



# Phylogenetic analysis of the *Belostoma plebejum* group *sensu* Nieser (Insecta, Hemiptera, Belostomatidae): the effect of adding continuous characters on its accuracy

José R. I. Ribeiro<sup>1</sup>, Augusto Ferrari<sup>2,3</sup>

<sup>1</sup> Universidade Federal do Pampa (UNIPAMPA), Campus São Gabriel. Rua Aluizio Barros Macedo, BR 290, km 423, São Gabriel, Rio Grande do Sul State, Brazil

<sup>2</sup> Laboratório de Entomologia, Sistemática e Biogeografia (LESB), Universidade Federal do Rio Grande (FURG), Rio Grande, Rio Grande do Sul State, Brazil

<sup>3</sup> Programa de Pós-Graduação em Biologia Animal - PPG-BAN, Universidade Federal do Rio Grande Sul, UFRGS, Av. Bento Gonçalves 9500, Porto Alegre, 91540-000, Brazil

<https://zoobank.org/49F18A20-78B9-430D-858B-15FFDC1B7E51>

Corresponding author: José R. I. Ribeiro ([joseribeiro@unipampa.edu.br](mailto:joseribeiro@unipampa.edu.br))

Received 02 July 2022  
Accepted 04 December 2022  
Published 16 January 2023

Academic Editors Bruno Clarkson, Mónica M. Solórzano-Kraemer

**Citation:** Ribeiro JRI, Ferrari A (2023) Phylogenetic analysis of the *Belostoma plebejum* group *sensu* Nieser (Insecta, Hemiptera, Belostomatidae): the effect of adding continuous characters on its accuracy. Arthropod Systematics & Phylogeny 81: 1–34. <https://doi.org/10.3897/asp.81.e87378>

## Abstract

The *Belostoma plebejum* group comprises nine species, and the most evident characteristic shared by all species of the group is a phallus that is strongly curved ventrally. The difficulty in studying its species is much aggravated by the scarcity of identified material in Brazilian collections, and this has negatively impacted phylogenetic studies within the group. We tested the monophyly of the *B. plebejum* group using discrete and continuous characters under different weighting schemes and inferences. We described *B. lanemeloï* **sp. nov.** and *B. nieseri* **sp. nov.** and they served as the basis to study the phylogenetic relationships. A strict-consensus tree recovered under maximum parsimony and with implicit weighting scheme is as follows: (*B. parvum*, ((*B. lanemeloï* **sp. nov.**, (*B. nessimiani*, *B. nieseri* **sp. nov.**)), (*B. micantulum* var1, (*B. micantulum* var2, (*B. estevezae*, ((*B. plebejum*, (*B. minusculum* var1, *B. minusculum* var2))), (*B. nicaeum* var1, *B. nicaeum* var2), ((*B. lariversi* var1, *B. lariversi* var2), (*B. pygmeum* var1, *B. pygmeum* var2)))))). The monophyly of the *B. plebejum* group is corroborated by four non-homoplastic synapomorphies, and the aforementioned condition of the phallus is one of them. We tested the phylogenetic integrity of some species of the *B. plebejum* group, and only the exemplars of *B. micantulum* did not constitute monophyletic clades. Comparing the topologies obtained by different approaches clearly showed the presence of different scenarios in terms of heterogeneity of evolutionary rates among characters, but this could also be influenced by the disproportionate number of discrete characters compared with continuous characters.

## Key words

aquatic insect, combinations of measures, combined analysis, giant water bug, male genitalia.

## 1. Introduction

Belostomatidae currently comprises 11 genera and approximately 150 species (Ribeiro et al. 2018). The New World comprises a large part of its diversity, with about two thirds of the total number of records of this fauna

(Polhemus and Polhemus 2008). The genus *Belostoma* Latreille, 1807 consists of 74 species distributed in the Neotropical Region, of which around 40 have been recorded in Brazil (Moreira et al. 2011; Ribeiro et al. 2018;

Stefanello 2021; Stefanello et al. 2021). Considering the existence of a few major available revisions on this genus, some old or even incomplete ones have represented the unique material available for the identification of these species, which has contributed to the still incipient establishment of the systematics of its species over the years. In David Lauck's seminal paper (Lauck 1962, 1963, 1964), directed to the genus *Belostoma*, the descriptions of the treated species were based on the morphological study of dorsal projections of the phallotheca ('dorsal arms of phallosoma' from the authors)—the arms of the phallotheca—as well as a ventrally positioned diverticulum ('ventral diverticulum' from the authors), both more sclerotised portions of the male genitalia in Belostomatidae (Dupuis 1955; Ribeiro et al. 2018).

Taking into account the presence of a very subtle variation found across the external morphology and male genitalia of the *Belostoma* species, coupled with the scarcity of taxonomists, the difficulty in discriminating individuals of the same species has made the elucidation of biological units at the species level and the taxonomy of the genus onerous and unsolved for many years. Identification keys to some groups of species can be difficult for beginners to understand and use (e.g., Ribeiro and Alecrim 2008; Ribeiro et al. 2017). However, the study of these genitalic structures and their use in the taxonomy of the genus has, to some extent, contributed to the more precise identification and description of species—especially those considered morphologically similar to each other—despite some successful attempts to establish a taxonomy without mentioning the aforementioned genitalic characteristics (e.g., Estévez and Polhemus 2007; Stefanello 2021). Recent revisions, such as Nieser and Melo (1997), Estévez and Polhemus (2001, 2007), Ribeiro and Estévez (2009), Ribeiro et al. (2017), and Stefanello (2021), have therefore apparently been solving some of the previous problems found through the review papers mentioned above.

