiens 2001; Goloboff *et al.* 2006; Baur & Leuenberger 2011), a problem that has been referred to as the 'issue of scaling'. Therefore, inverting the way in which characters are coded will transform uninformative characters into dominant ones and vice versa, modifying their relative influences on the final topology. However, range rescaling (*i.e.* the standardization of the ranges of variation of all continuous characters to a common magnitude) has been shown to decouple the character's original magnitude from its influence in determining the optimal topology (see Mongiardino Koch *et al.* 2015). Nonetheless, if both matrices in our example are rescaled so that the overall range of variation (*i.e.* maximum–minimum value) of all characters is equal to 1, the same two different trees are obtained as the sole optimal topologies. Therefore, a second issue that is unrelated to differences in scale is determining the discrepancy.

When carefully inspected, the two characters used by both systematists present conflicting evidence for the relationships of the organisms under study. Character 1 proposes D as the basalmost species of the genus, while character 2 places species C closest to the root. Under the coding preferred by Anna, the conflict is resolved in favour of character 1, given that contradicting this character

results in 0.12 homoplastic steps, while contradicting character 2 results in only 0.1 extra steps (number of steps for continuous characters on a given branch are calculated as the difference in character states, the natural extension of Farris' (1970) algorithms for continuous variables; see Goloboff *et al.* 2006). This renders the first phylogenetic hypothesis as the most parsimonious one with a total length 0.02 steps shorter. However, when characters are inverted, the cost of contradicting character 2 becomes larger (0.833 vs. 0.631) and the second hypothesis is therefore preferred, by a total cost difference of 0.202. The reason why this happens is that, given the nonlinear property of ratios, the magnitude of the differences between character states is modified depending on the way ratios are built. For example, the magnitude of the difference between species A and C for character 1 (Fig. 1, Anna's matrix) is larger (0.4 steps) than the one that separates C and E (0.3 steps). However, when characters are inverted (Fig. 1, John's matrix), the second distance becomes larger than the first one (3.00 vs. 0.89 steps). This does not modify the relationships that are inferred using this sole character, but will affect the result when conflicting characters are included in the matrix, as is the case for the matrices in Fig. 1 (and all real data sets as well). Linearly rescaling characters does not eliminate such change in the magnitude of the differences between states of the same character, which is the reason why the two matrices continue to produce different trees even after characters are rescaled.

However, applying logarithms to ratio characters results in achieving linearity between a quotient and its inverse form. As a consequence, absolute differences between character states no longer depend on the way ratios are built. After eliminating the arbitrariness of ratio construction by using log ratios as characters, both systematists would have obtained the tree on the right of Fig. 1.

Materials and methods

In order to analyse the extent to which this phenomenon is affecting phylogenetic inference, eight real data sets (listed in Table 1) were explored using the software TNT (Goloboff *et al.* 2008). Two of these matrices (taken from Clouse *et al.* 2010 and de Bivort *et al.* 2010) were built using exclusively continuous characters, whereas the remaining ones (taken from Coyle 1995; Pereyra & Mound 2009, 2010; Kammerer *et al.* 2013; Mannion *et al.* 2013 and Mongiardino Koch *et al.* 2015) combined continuous and discrete morphological characters (see Table 1 for references), with quantitative information representing between 12 and 40% of the data. In all cases, we retained only characters coded as ratios. Furthermore, some of the matrices employed showed a large portion of taxa with high levels of missing data (sometimes as much as 100%

Data set		Matrix size (characters × taxa)	Shared internal node similarity			
Author	Group studied		Normal	Original rescaling	Range rescaling	Z-scoring
Coyle (1995)	Ischnothelinae, Araneae	19 × 19	0.402	0.564	0.771	0.790
Pereyra & Mound (2009)	<i>Cranothrips</i> , Thysanoptera	18 × 14	0.289	0.432	0.540	0.683
de Bivort <i>et al.</i> (2010)	Pettalidae, Opiliones	84 × 52*	0.163	0.302	0.543	0.604
Clouse <i>et al.</i> (2010)	Stylocellidae, Opiliones	60 × 66	0.131	0.413	0.437	0.425
Pereyra & Mound (2010)	<i>Desmothrips</i> , Thysanoptera	16 × 27	0.223	0.425	0.948	0.534
Mannion <i>et al.</i> (2013)	Titanosauriformes, Sauropoda	68 × 35*	0.133	0.262	0.350	0.379
Kammerer <i>et al.</i> (2013)	Anomodontia, Therapsida	18 × 36*	0.156	0.482	0.566	0.612
Mongiardino Koch <i>et al.</i> (2015)	Neriidae, Diptera	70 × 54	0.154	0.383	0.702	0.660
		Mean	0.206	0.408	0.607	0.586

*Mark datasets for which the number of taxa was reduced.

for the continuous partition), or had such a large taxa-to-characters ratio that the tree search algorithms used were unable to find the most parsimonious trees. Therefore, some taxa were eliminated for some of the matrices included in the study. In all cases, taxa that were pruned were those with the largest amount of missing/inapplicable data. The final dimensions of all matrices are shown in Table 1. TNT can deal with continuous characters with values ranging from 0 to 65, using up to three decimal places. This means that the smallest state value that can be inverted without exceeding the upper limit is 0.016. In case characters showed state values lower than that, the magnitude needed to transform the smallest state value into 0.016 was added to all taxa for that specific character. This allowed us to include as much information as possible without significantly modifying the relationships between states for different species.