On the other hand, male genitalia have become a limiting factor for the determination of *Belostoma* species, as some of them cannot be identified without its observation (e.g., Ribeiro and Estévez 2009; Ribeiro et al. 2017). For example, based on an assessment of the general morphological similarity between these taxa and the morphology of the male genitalia, Lauck (1962) organised *Belostoma* into 16 species groups. He considered the differences between the shapes of the arms of the phallotheca and the diverticulum as important criteria for separating 10 of the 16 established species groups. He also recognised that some species could comprise groups based on the body size of their members (Estévez and Polhemus 2001, 2007). In fact, untreated groups comprised those whose species were represented by small individuals (i.e., the *B. denticolle* Lauck, 1962, *B. oxyurum* Lauck, 1962, *B. plebejum* Lauck, 1962, and *B. pygmeum* Lauck, 1962 groups), possibly due to the author's difficulty in listing those characteristics used for differentiating groups, given the morphological proximity among their members.

When established by Lauck (1962), the *B. plebejum* group comprised only three species: *B. minusculum*

(Uhler, 1884) occurring from Honduras to Venezuela; *B. micantulum* (Stål, 1860), species with the largest distribution area in South America, occurring from Venezuela to Argentina (Estévez and Polhemus 2007); and *B. plebejum* (Stål, 1860), recorded in southern Brazil (Paraná), Uruguay, and Paraguay, extending to Bolivia and Peru (Table S1). Lauck (1962) defined the group based on the existence of a ventroapical protuberance on the ventral diverticulum of the phallus, a condition that was otherwise considered doubtful at that time. The author never presented detailed descriptions of these species in his works, although he had included the group in his proposed identification key to the groups of *Belostoma* species. In addition, based on the conspicuous thickening of the dorsolateral margins of the ventral diverticulum, Lauck (1962) defined the *B. pygmeum* group, which included *B. lariversi* De Carlo, 1960 and *B. pygmeum* (Dufour, 1863), both recorded in southern Brazil (Paraná), Uruguay, and Paraguay, extending to Bolivia and Peru. Despite having established its members as different from those of the *B. plebejum* group, he did not work thoroughly on the taxonomy of these species, although he had already included this group in the same identification key to the species groups he had proposed.

The *B. plebejum* group, as conceived by Nieser (1975), now comprises the members of the two groups mentioned in the previous paragraphs (Table S1). At that time, however, nine species were known, occurring from Central America (Honduras) to southern South America (Argentina) (Estévez and Polhemus 2007; Ribeiro and Alecrim 2008). According to Nieser (1975), the conspicuous condition of the ventral diverticulum in male specimens of two species of the *B. pygmeum* group would represent an extreme case of specialisation of the representatives of this group. Later, Estévez and Polhemus (2007) described and added, in agreement with Nieser's study (1975), *B. nicaeum* Estévez and Polhemus, 2007 and *B. parvum* Estévez and Polhemus, 2007 to the group (Table S1), species restricted to northern South America. Furthermore, Nieser (1975) commented that the rounded and slightly prominent aspect of the prosternal carina is a characteristic shared by all species of that group. However, the most evident characteristic is the condition of the phallus, which is strongly curved ventrally (Nieser 1975; Ribeiro 2007), differentiating all members of the *B. plebejum* group from the other species of *Belostoma* that are represented by small individuals.

The aforementioned difficulty in studying the species of the *B. plebejum* group is much aggravated by the scarcity of identified material in Brazilian collections, which is, in turn, an outcome of the impossibility of identifying females of certain species. The main problem of females is the absence of important characteristics like male genitalia (Ribeiro 2007). This has negatively impacted phylogenetic studies within the group, becoming the understanding of the phylogenetic relationship among species of the *B. plebejum* group necessary. Besides, testing its monophyly—as suggested by Nieser (1975) and Estévez and Polhemus (2007)—is also important, especially if these efforts are linked to a better understanding

of the morphological variation, focussing on the identification of characteristics that bear phylogenetic signal and that can substantially contribute for the taxonomy of the group. Since the size of character matrices for the construction of phylogenetic relationships directly and proportionally influences the accuracy of these relationships (Parins-Fukuchi 2018), it is particularly welcome to include measures and metric relations that could supposedly increase the accuracy in the definition of these species. Intraspecific variations observed during studies of phylogeny are often, by tradition among systematists, ignored (Pollock et al. 1998). Some systematists even sample a few individuals of each species in order to minimise the number of different states of a multistate character (Wiens 1999). This neglect, however, corresponds to a potential loss of information (Amorim 1997; Escapa and Catalano 2013; Mounce et al. 2016) because errors in the reconstructions inferred by this type of information added to matrices of discrete characters tend to be generally smaller, especially in scenarios with high evolutionary rates (Parins-Fukuchi 2018). If we still consider a homology hypothesis as a particular distribution of frequency values in a sample of taxon specimens (Thiele 1993), we can assume that this particular distribution or trend is inherited from its ancestor (Rae 1998), something usually neglected.