A TNT script was programmed which explored the phylogenetic effect of the way in which ratios were built. Given that the choice of numerator and denominator for each character is both arbitrary and independent of the choice made for the other characters, 2^n matrices can be built for the same information, where n is the number of characters derived from ratios. For each data set, state values for all characters were inverted, and 100 matrices were built by randomly choosing between the original and the inverted coding for each character. Given that rescaling techniques may eliminate one of the issues that arise from inverting characters, these 100 matrices per data set were then subjected to four different treatments. Tree searches were either performed with the unmodified state values (henceforth, the ‘normal’ treatment), or after one of three different rescaling procedures were applied. These included (i) rescaling only inverted characters so as to present the same range of variation as the original corresponding

Table 1 Matrices employed in the analysis and the resulting values of topological differences obtained by modifying the way in which ratios were coded. Values of similarity are expressed as mean number of shared internal nodes

character (‘original rescaling’); (ii) ‘range rescaling’, standardizing the ranges of variation of all characters to unity; and (iii) ‘z-scoring’, normalizing the data by conversion into z-scores (*i.e.* subtracting the character’s mean from every state value and dividing the result by the character’s standard deviation). The resulting z-values were made positive by subtracting the smallest state value to all of them. In this last case, all state values expressed as ranges were first transformed into single values by averaging the extremes of the interval. Range rescaling and z-scoring were explored as they are two of the most widely used rescaling techniques. On the other hand, the ‘original rescaling’ treatment was only included to show that differences in scale produced by inverting characters are responsible for only part of the incongruence, but that the elimination of the scale difference introduced does not solve the issue. A driven tree search (Goloboff *et al.* 2008) using new technologies was then performed under equal weights, using each of the 100 matrices per data set under all four treatments, starting in each case with 10 initial sequences and continuing to search until minimum length was found 10 times. Parsimony analysis of matrices built from continuous characters usually result in only one optimal topology (Bardin *et al.* 2014); hence, only one tree per matrix was saved.

All 4950 pairwise comparisons of the resulting 100 trees per treatment were explored with two measures of topological congruence: unweighted SPR distances (Goloboff 2007) using 15 000 replicates and 10 stratifications, and number of shared internal nodes. Compared trees were always binary, so this last measure is similar to the Robinson–Foulds distance (Robinson & Foulds 1981). As the absolute value of both measures depends on the number of taxa included in the data set, both were standardized and expressed as values of tree similarity. This was performed by dividing both measures by the largest possible value,

and in the case of SPR distances, subtracting the resulting value from 1. A value of similarity equal to 1 is therefore obtained when comparing identical trees, and the number becomes smaller as trees have less in common. Furthermore, matrices were also directly compared using the incongruence length difference (ILD; Mickevich & Farris 1981). The same 100 matrices were fused in pairs in all 4950 combinations, and a tree search under the same parameters was performed using each of these combined matrices. ILD values were calculated as in Wheeler *et al.* (2006). By expressing values of incongruence as a percentage of length increase, this measure allowed us to compare trees built using different matrices. The ILD was only applied under the three rescaling treatments, given that length increases are only comparable when ranges of continuous characters have been standardized. These three values were taken to represent the degree of discordance between matrices, and the mean value of the first two measures can be interpreted as the average topological difference that would have resulted from choosing a different, equally valid way in which to code the characters employed. Furthermore, the degree to which a given strategy can cope with the arbitrariness involved in the coding of ratios can be seen by comparing the resulting distributions of SPR distances, number of shared internal nodes and ILD scores. In this sense, higher mean values of tree similarity are the result of obtaining more congruent topologies, while lower mean values of ILD reflect reduced conflict between alternative coding schemes. This was also visually explored with the use of within-strategy majority-rule consensus.

Besides studying the behaviour of ratio characters in isolation, we studied whether this phenomenon had an impact on phylogenetic analyses performed with combined discrete-continuous data sets. Given that the final effect on combined data sets likely depends on too many factors (such as the size of the data set, the proportion of ratio characters, the phylogenetic signal of the characters used and their interaction with weighting, scaling and other treatments), we chose to use only the data sets of Pereyra & Mound (2009, 2010), dealing with the relationships of two different genera of Thysanoptera, as examples. The small size of these matrices (14–27 terminals), their high percentage of continuous characters (33–40%) and the fact that all of these are derived from ratios, made them suitable for this sort of analyses. Two aspects of the interaction between the discrete partition and alternative codings of the quantitative partition were studied (i) the phylogenetic hypothesis supported by a simultaneous analysis of both partitions and (ii) the degree to which continuous characters provide additional support for the tree found using only the discrete partition. Both analyses were performed

under implied weighting (Goloboff 1993; using the same values of the constant of concavity k as the authors did) and were replicated after quantitative characters were range rescaled. In the first case, the discrete partition was combined with 100 alternative codings of the continuous characters, resulting from randomly inverting characters (as explained above). Tree searches were then performed using 100 replicates of RAS + TBR followed by ratchet. Resulting trees were summarized using majority-rule consensus. In the second case, symmetric resampling values (Goloboff *et al.* 2003) were calculated for the tree derived from the discrete partition, using both discrete characters only and discrete + continuous characters and summarizing results as changes in node frequency after each of the 100 pseudoreplicate matrices of continuous characters were added. This technique was originally proposed by Goloboff *et al.* (2006) as a way to evaluate the conflict/concordance between subsets of data. Resampling values were estimated for each pseudoreplicate as in Pereyra & Mound (2009, 2010), using 1000 replicates of 20 RAS + TBR and saving up to 5 trees per replicate, and were presented as absolute frequencies.

Results

Both measures of topological similarity rendered very similar results. In fact, the two measures were found to be significantly correlated ($P < 0.0001$), showing a very strong covariation, with an R^2 of 0.93 (Fig. 2A). This means that, at least for the data sets employed in this study, the significant investment of time in calculating SPR distances (a computationally intensive algorithm that produces heuristic results, see Goloboff 2007) was largely redundant, providing the same information as the tree similarity derived from the number of shared internal nodes, which can be calculated orders of magnitude faster. Because of this strong correlation, and given that the number of shared internal nodes is a much more intuitive quantity, only the values of similarity derived from this measurement are reported below. Furthermore, both measures of topological comparison were found to be significantly correlated ($P = 0.032$ for SPR distances, $P = 0.003$ for shared nodes) with the values obtained using ILD (Fig. 2B). Although these covariations were not as strong as that between topological quantities (with $R^2 \leq 0.35$), the significant regressions prove that more dissimilar tree topologies are, at least in part, the result of higher conflict at the data set level.