The usual tradition of excluding characters that show more explicit intraspecific variation has generally been justified by the presence of some degree of overlap of their values among the studied taxa (Thiele 1993; Wiens 1999). These characters have been called ‘continuous’ or ‘quantitative’ (Wiley 1981), as well as ‘polygenic’ (Felsenstein 2004) and even ‘continuous and quantitative’ (Lawing et al. 2008), and many phylogeneticists still treat them as ‘controversial’ (Hauser and Preschi 1991) or ‘non-cladistic’ (Pimentel and Riggins 1987). In fact, one could consider that characters supposedly without overlapping values among the taxa—the so-called ‘discrete’ characters—are characters whose frequency distributions (or their character states) may not have been well sampled, characterising an error of sampling (Thiele 1993). Measures, counts, and metric relationships are even better utilised when treated as such, rather than being converted into discrete, non-overlapping entities (Farris 1990; Goloboff et al. 2006), as well as when treated as additive or ‘minimally connected’ in the terminology of Slowinski (1993). Using metric characters as such brings greater objectivity (Parins-Fukuchi 2018) and reduces the possibility of joining different discrete states to terminals that are not different (Farris 1990), as in the case of transforming them into discrete entities (e.g., Thiele 1993). The use of minimally connected characters reinforces the fact that a metric or meristic state is intermediate to the others and conforms itself to the idea of ‘morphological intermediacy’ (Gutberlet 1998). Furthermore, the intervals between measurements can be considered polymorphisms (Goloboff et al. 2006).

Character matrices subjected to maximum parsimony (MP) analysis are supposed to assess organismic disparity well, even if different body portions (or modules) of

the studied individuals of each species with different homoplasy levels are used and grouped in the same matrix (Gatesy et al. 1999; Gatesy and Arctander 2000), or if different coding strategies are used to create these matrices (Hetherington et al. 2015). If different parts of the body have different homoplasy levels, the incongruence between the partitions associated with these different body parts will favour particular patterns of convergence and homoplasy, therefore producing different sets of topologies (Clarke and Middleton 2008). Thus, the use of measures as continuous series not divided into an arbitrary set of states, together with the use of discrete morphological characters (non-overlapping frequency distributions), is a more appropriate approach because it will generally not impact the actual distances between the taxonomic units used, implying the existence of morphological integration between these different body parts (Hetherington et al. 2015). Furthermore, this procedure allows for establishing of a stronger phylogenetic signal, which overrides the local conflicts caused by existing homoplasies (Gatesy et al. 1999). Therefore, exploring these behaviours, by adopting different implicit weighing schemes for characters and different quantities and combinations of measures during the tree search analyses (such as MP) intends to investigate the effect of these practices on the accuracy and stability of the topologies obtained (Goloboff et al. 2008). Besides, if we compare the results of these analyses under MP to those obtained with explicitly statistical methods (maximum likelihood [ML] and Bayesian inference [BI]), considering the resolution, disparity, and support, the topologies obtained by different approaches may indicate the existence or not of different scenarios, in terms of heterogeneity of evolutionary rates. In those cases, a certain disparity between the results obtained with different reconstruction modes can be evident in a scenario with different evolutionary rates among the proposed characters, even with the use of discrete characters only (Wright and Hillis 2014). The combination of different approaches, therefore, could indicate whether the evolution of morphological characters is somehow under the influence of the long-branch attraction phenomenon (e.g., Wiens and Hollingsworth 2000; Magalhães and Santos 2012; Ribeiro et al. 2018). Finally, topologies obtained under such parametric methods provide information not only on relationships but also on distances between species, the information provided by estimating branch lengths (Yang 2014).

In this study, we have conducted a phylogenetic analysis of the species of the *B. plebejum* group, according to the most recent and updated group configuration found in Ribeiro and Alecrim (2008). During the survey of specimens for that purpose, as well as of characters, we have described two new species occurring in Brazil. They have served as the basis to study the phylogenetic relationships among the species of *B. plebejum* group. We have included taxonomic notes and have provided an identification key to the species. Furthermore, we have submitted some specimens, for which we observed intraspecific variations, to the phylogenetic integrity test, under MP, according to the procedures described by de Pinna (1999). More-

over, based on Goloboff et al. (2006), we have developed an evaluation procedure in order to verify the effect of progressively adding measures and metric relationships to discrete character matrices of MP analyses. We have also included parametric topology searching methods to better assess the effects of homoplasy accumulation in this group of cryptic species (in terms of external appearance) because a matrix of a few characters was constructed (see below). Furthermore, we have reconstructed the evolution of characters obtained from the study of male genitalia, whose conditions were mutually exclusive and exclusively comparable with other conditions of the male genitalia characters (comparable characters) considered to be important in the taxonomy of the group, after optimisation over the trees obtained under MP and BI with the matrix combining all these characters, according to the logic of ‘total evidence’ (de Queiroz et al. 1995). In addition, we have studied the pattern of diversification of these characteristics throughout the obtained trees and have determined their homoplasy levels. By optimising these characters via ML in the topology obtained under BI, we also explore the probability that the reconstructed ancestral states (via MP) are true, due to the possible discrepancy between phylogenetic relationship and distance between species.

## 2. Methods

### 2.1. Specimens

We studied about 130 adult specimens for this work. We evaluated 18 species and included 30 terminal taxa in the analyses (see the section on the phylogenetic integrity test), with 11 species comprising the ingroup, including the two new species; all species are listed in Table S2. We used the following species as outgroups: *B. amazonum* Estévez and Polhemus, 2001, *B. denticolle* Montandon, 1903, from *B. denticolle* group Lauck, 1962; *B. minor* (Pasilot de Beauvois, 1805), from *B. minor* group Lauck, 1962; *B. candidulum* Montandon, 1903, *B. horvathi* Montandon, 1903, *B. oxyurum* (Dufour, 1863) and *B. sanctulum* Montandon, 1903, from *B. oxyurum* group Lauck, 1962. We chose the outgroups based on the criteria of Nixon and Carpenter (1993), more precisely when taking into account the existence of some level of morphological proximity as suggested by Lauck (1962), Nieser (1975), Estévez and Polhemus (2001), Estévez and Polhemus (2007), and Ribeiro and Estévez (2009).