The values of average tree similarity per strategy for each data set are shown in Table 1. In all cases, the smallest values of similarity were obtained when characters were not rescaled, with values ranging from 0.131 to 0.402. On the other hand, the highest values were always obtained either after applying range rescaling or after transforming

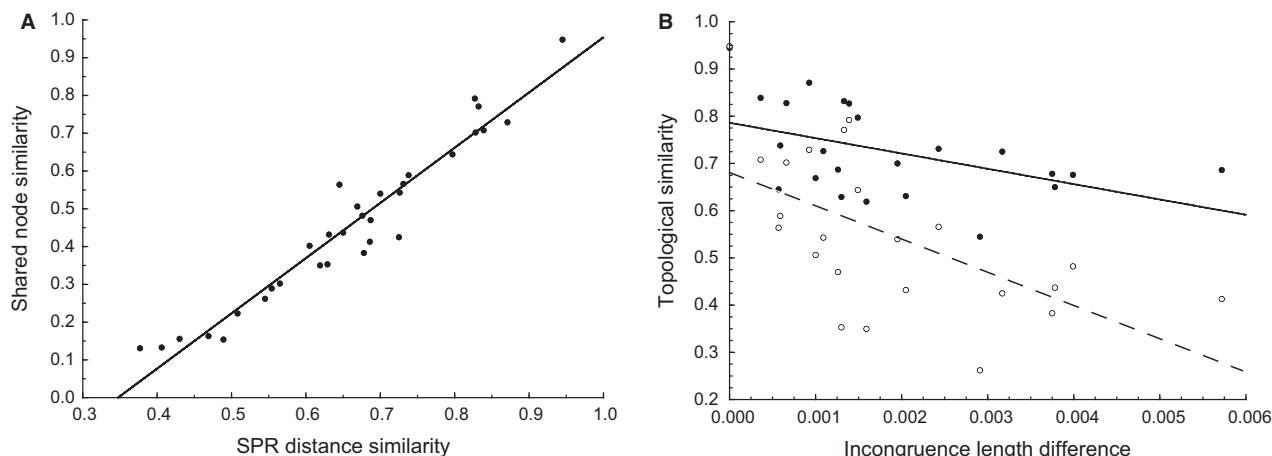


Fig. 2 A–B. Correlation analysis between measures of congruence employed (lines represent significant linear regressions). —A. Regression between topological measures of tree similarity. —B. Regressions between measures of tree similarity and values of ILD (the solid line and black circles represent the data drawn from SPR distances, while the dashed line and empty circles represent the data points obtained using the number of shared internal nodes). Datapoints represent the mean value for each measure of congruence obtained for all dataset under the different treatments ($n = 32$).

characters to z -scores, with the ‘original rescaling’ treatment consistently rendering an intermediate improvement in average tree similarity. The intermediate values of the ‘original rescaling’ indicate that only part of the incongruence arises from the modification of the relative influences of characters introduced by the inversion procedure. The average similarity improvement for the two common rescaling techniques implemented was fairly similar, with an average increase of 40% for range rescaling and 38% for z -scoring (although in 6 of 8 cases z -scoring outperformed range rescaling). Beyond these common trends, the behaviour of each data set was largely individual, with some responding better to range rescaling (e.g. *Desmothrips* data set), others to z -scoring (e.g. *Chranothrips* dataset), and yet others performing almost identically under both treatments (e.g. the *Styllocelidae* and *Ischnothelinae* matrices). The increase in magnitude of the congruence parameters after characters are rescaled is not significantly correlated with the matrices’ dimensions (*i.e.* number of taxa, number of characters and their product) and probably depends on the information content of each matrix. The only relationship found was a significant and negative correlation between the number of taxa and the mean similarity under the ‘normal’ treatment ($P = 0.03$). This reveals that, unless characters are rescaled, data sets with many terminals may be more susceptible to differences in the way ratios are built.

When a majority-rule consensus is applied to the 100 trees obtained for each treatment, a similar pattern arises. When characters are not rescaled, the resulting consensus is generally found to consist of very few resolved nodes (as seen in Fig. 3A), most of which have only marginal

frequencies. In contrast, consensus obtained after applying range rescaling or z -scoring show a dramatic increase in both node resolution and frequency (see Fig. 3B–C and values in caption). Nonetheless, it should be noted that both rescaling techniques are still far from completely eliminating the noise introduced by the decisions made during ratio construction. After this noise is eliminated through the use of log ratios, the resulting tree has a similar topology to the ones found using regular quotients (see black dots in Fig. 3B–C), yet stripped from the underlying arbitrariness.

The analysis of the *Desmothrips* and *Chranothrips* matrices are illustrative of the range of effects that different codifications of ratio characters can produce on the analysis of the combined data. On one side, the majority-rule consensus of the trees corresponding to the 100 pseudoreplicates of combined *Chranothrips* matrices was poorly resolved when characters were not rescaled and the search was performed under equal weights. Under such conditions, only 15% of internal nodes were resolved in the consensus, with most taxa collapsing in a single polytomy. However, both range rescaling and implied weighting, as well as the simultaneous use of the two, eliminated all incongruence between pseudoreplicates, and a unique tree was always found for the range of concavities used by the authors ($k = 4–7$). On the other hand, when using the *Desmothrips* matrix, similar levels of topological differences were obtained when characters were not rescaled, independent on whether implied weighting was used ($k = 10–19$, as the authors) or not (see Fig. 4A as an example of the topological differences found). After rescaling, a unique topology

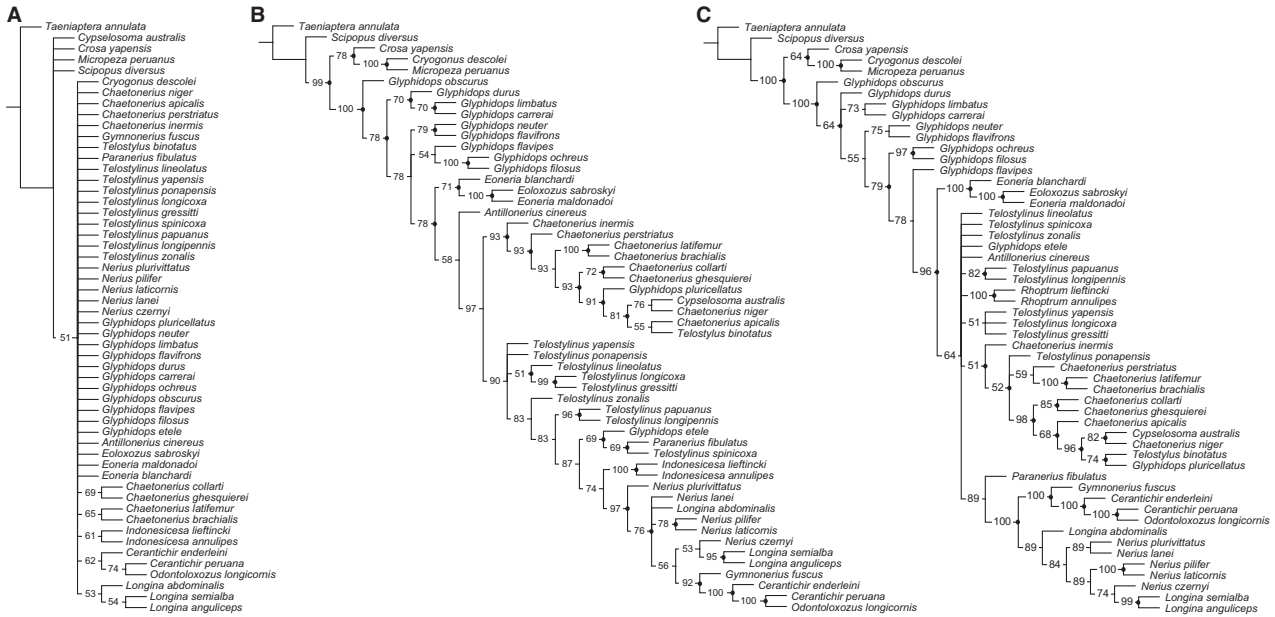


Fig. 3 A–C. Majority-rule consensus of the trees derived from the 100 pseudoreplicates of the continuous partition of the Neriidae dataset under different treatments. —A. Normal treatment. —B. Range rescaling treatment. —C. Z-scoring treatment. Both rescaling strategies increase the robustness of the phylogenetic inference towards changes in the way ratio characters are built, yet falling short of eliminating the problem altogether. From left to right, number of resolved nodes: 8, 46 and 40; mean node frequency: 61, 83 and 84. Black dots show the groups that are also recovered when using log ratios and applying the same rescaling techniques. Although not shown, the ‘original rescaling’ treatment once again shows limited improvement in tree congruence (number of resolved nodes = 23, mean node frequency = 66).

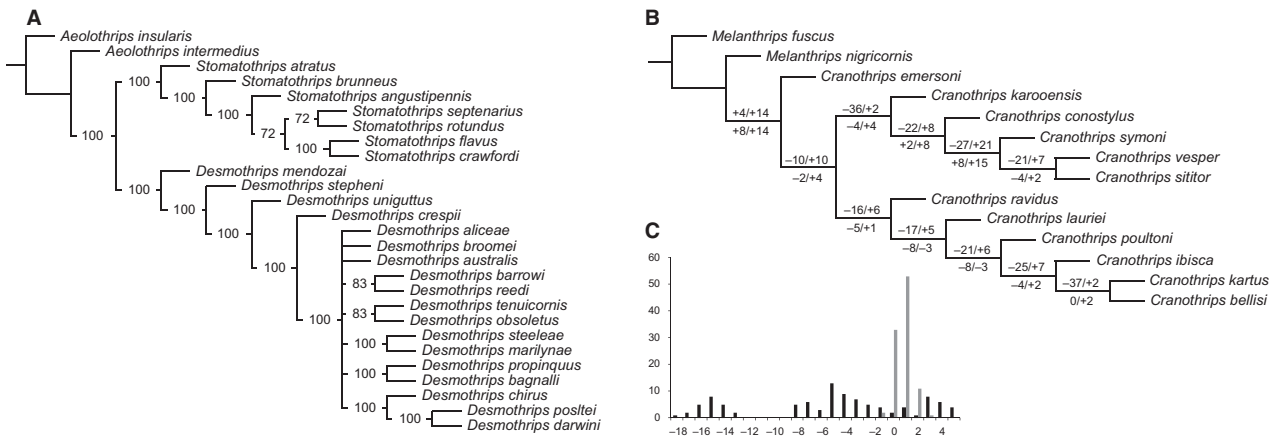


Fig. 4 A–C. Interaction between discrete and alternative pseudoreplicates of the continuous partition. —A. Majority-rule consensus of the 100 optimal trees obtained from the pseudoreplicates of the *Desmothrips* dataset under $k = 19$ and without rescaling. After range rescaling characters and searching under the same parameters, a unique topology was always found. —B. Tree obtained using the discrete partition of the *Cranathrips* dataset, with the effect on node support after adding the continuous partition. Values are shown as ranges of change in group support, expressed as raw frequencies under symmetric resampling. Numbers below and above branches represent, respectively, changes obtained with and without range rescaling characters. —C. Mean change in group support after adding continuous characters. Black bars are values obtained without rescaling characters, grey ones are values after rescaling.

was found both under equal weights and implied weighting using the largest values of k employed, although some alternative topologies were still found, for example, under

$k = 10$. The analysis of these data sets confirmed that, at least under certain conditions, the differences in the phylogenetic signal contained by alternative constructions

ratio characters (which result in the topological differences listed in Table 1) are strong enough to modify the result of combined studies. In both cases, range rescaling largely mitigates the effect, but whether or not this results in the same topology being optimal under different ratio constructions also depends on additional parameters, such as the constant of concavity of implied weighting.