The material studied comes from the collections of the American Museum of Natural History, New York, United States of America (AMNH); Departamento de Parasitologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Minas Gerais, Brazil (DPIC); Departamento de Zoologia, Instituto de Biologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil (DZRJ); the Hungarian Natural History Museum, Budapest, Hungary (HNHM); Institut Royal des Sciences

Naturelles de Belgique, Brussels, Belgium (ISNB); Instituto Nacional de Pesquisas da Amazônia, Manaus, Brazil (INPA); Museo Argentino de Ciencias Naturales, Buenos Aires, Argentina (MACN); Museu Nacional, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil (MNRJ); Museum National d’Histoire Naturelle, Paris, France (MNHN); Museu de Zoologia da Universidade de São Paulo, São Paulo, Brazil (MZSP); and the Snow Entomological Museum, University of Kansas, Lawrence, United States of America (SEMC). All acronyms of depositories are based on Arnett-Jr. et al. (1993), except DZRJ and MNRJ. The list of examined material of the two new species is associated with the formal description of each one presented in the ‘Systematics’ section of our results.

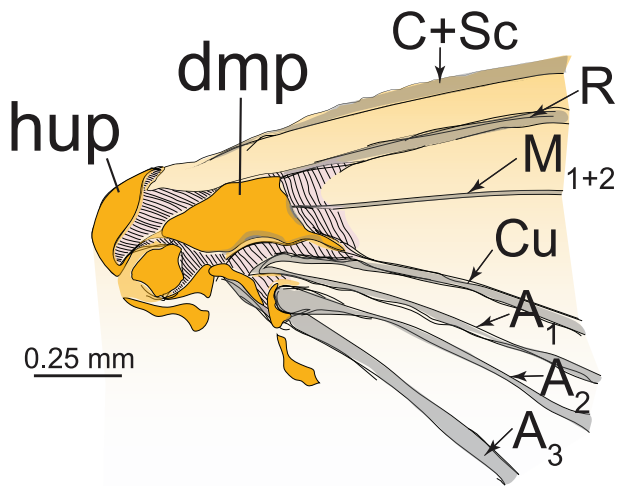
### 2.2. Measurements, illustrations, terminology, and techniques

To obtain the total body length and the greatest width of the body of adults, we used a manual calliper with an accuracy of 0.1 mm. We took all other measurements with the aid of a 0.01 mm accuracy millimeter ocular micrometer on a stereoscopic microscope. The measures taken are based on the measurement according to Ribeiro (2007) and Ribeiro and Alecrim (2008). We used all available materials of each species in this study to obtain the measurements. The referred measurements and relations are summarised as mean values in Table S3, and are as follows: **TBL**, total body length (from the apex of the head to the end of the abdomen); **T/G**, ratio between the total body length and the greatest width of body; **R1/R2**, ratio between the lengths of first and second visible segments of the rostrum; **A/I**, ratio between the median lengths of the anteoculus and interoculus; **IW/A**, ratio between the posterior interocular width and the median length of anteoculus; **IW/O**, ratio between the posterior interocular width and the width of an eye; **W/L**, ratio between the width and median length of the ventral diverticulum of the phallosoma.

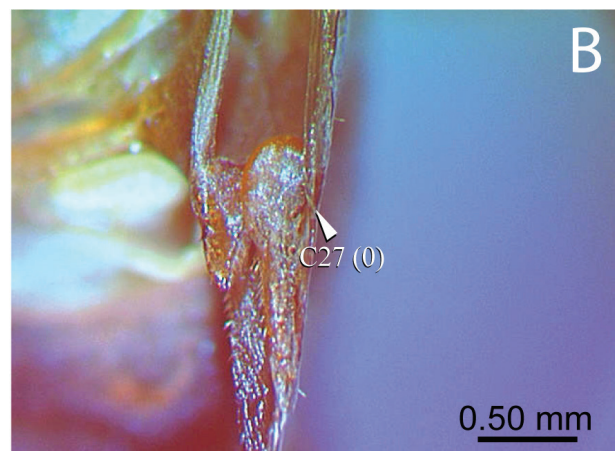
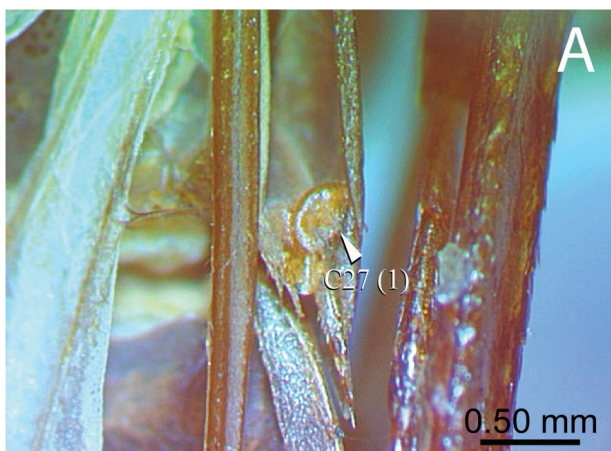
We prepared all line drawings and sketches with a drawing tube mounted on a ZEISS (AXIOLAB) optic microscope and a ZEISS (SV-6) stereo microscope. We took photographs with a ZEISS stereo microscope with a MC 80 camera (with AXIOVISION 3.0 SP4 software), housed at Laboratório de Captação de Imagens do Instituto de Biologia, Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil.

The terminology of the head, thorax, and abdomen follows Hamilton (1981), Lauck (1962), Nieser (1975), Parsons (1964, 1974), Estévez and Polhemus (2007), Ribeiro (2007), Ribeiro and Alecrim (2008), and Ribeiro and Estévez (2009) for specific areas, sclerites, and sutures of the head, thorax, and abdomen; Betts (1986) for basal portions of fore and hind wings; Gorb and Perez-Goodwyn (2003) for the *Druckknopf* system, which is responsible for coupling of the hemelytra and the body; and Estévez and Polhemus (2007) and Ribeiro and Estévez (2009) for male genitalia.