Not only do alternative ratio characters modify the result of combined analysis, they also differ widely on the support they provide to groups defined using only the discrete partition. As shown in Fig. 4B–C, some of the 100 alternative continuous matrices built using the *Chranothrips* dataset improved the average node support of clades obtained after analysing the discrete partition, while others strongly decreased it. On average, the effect of combining partitions proved to be negative when characters were not rescaled, and slightly positive after range rescaling (with a mean change in node frequency of -4.5 and $+1.3$, respectively). However, the striking result is the degree to which different pseudoreplicates differ in the support provided. When using characters without rescaling them, the change in mean node support can vary tremendously (Fig. 4C), with approximately 30% of pseudoreplicates increasing mean node frequency and 70% decreasing it. The magnitude of these changes can be even more dramatic when considering the effect produced on particular nodes (Fig. 4B). After characters are rescaled, pseudoreplicates tend to perform more similarly, with the standard deviation of the distributions in Fig. 4C decreasing nearly 10 times. Nonetheless, it should be noted that the change in node support after adding rescaled continuous characters can still be strongly dependent on character coding (Fig. 4C, range of values below branches). The *Desmothrips* data set shows a similar pattern, although results are less extreme. The average change in node frequency after adding the pseudoreplicates of the continuous partition has a mean value of $+1.58$ when characters are not rescaled, and $+0.30$ after range rescaling. This decrease is accompanied by a threefold reduction in the standard deviation of distributions. Finally, when ratios are log-transformed, the mean change in node frequency obtained is $+2.45$ for the *Chranothrips* data set and $+1.38$ for the *Desmothrips* data set.

Discussion

The so-called continuous characters were generally disregarded as sources of information during the early days of phylogenetic systematics. In spite of this, the use of morphometric measurements in phylogenetic analysis has increased greatly in the past few years, as shown by the number of recently published data sets that are readily available (see de Bivort *et al.* 2010 for a synthetic description of the developments leading to this change in

perspective). Despite this trend in use increase, methodological consensus on how to implement morphometric characters has not been reached. For example, most work nowadays focuses on the use of continuous characters as such (Goloboff *et al.* 2006) although some researchers still prefer to use discretizing techniques (see e.g. Vargas *et al.* 2010; Castañeda & de Queiroz 2013; Rae & Buckley 2014). Furthermore, among the first group, some favour the use of independence and principal component analyses (de Bivort *et al.* 2010; Clouse *et al.* 2010), others choose to apply only rescaling techniques (Carrasco *et al.* 2012; Mongiardino Koch *et al.* 2015), and yet others use quantitative data without any special treatment (Bert & Bersac 2013; Kammerer *et al.* 2013).

Beyond the precise methodology used for continuous character implementation, most of the recent phylogenetic analyses that use morphometric measurements include these in the form of ratios. This approach has raised several concerns, among which the lack of character independence (Atchley *et al.* 1976) and the fact that ratios may conceal information regarding the evolution of characters (given that different evolutionary changes can possibly lead to the same ratio value, for example Hormiga *et al.* 2000), are among the most serious. Nonetheless, researchers either consider that their techniques successfully deal with these issues or are confident that the benefit of including this kind of data is greater than the harm (Lopardo *et al.* 2011).

We discuss here a different problem, that of the arbitrariness that is introduced into the study when deciding which measure constitutes the numerator and which the denominator of ratio characters. As stated above, whether one is describing the shape of a structure (e.g. length/width) or the relative dimensions or positions of different structures (e.g. length of A/length of B), the numerator/denominator choice is, in many cases, disputable. Given this general lack of evidence that can be used in order to prefer one coding over the other, the eventual decision becomes largely arbitrary. At first sight, this distinction seems completely trivial. Describing structure A as having half the size of structure B should render the same result than defining structure B as being twice as large as structure A. However, as shown by the example provided before, in case contradicting information is included in the data set, alternative codings will in fact result in different phylogenetic hypotheses. This occurs as a consequence of the nonlinearity of quotients, a phenomenon that is not modified by the standardization of the magnitude of characters. Although many have acknowledged the impact that alternative coding schemes have on phylogenetic analyses, a discussion that is certainly not restricted to the use of continuous data (e.g., Maddison 1993; Pleijel 1995; Wiens 2001), the sort of arbitrariness

described here seems to be even more pervasive than that shown by other sorts of characters. All of this leads to the following question: to what extent do alternative and seemingly equally valid options of character coding impact the phylogenetic outcome of the analysis?

As can be seen from the values shown in Table 1, the obtained topology is highly dependent on the way characters were coded. Alternative ways in which to build ratios from the same morphometric information result in very different phylogenetic hypothesis. Randomly inverting characters of the matrices studied results in obtaining, on average, trees that agree only on 21% of the total number of internal nodes (relative to a perfect congruence). The magnitude of these topological differences seriously underscores the validity of any individual tree obtained under a particular coding scheme. In fact, when a majority-rule consensus is applied to the 100 trees obtained under the different codings explored, the resulting tree is generally found to consist of very few resolved nodes, with most taxa collapsing into large polytomies (as seen in Fig. 3A). The particular phylogeny that is inferred from the use of continuous characters built from ratios is therefore determined both by the information content of the characters employed and by the coding scheme applied by the systematist. The relative importance of these two factors is difficult to evaluate, but evidence seems to support that the second one plays a significant role.

Once scale differences among characters are eliminated, average tree similarity increases drastically. The average magnitude of this effect is similar, independent of whether characters are rescaled to a common range or transformed to z-scores, both techniques increasing mean tree similarity in approximately 38–40%. This effect can also be appreciated by the difference in resolution of the majority-rule consensus (Fig. 3B–C). It is therefore evident that rescaling characters canalizes phylogenetic inference towards more similar results. However, and in congruence with the existence of two separate issues involved in inverting ratio characters, average tree similarity seldom approaches unity. In fact, mean values obtained when pulling together the data of the eight data sets employed show that, even after eliminating scale differences, tree similarity will only rise to about 60% of a perfect congruence. As a consequence, rescaling procedures, on which many phylogenetic studies rely, cannot be considered to solve the arbitrariness that is incorporated into the process of phylogenetic inference through the coding of ratio characters.