**Figure 1.** Basal portion of the right hind wing and its axillary sclerites. Six characters were obtained based on the study of sclerite 1 of the distal median plate (characters 21–26), which was the only one to show useful variation through the studied specimens. By using these sclerites and veins as ‘markers’, we established all other units of comparison. These veins (and their branches) are named according to a system devised by John Comstock and George Needham: the Comstock–Needham system. Abbreviations: dmp, sclerite 1 of the distal median plate; hup, humeral plate. Veins: A<sub>1</sub>, anal 1; A<sub>2</sub>, anal 2; A<sub>3</sub>, anal 3; C + Sc, costa + subcosta; Cu, cubitus; M<sub>1+2</sub>, media 1 + media 2; R, radius.



**Figure 2.** Mesepimeron and part of wing-to-body coupling mechanism. Part of the ‘Druckknopf’ system component, located in the thorax. The system allows for the complete coupling of hemelytra to the dorsum. Character 27 was proposed by studying only the complementary structure of the apparatus, located in mesepimeron, not that on the wing apparatus. **A** Knobshaped ‘Druckknopf’ system rounded, as long as wide **B** knobshaped ‘Druckknopf’ system piriform, longer than wide.

More specifically, the general terminology of the axillary sclerites used in this study follows Betts (1986) and Yoshizawa and Saigusa (2001). According to Betts (1986), there is a reduction in the number of axillary sclerites through Heteroptera, when compared with the possible ancestral state found in Homoptera *s.l.* However, the location and number of described sclerites remain constant through the representatives of Heteroptera, except for those of the genus *Notonecta* Linnaeus, 1758 (Notonectidae). Due to their easy visualisation and handling, we preferred to use the sclerites of the basal portion of the hind wing of the specimens studied after having certified that both wings have the same pattern regarding the distribution and shape of the sclerites. This verification agrees with the pattern found in the arrangement and shape of the axillary sclerites, as reported by Betts (1986) and Yoshizawa and Saigusa (2001). It could be possible to obtain hypotheses of homology (‘primary homologies’ *sensu* de Pinna 1991) for the sclerites found in the specimens studied, identifying the humeral plate; cubitus; and anal 1, 2, and 3 veins (hup, Cu, A<sub>1</sub>, A<sub>2</sub>, and A<sub>3</sub>, respectively, Fig. 1). By using these sclerites and veins as ‘markers’, we were able to establish all other units of compar-

ison. In this study, we obtained six characters (Table S4, characters 21–26) based on the study of sclerite 1 of the distal median plate (‘DMP1’ for Yoshizawa and Saigusa 2001; dmp, Fig. 1), which was the only one that showed useful variation through the studied specimens.

The study of the *Druckknopf* system components, according to Weber (1930) and Schuh and Slater (1995), or the ‘knob-and-socket’ system, according to Gorb and Perez-Goodwyn (2003), which allows for the complete coupling of hemelytra to the dorsum, follows A. L. Estévez’s ideas (*pers. comm.*) regarding the purpose of homologies and coding characters. The system is known to occur in several families of aquatic Heteroptera, and the following authors have described it in detail: Cobben (1957) for Saldidae Amyot and Serville, 1843; Gorb and Perez-Goodwyn (2003) for Corixidae Leach, 1815, Naucoridae Fallén, 1814, Nepidae and Notonectidae Leach, 1815; Larsén (1945) for Nepidae; and Perez-Goodwyn and Gorb (2003) for Belostomatidae. To study the knob-and-socket apparatus, it was necessary to carefully displace the hemelytra, by pulling the anterior portion along their lateral margins near to scutellum after moistening the insertion point of the hemelytra on the thorax. Con-

cerning the procedure, it neither damaged the forewing nor completely decoupled the forewings from the body. Moreover, the involved musculature kept itself intact, allowing us to return the hemelytra to its initial position after the procedure. We have proposed characters by studying only the complementary structure of the apparatus, located in the thorax, and not that on the wing apparatus (Fig. 2).

The procedures for the preparation of male genitalia follow the same ones described by Ribeiro (2007); we adapted them for small insects. After the examination, we stored the dissected parts in microvials with glycerine, properly sealed, and then kept together with the voucher specimens.

### 2.3. Character coding and phylogenetic integrity

Our proposition of homologies is based on Brower and Schawaroch (1996), in which the hypothesis of primary homology (de Pinna 1991) is an outcome of topographical identity between structures. Furthermore, based on Thiele's (1993) ideas, we considered putative homologies from measurements and metric relations to be a particular distribution of frequency values in a sample of specimens of a taxon in this study. Such a distribution (or particular trend) is considered to be inheritable from its ancestor (Rae 1998). Forty-nine characters used in the phylogenetic analyses are listed in supplementary file 3, seven of which are continuous (Table S3, Table S4). We grouped the characters into the following categories: body size (1–2), head (3–8), thorax (9–27), abdomen (28–49), and male genitalia (30–49). Character coding was contingent for discrete data (Forey and Kitching 2000) and the character statements were formulated following the logical basis of Sereno (2007). We scored character states as dashes (-) if inapplicable and as question marks (?) if ambiguous or missing. We constructed the matrix of discrete characters in Mesquite v.3.61 (Maddison and Maddison 2019) and reconstructed character optimisations in MacClade (Maddison and Maddison 2005) using trees based on parsimony analyses with discrete and continuous characters combined. Only unambiguously optimised characters are discussed in a more detailed way through the text. We performed character optimisations on a morphological strict consensus tree from MP analyses with implicit weightings assigned to characters (analysis six, see below). Apart from MP optimisations, we optimised comparable characters based on male genitalia with Bayesian maximum credibility trees. When any character optimisation on trees was different, we have discussed such conflicts. For each character, we calculated the number of steps and the consistency (CI) and retention (RI) indices (Kluge and Farris 1969). Of the 42 discrete characters, 13 are multistate characters and have been coded as unordered (or 'maximally connected' in the terminology of Slowinski 1993) (Fitch 1971). Otherwise, we coded continuous characters as ordered (Farris 1970) and treated them as such, according

to Goloboff et al. (2006). To be used as such, however, we ln-transformed the mean values of measurements and metric relations to reduce the supposed overdominance that they may present during implicit weighing analyses under MP (Koch et al. 2014).

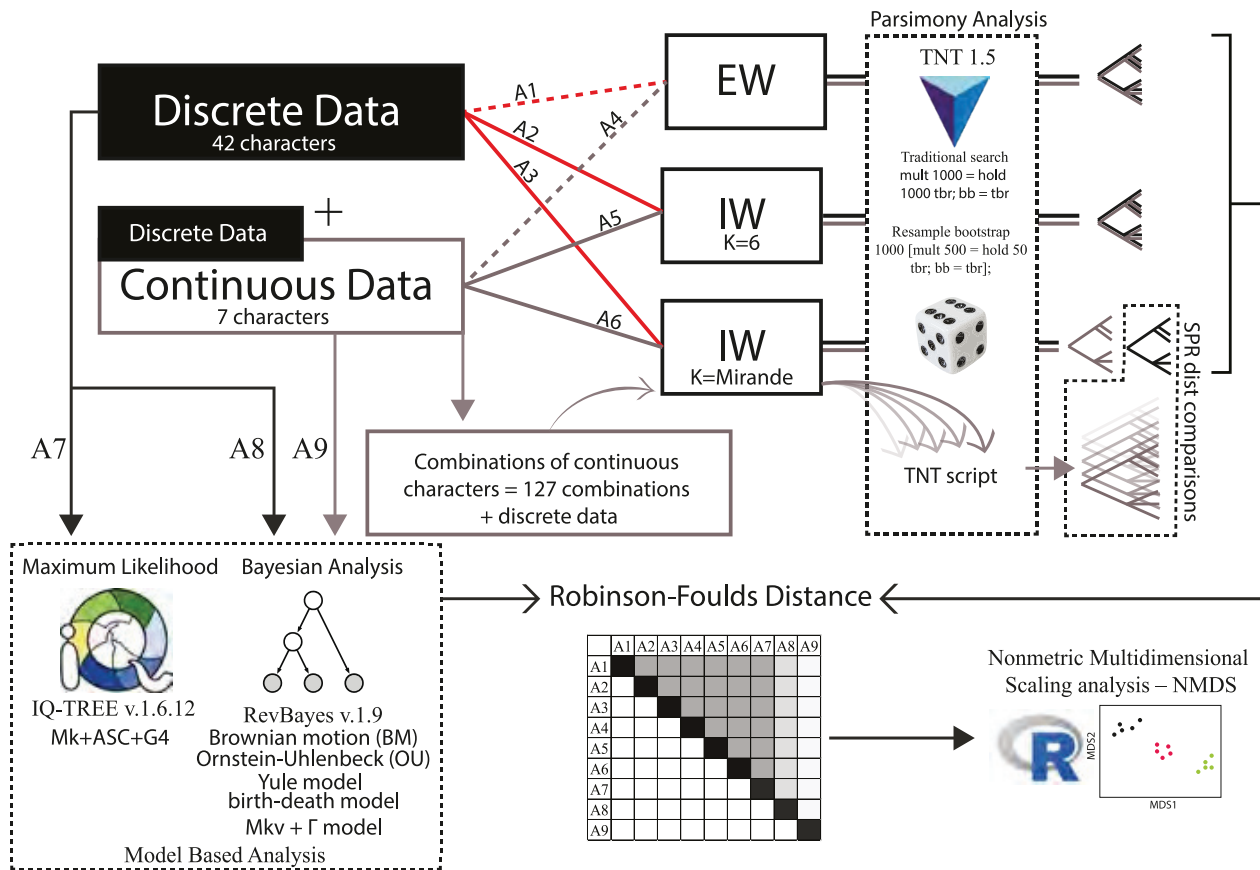
We treated specimens with intraspecific variations ('exemplars', in the terminology of de Pinna 1999) as terminal taxa, receiving the word 'var' followed by a number, and we thus considered intraspecific variations as distinct character states. This procedure allowed us to test the phylogenetic integrity (de Pinna 1999) of some putative species represented by individual organisms. Following the protocol described by de Pinna (1999), we corroborated the phylogenetic integrity of a species containing individual organisms with intraspecific variation if individuals did not themselves constitute para- or polyphyletic groups or even monophyletic groups with subclades in a cladogram. The following taxa had specimens with the intraspecific variation being evaluated: *B. amazonum*, *B. denticolle* (with three variations), *B. minor*, *B. candidulum*, *B. oxyurum*, and *B. sanctulum* (outgroups); *B. lariversi*, *B. micantulum*, *B. minusculum*, *B. nicaeum*, and *B. pygmeum* (ingroup). We have only discussed variation found among representatives of the ingroup. Especially regarding *B. micantulum*, studied specimens may apparently constitute distinct populations, and this seems to be the case here because of the existence of variation in the shape of their arms of phallotheca. Besides, many authors had cited this species as actually a species complex (see Stål 1860; Nieser 1975; Ribeiro 2007;). In this study, we partly adapted the original method by using in some cases groups of individuals rather than only one specimen per putative species. The species treated in this way were *B. lariversi*, *B. micantulum*, *B. minusculum*, *B. nicaeum*, and *B. pygmeum*. If we could refute the phylogenetic integrity of some species-candidate assemblage, then we mention the species name inside quotation marks throughout the text.