The degree to which this issue has shaped the phylogenetic reconstruction supported by the authors of the different data sets depends on multiple factors, including the relative proportion of continuous and discrete characters used, the treatment of the data (*i.e.* the use of rescaling

techniques, implied weighting, independence analyses), as well as the intrinsic robustness of each data set to changes in codification. Nonetheless, the two matrices we chose for comparison (*Chranothrips* and *Desmothrips*) provide evidence that coding strategies have a significant effect. Alternative ratio constructions can in fact modify phylogenetic inference in combined discrete + continuous datasets. Despite discrete characters providing a stable topological scaffold, some relationships show alternative resolutions depending entirely on the way ratios are built. This results in the collapse of several nodes of the ingroup in the majority-rule consensus shown in Fig. 4A. These topological changes do not modify the conclusions drawn by the authors of the study, but one can easily envision how this phenomenon could impact the taxonomic and evolutionary conclusions drawn from a phylogenetic analysis. Furthermore, alternative quantitative data sets can widely differ in terms of the support they provide to a phylogeny obtained using other sort of characters (Fig. 4B–C). In case characters are expressed as ratios, the resulting change in node support due to their inclusion can be completely determined by the way in which they are built. As can be seen in Fig. 4C, we detected a change in average node support (*i.e.* averaging the change in support across all nodes of the tree) that ranged from -17 to $+5.6$ when continuous characters were added. Furthermore, individual nodes were subject to almost equiprobable increases and decreases in values of support. Such differences in outcome do not allow drawing any conclusions regarding the phylogenetic information of the quantitative characters employed. Only after characters were range rescaled did topologies resulting from combined analysis become alike, and mean support values obtained from different pseudoreplicates converge to similar values. How much of the original noise is eliminated by the rescaling techniques will however depend on the dataset employed, with results ranging from almost imperceptible to those shown in Fig. 4.

However, it should be pointed out that not all ratios are equally susceptible to be coded in different ways. These issues are mostly relevant for characters representing shapes and relative proportions or positions, and do not apply as much to those built trying to reduce allometric variation (see Albrecht *et al.* 1993 for a similar distinction). Here, we preferred to treat them equally, as there is no mathematical impediment to do so, although the actual issue may only apply to a subset of the characters evaluated. However, the recent increase of interest in the use of continuous characters is driving more authors towards including undiscretized morphometric information in their datasets, increasing the relevance of studies dealing with the issues intrinsic to the different ways in which these may be analysed.

It could be argued that, with the increase in genetic information, morphometric characters will be a decreasing fraction of data, thus uninfluential in phylogenetic reconstructions (although there are many good reasons why morphology should not be left aside in the era of genomics, starting with fossil taxa and many museum specimens; see Giribet 2015). Furthermore, given that a ratio and its inverse are monotonically related, the optimizations of either way of scoring on a tree constructed from other evidence should be highly congruent. This may be so, but it is also true that many studies that employ morphometric data focus on character correlation, in which case the magnitudes of change are influential on the outcome of a correlation test, hence the issues here presented may have practical implications beyond the utilization of ratios as indicators of relationships.

Although rescaling techniques strongly reduced the differences in results obtained under different coding strategies, most of the times the effect was still far from disappearing. One way we envision of completely eliminating this confounding effect is to log-transform ratio characters, and use the resulting absolute magnitudes as state values. Such transformation has been already applied by some researchers using continuous data (e.g. Hornung-Leoni & Sosa 2008; Lehtonen 2009; Marks *et al.* 2011; De Gennaro & Scatagliini 2012) either due to differences in the variance of characters or in order to reduce the differences in scale. Although the effect of such transformation on the phylogenetic information contained by continuous characters has never been studied, applying logarithms to ratio characters would eliminate the need to choose between alternative codifications, allowing classic morphometric variables to be included in phylogenetic analysis without the associated issues described here. Nonetheless, log transformation of ratio characters should only be viewed as a way to avoid the problems associated with the numerator/denominator decision, not as an absolute solution to all problems that may arise during the codification and use of morphometric data. Another option could be to avoid the use of ratios during phylogenetic inference, replacing such information with the use of more sophisticated representations of biological shape. The arbitrariness involved in the election of the numerator/denominator is likely a consequence of the fact that ratios are very poor representations of shape, a limitation that originally contributed to the development of geometric morphometrics (Adams *et al.* 2013). Methods to incorporate morphometric information in parsimony analysis through the use of landmarks are now readily available (Catalano *et al.* 2010, 2015) and could be regarded as appropriate replacements for traditional morphometric variables in a phylogenetic framework.

To our knowledge, the only work that directly referred to the possibility of building two equally valid ratios from two morphometric variables is that of Frampton & Ward (1990). In a response to Atchley *et al.* (1976), these authors showed that the magnitude of the “undesirable properties” of ratios (*i.e.* non-normality and incorporation of both redundant information and non-useful variability) depend on how the numerator and denominator are chosen. Such decision, they proposed, should therefore be taken considering the lessening of these effects. However, they did not study whether such choice affected the resulting phylogenetic tree. The otherwise complete lack of references to this topic is seemingly a consequence of the fairly recent change in perception relating to the use of continuous characters in parsimony analysis.

Acknowledgements

The authors would like to thank Verónica Pereyra and Benjamin L. de Bivort for providing some of the matrices employed, Santiago Catalano and Benjamin L. de Bivort for their comments and ideas relating to this work, and Pablo Goloboff and Sebastián L. Rodríguez for assistance with the programming of scripts. We would also like to thank Pablo Goloboff and an anonymous reviewer for vital comments on an early version of this work. The Willi Hennig Society is acknowledged for the free version of the software TNT. This work was supported by grants from ANPCyT, Universidad de Buenos Aires and CONICET (PIP 943 to MJR). NMK was supported by a postgraduate fellowship from CONICET, IMS and MJR are members of Carrera del Investigador Científico (CONICET).