In *B. candidulum* var2, the character referring to the apex of the prosternal carina (character 12) could not be coded, as the apex of carina was damaged. In *B. sanctulum* var2, male genitalia characters (30–49) could not be coded, as unfortunately we only studied females (see below).

## 2.4. Phylogenetic analyses

### 2.4.1. Maximum parsimony analyses

First, we carried out a comparative study of the characters by using the cladistic method (Brower and Schuh 2021). We performed character polarisation *a posteriori*, using the outgroup comparison method (Nixon and Carpenter 1993). We performed MP phylogenetic analyses in TNT v.1.5 (Goloboff and Catalano 2016) under two weighting schemes: equal weights (EW) and implied weights (IW) (Goloboff 1993). We conducted groups of analyses on either the character matrix considering only the discrete characters (DISC) or on a combined morphologi-



**Figure 3.** Scheme of all analyses performed with different matrices, combinations of continuous characters and inferences. First, maximum parsimony analyses were performed in TNT v.1.5, under two weighting schemes: equal weights (dotted lines) and implied weights (solid lines). Groups of analyses were conducted on either the character matrix considering only the 42 discrete characters (red lines) or on the combined dataset, including seven continuous and 42 discrete characters (grey lines). Three schemes combining character matrices with different weights were applied, totalling six analyses: (1) the matrices of analyses one (A1) and four (A4) were assigned equal weights; (2) the same matrices, now in the analyses two (A2) and five (A5), respectively, were submitted to an implicit weighting scheme, with constant  $K$  equal to six; (3) the same matrices, now in the analyses three (A3) and six (A6), respectively, were submitted to implicit weightings, but with a constant  $K$  estimated after exploring different concavities. Second, a total of 127 combinations were added gradually to the matrix of discrete characters submitted to A3, because the estimated  $K$ -value was quite close to that of A6. The Subtree Pruning Regrafting distance (SPR dist) and parsimony character bootstrap were used in a way to account for differences, similarities, and special behaviours among topologies (and nodes) found in each iteration of the procedure and the topology (or consensus topology) found in A3. Apart from maximum parsimony analyses, in analysis seven (A7), maximum likelihood inference was conducted for the matrix of only discrete characters using W-IQ-TREE v.1.6.12, with default parameters (see the text for more details). For Bayesian inferences, analysis eight (A8) was performed on the matrix with only discrete characters, while the analysis nine (A9) used the matrix with all datasets combined. It was partitioned into discrete and continuous subsets, and the appropriate evolutionary models were applied to each partition separately. The partitioned matrices were used in REVBayes v.1.9, and six evolutionary scenarios were established, given the presence of continuous data and its phenotypic variance (see the text for more details). The Robinson–Foulds symmetric absolute distance was used to compare the topologies obtained. Accordingly, a representation of differences and similarities between the topologies found under different approaches was summarised in a nonmetric multidimensional scaling analysis biplot, after calculating a triangular Robinson–Foulds distance matrix.

cal dataset, that is, including continuous ones (DISC+CONT) (Fig. 3). In each group of analyses, we applied three schemes combining character matrices with different weights, totalling six analyses: for the first approach (EW analyses), we assigned equal weights to the matrices of analyses one (DISC) and four (DISC+CONT). For the second approach (IW analyses), we submitted the same matrices, now in analyses two (DISC) and five (DISC+CONT), to an IW scheme with a constant  $K$  equal to six, downweighting the fit of homoplastic characters in the

obtained trees more slowly. For the third approach, we submitted the same matrices, now in the analyses three (DISC) and six (DISC+CONT) to IW, but with a constant  $K$  estimated after exploring different concavities based on the method developed by Mirande (2009) (Table 1).

The constant  $K$  determines the weight to which characters will be submitted according to the existing homoplasies in a given topology (Goloboff 1993). Because there is no objective criterion for choosing an ideal  $K$  value, we decided after exploring different concavities; this



**Table 1.** Summary of the results of weighing combinations, totaling six analyzes of maximum parsimony inference.

Analysis	Dataset	Weighting scheme	Number of trees	Number of steps	CI	RI	Fit	$K_{refM}$ (K)	K ranging
A1	DISC	EW	53	130	0.45	0.75	—	—	—
A2	DISC	IW6	1	132	0.45	0.75	34.27	6	—
A3	DISC	IW	16*	132	0.43	0.74	15.94	2.2 (3)	1.2–3.7
A4	DISC+CONT	EW	1	139.9	0.44	0.74	—	—	—
A5	DISC+CONT	IW6	1	142.7	0.44	0.74	16.80	6	—
A6	DISC+CONT	IW	4*	142.3	0.44	0.74	14.61	2.8 (10)	2.5–3.2

\*Total number of trees recovered by  $K$ -values ( $K$  ranging) with the smallest SPR distance.