References

- Adams, D. C., Cardini, A., Monteiro, L. R., O’Higgins, P. & Rohlf, F. J. (2011). Morphometrics and phylogenetics: principal components of shape from cranial modules are neither appropriate nor effective cladistic characters. *Journal of Human Evolution*, *60*, 240–243.
- Adams, D. C., Rohlf, F. J. & Slice, D. E. (2004). Geometric morphometrics: ten years of progress following the ‘revolution’. *Italian Journal of Zoology*, *71*, 5–16.
- Adams, D. C., Rohlf, F. J. & Slice, D. E. (2013). A field comes of age: geometric morphometrics in the 21st century. *Hystrix, the Italian Journal of Mammalogy*, *24*, 7–14.
- Atchley, W. R., Gaskins, C. T. & Anderson, D. (1976). Statistical properties of ratios. I. Empirical results. *Systematic Zoology*, *25*, 137–148.
- Albrecht, G. H., Gelvin, B. R. & Hartman, S. E. (1993). Ratios as a size adjustment in morphometrics. *American Journal of Physical Anthropology*, *91*, 441–468.
- Bardin, J., Rouget, I., Yacobucci, M. & Cecca, F. (2014). Increasing the number of discrete character states for continuous characters generates well-resolved trees that do not reflect phylogeny. *Integrative Zoology*, *9*, 531–541.

- Baur, H. & Leuenberger, C. (2011). Analysis of ratios in multivariate morphometry. *Systematic Biology*, *60*, 813–825.
- Bert, D. & Bersac, S. (2013). Evolutionary patterns - tested with cladistics - and processes in relation to palaeoenvironments of the Upper Barremian genus *Gassendicerias* (Ammonitina, Lower Cretaceous). *Paleontology*, *56*, 631–646.
- Carrasco, P. A., Mattoni, C. I., Leynaud, G. C. & Scrocchi, G. J. (2012). Morphology, phylogeny and taxonomy of South American bothropoid pitvipers (Serpentes, Viperidae). *Zoologica Scripta*, *2*, 109–124.
- Castañeda, M. D. R. & de Queiroz, K. (2013). Phylogeny of the *Dactyloa* clade of *Anolis* lizards: new insights from combining morphological and molecular data. *Bulletin of the Museum of Comparative Zoology*, *160*, 345–398.
- Catalano, S., Ercoli, M. D. & Prevosti, F. J. (2015). The more, the better: the use of multiple landmark configurations to solve the phylogenetic relationships of musteloids. *Systematic Biology*, *64*, 294–306.
- Catalano, S., Goloboff, P. A. & Giannini, N. P. (2010). Phylogenetic morphometrics (I): the use of landmark data in a phylogenetic framework. *Cladistics*, *26*, 539–549.
- Clouse, M. C., de Bivort, B. L. & Giribet, G. (2010). A phylogenetic analysis for the South-east Asian mite harvestman family Stylocellidae (Opiliones: Cyphophthalmi) - a combined analysis using morphometric and molecular data. *Invertebrate Systematics*, *23*, 515–529.
- Clouse, R. M., de Bivort, B. L. & Giribet, G. (2011). Phylogenetic signal in morphometric data. *Cladistics*, *27*, 337–340.
- Corruccini, R. S. (1977). Correlation properties of morphometric ratios. *Systematic Zoology*, *26*, 211–214.
- Coyle, F. A. (1995). A revision of the funnel web mygalomorph spider subfamily Ischnothelinae (Araneae, Dipluridae). *Bulletin of the American Museum of Natural History*, *226*, 1–133.
- de Bivort, B. L., Clouse, R. M. & Giribet, G. (2010). A morphometrics-based phylogeny of the temperate Gondwanan mite harvestmen (Opiliones, Cyphophthalmi, Pettalidae). *Journal of Zoological Systematics and Evolutionary Research*, *48*, 294–309.
- De Gennaro, D. & Scatagliini, M. A. (2012). Análisis filogenético del género *Sacciolepis* (Poaceae, Panicoideae, Paniceae) con caracteres morfológicos. *Darwiniana*, *50*, 81–96.
- Doyle, J. J. (1997). Trees within trees: genes and species, molecules and morphology. *Systematic Biology*, *46*, 537–553.
- Farris, J. (1970). Methods for computing Wagner trees. *Systematic Zoology*, *19*, 83–92.
- Frampton, C. M. & Ward, J. M. (1990). The use of ratio variables in systematics. *Taxon*, *39*, 586–592.
- Giribet, G. (2015). Morphology should not be forgotten in the era of genomics—a phylogenetic perspective. *Zoologischer Anzeiger*, doi:10.1016/j.jcz.2015.01.003.
- Goloboff, P. A. (1993). Estimating character weights during tree search. *Cladistics*, *9*, 83–91.
- Goloboff, P. A. (2007). Calculating SPR distances between trees. *Cladistics*, *23*, 1–7.
- Goloboff, P. A., Farris, J. S., Källersjö, M., Oxelman, B., Ramírez, M. J. & Szumik, C. A. (2003). Improvements to resampling measures of group support. *Cladistics*, *19*, 324–332.
- Goloboff, P. A., Farris, J. S. & Nixon, K. C. (2008). TNT, a free program for phylogenetic analysis. *Cladistics*, *24*, 774–786.
- Goloboff, P. A., Mattoni, C. I. & Quinteros, A. S. (2006). Continuous characters analyzed as such. *Cladistics*, *22*, 589–601.
- González-José, R., Escapa, I., Neves, W. A., Cúneo, R. & Pucciarelli, H. M. (2008). Cladistic analysis of continuous modularized traits provides phylogenetic signals in *Homo* evolution. *Nature*, *453*, 775–778.
- Hills, M. (1978). On ratios - A response to Atchley, Gaskins and Anderson. *Systematic Zoology*, *27*, 61–62.
- Hornung-Leoni, C. T. & Sosa, V. (2008). Morphological phylogenetics of *Puya* subgenus *Puya* (Bromeliaceae). *Botanical Journal of the Linnean Society*, *156*, 93–110.
- Hormiga, G., Scharff, N. & Coddington, J. A. (2000). The phylogenetic basis of sexual size dimorphism in orb-weaving spiders (Araneae, Orbicularia). *Systematic Biology*, *49*, 435–462.
- Jensen, R. J. (2003). The conundrum of morphometrics. *Taxon*, *52*, 663–671.
- Kammerer, C. F., Fröbisch, J. & Angielczyk, K. D. (2013). On the validity and phylogenetic position of *Eubrachiosaurus browni*, a kannemeyeriiform dicynodont (Anomodontia) from Triassic North America. *PLoS ONE*, *8*, e64203.
- Lehtonen, S. (2009). Systematics of the Alismataceae - A morphological evaluation. *Aquatic Botany*, *91*, 279–290.
- Leirmann, M., Steel, A., Rosing, M. & Guttorp, P. (2004). Random denominators and the analysis of ratio data. *Environmental and Ecological Statistics*, *11*, 55–71.
- Lopardo, L., Giribet, G. & Hormiga, G. (2011). Morphology to the rescue: molecular data and the signal of morphological characters in combined phylogenetic analyses - a case study from mysmenid spiders (Araneae, Mysmenidae), with comments on the evolution of web architecture. *Cladistics*, *27*, 278–330.
- MacLeod, N. (2001). Landmarks, localization, and the use of morphometrics in phylogenetic analysis. In G. Edgecombe, J. Adrain & B. Lieberman (Eds) *Fossils, Phylogeny, and Form: An Analytical Approach* (pp. 197–234). New York: Plenum.
- MacLeod, N. (2002). Phylogenetic signals in morphometric data. In N. MacLeod & P. L. Forey (Eds) *Morphology, Shape and Phylogeny* (pp. 45–52). London: Taylor & Francis.
- MacLeod, N. & Forey, P. L. (2002). Introduction: morphology, shape & phylogeny. In N. MacLeod & P. L. Forey (Eds) *Morphology, Shape and Phylogeny* (pp. 45–52). London: Taylor & Francis.
- Maddison, W. P. (1993). Missing data versus missing characters in phylogenetic analysis. *Systematic Biology*, *42*, 576–581.
- Mannion, P. D., Upchurch, P., Barnes, R. N. & Mateus, O. (2013). Osteology of the late Jurassic Portuguese sauropod dinosaur *Lusotitan atalaiensis* (Macronaria) and the evolutionary history of basal titanosauriforms. *Zoological Journal of the Linnean Society*, *168*, 98–206.
- Marks, C. E., Newbigin, E. & Ladiges, P. Y. (2011). Comparative morphology and phylogeny of *Nicotiana* section *Suaveolentes* (Solanaceae) in Australia and the South Pacific. *Australian Systematic Botany*, *24*, 61–86.
- Mickevich, M. F. & Farris, J. S. (1981). The implications of congruence in *Menidia*. *Systematic Zoology*, *30*, 351–370.
- Mongiardino Koch, N., Soto, I. M. & Ramírez, M. J. (2015). First phylogenetic analysis of the family Neriidae (Diptera), with a study on the issue of scaling continuous characters. *Cladistics*, *31*, 142–165.

- Monteiro, L. R. (2000). Why morphometrics is special: the problem with using partial warps as characters for phylogenetic inference. *Systematic Biology*, 49, 796–800.
- Naylor, G. J. P. (1996). Can partial warp scores be used as cladistic characters? In L. F. Marcus (Ed) *Advances in Morphometrics* (pp. 519–530). New York: Plenum Press.
- Pereyra, V. & Mound, L. A. (2009). Phylogenetic relationships within the genus *Cranothrips* (Thysanoptera, Melanthripidae) with consideration of host associations and disjunct distributions within the family. *Systematic Entomology*, 34, 151–161.
- Pereyra, V. & Mound, L. A. (2010). Phylogenetic relationships within the genus *Desmothrips* (Thysanoptera, Aeolothripidae), an Australian genus of facultative flower-living predators. *Systematic Entomology*, 35, 306–317.
- Pleijel, F. (1995). On character coding for phylogeny reconstruction. *Cladistics*, 11, 309–315.
- Rae, T. C. (1998). The logical basis for the use of continuous characters in phylogenetic systematics. *Cladistics*, 14, 221–228.
- Rae, T. C. (2002). Scaling, polymorphism and cladistic analysis. In N. MacLeod & P. L. Forey (Eds) *Morphology, Shape and Phylogeny* (pp. 45–52). London: Taylor & Francis.
- Rae, T. C. & Buckley, A. (2014). TaxMan: An online facility for the coding of continuous characters for cladistic analysis. *arXiv preprint arXiv: 1411.4550*.
- Reyment, R. A., Blacklith, R. E. & Campbell, N. A. (1984). *Multivariate Morphometrics*, 2nd edn. New York: Academic.
- Robinson, D. & Foulds, L. (1981). Comparison of phylogenetic trees. *Mathematical Biosciences*, 53, 131–147.
- Rohlf, F. J. (1998). On applications of geometric morphometrics to studies of ontogeny and phylogeny. *Systematic Biology*, 47, 147–158.
- Rohlf, F. J. & Marcus, L. F. (1993). A revolution in morphometrics. *Trends in Ecology & Evolution*, 8, 129–132.
- Schuh, R. T. (2000). *Biological Systematics: Principles and Applications*. New York: Cornell University Press.
- Smith, U. E. & Hendricks, J. R. (2013). Geometric morphometric character suites as phylogenetic data: extracting phylogenetic signal from gastropod shells. *Systematic Biology*, 62, 366–385.
- Thiele, K. & Ladiges, P. Y. (1988). A cladistic analysis of *Angophora* Cav (Myrtaceae). *Cladistics*, 4, 23–42.
- Vargas, S., Breedy, O. & Guzman, H. M. (2010). The phylogeny of *Pacifigorgia* (Coelenterata, Octocorallia, Gorgoniidae): a case study of the use of continuous characters in the systematics of the Octocorallia. *Zoosystema*, 32, 5–18.
- Wheeler, W. C., Ramirez, M. J., Aagesen, L. & Schulmeister, S. (2006). Partition-free congruence analysis: implications for sensitivity analysis. *Cladistics*, 22, 256–263.
- Wiens, J. J. (2001). Character analysis in morphological phylogenetics: problems and solutions. *Systematic Biology*, 50, 689–699.
- Winston, J. E. (1999). *Describing Species: Practical Taxonomic Procedure for Biologists*. New York: Columbia University Press.