**Abbreviations:** CI, consistency index; CONT, continuous characters; DISC, discrete characters; EW, equal weight; IW, implicit weighting with  $K$ -values defined by the Mirande's (2009) method; IW6, implicit weighting ( $K = 6$ ); RI, retention index.

decision depended on the type of data coded into character matrices (Mirande 2009). We employed Mirande's (2009) routine with 25 distortion values, adjusting the minimum and maximum distortion degree thresholds between 50% and 90% respectively. We further carried out the whole routine following the procedures described in a more detailed way by Weiler et al. (2016). To obtain a better value of  $K$  separately for analyses three and six, each with IW, we compared all topologies resulting from the procedure with IW to each other. If we obtained more than one topology in each of the suggested  $K$ -values, we used strict consensus trees instead (Tables S5 and S6). Then, we compared the most parsimonious topologies to each other using the Subtree Pruning Regrafting (SPR) distance (Goloboff 2007), which calculates the minimum number of SPR moves required to transform one tree into another (Goloboff et al. 2008; Kuhner and Yamato 2015). We defined the best  $K$ -value as that value (interval of  $K$ -values) of the topology with the smallest SPR distance among all trees produced by an interval of  $K$ -values defined by the method of Mirande (2009). The calculation of the SPR distance for each of the trees produced is usually laborious and is assumed to produce typing errors because the outcome of the routine usually is converted into triangular matrices, which are supposed to be built in software such as Excel. We thus developed a routine for the automatic calculation of these distances in TNT: it is independent of the chosen number of  $K$  values. It is available in <https://github.com/jozecaricardo/TNT-routines> (spr.run file).

Second, to evaluate the congruence between characters based on measures, metric relations, and discrete characters, we established a procedure that progressively compares the effect of adding measures and their impact on topologies and support measures resulting from different resulting matrices. This procedure is based on all possible combinations of characters based on measures and relations (a total of 127 combinations), which we added gradually to the matrix of discrete characters submitted to analysis three, with  $K$ -values defined by the method of Mirande (2009) (Fig. 3). Because the estimated  $K$ -value in analysis three (DISC) was quite close to that of analysis six (DISC+CONT), we used the  $K$ -value from the former throughout the whole procedure. Similarly, we used the SPR distance—a measure of access to the general similarity between different topologies—and parsimony character bootstrap (Felsenstein 1985)—a way to access

the node stability (Barão et al. 2020)—to account for differences, similarities, and special behaviours among those topologies (and nodes) found in each iteration of the procedure and the topology (or consensus topology) found in analysis three. We developed the routines used for the procedure for TNT and they are available <https://github.com/jozecaricardo/TNT-routines> (tree\_comb.txt file). Investigating the effect produced by the addition of measures in the matrix of discrete characters herein constructed also allowed us to confirm or reject whether or not similar topologies would be obtained, observing the existence or not of repeated relationships between the taxa (following the ideas of Kitching et al. 1998).

We performed heuristic tree searches using the 'tree-bisection-reconnection' algorithm, with 1,000 replications (RAS), keeping 1,000 trees at each step of the search (mult 1000 = hold 1000 tbr; bb = tbr). Cladograms obtained in each analysis were given the relative reliability of each node with standard bootstrapping (Felsenstein 1985)—the outcome of resampling with substitution (with bootstrap option) (BS)—by using 1,000 pseudoreplicates. We used the same parameters previously used in TNT under EW when resampling (resample boot frequency replications 1000 [mult 500 = hold 50 tbr; bb = tbr]). Conversely, under IW we applied a Poisson correction due to the known effect of IW in resampling (Goloboff et al. 2003) (resample poisson replications 1000 [mult 1000 = hold 100 tbr; bb = tbr]). In addition, we estimated relative Bremer decay supports (BRel) (Goloboff and Farris 2002), which measure relative fit differences between optimal and suboptimal trees—with supports being the number of additional steps required for the disappearance of a corresponding clade (Bremer 1994)—by producing suboptimal trees with up to 15 extra steps and the proportion of supporting values to a clade (relative fit differences) equals 0.9 (subopt 15 × 0.9). BRel shown throughout the results refer only to those obtained in analysis six. We included autapomorphies in the matrix and therefore considered them when calculating the consistency index of trees as recommended by Yeates (1992).

#### 2.4.2. Model-based methods

Apart from MP analyses—which assume short and similar branch lengths across lineages and may be thus inconsistent when dealing with groups of species with su-



perimposed homoplastic changes ('long-branch attraction phenomenon', see Bergsten 2005 for more details; Wright and Hillis 2014; Wright 2019)—we adopted model-based methods. Their use greatly improves tree searches when a morphological dataset comprises a large amount of homoplasy—which is supposed to exist in this group of *Belostoma* species (J.R.I. Ribeiro unpubl. data)—as well as when morphological character matrices are coded based on a small number of characters (Wright and Hillis 2014) because these methods are known to have a great ability to explain parallel or convergent evolution (Wright 2019). Besides, as long as changes are small between terminals—that is, when characters exhibit low evolution rates—and the rate of evolution among characters is also similar to each other, topologies obtained by MP, ML, and BI tend to be similar to each other (Vernygora et al. 2020).

We conducted ML analyses for the matrix of only discrete characters using W-IQ-TREE v.1.6.12 (Nguyen et al. 2015), with default parameters; i.e., 100 parsimony starting trees and a BIONJ tree, as well as stochastic and iterative Nearest-Neighbour-Interchange (NNI) rearrangements of candidate trees produced by the heuristic search algorithm (a greedy hill-climbing tree search). During the searches, we adopted a value of 0.5 as the proportion of starting trees to be disturbed (perturbation strength = 0.5) and a limit of 100 iterations for another run to start, if any greedy search for better ML trees failed (IQ-TREE stopping rule = 100). The model applied to the matrix was Mk+ASC+G4, which considers changing one character state for another as common as the opposite, and assumes equal state frequencies among character states (Lewis 2001). Since the occurrence of a character state can be much high in frequency through the matrix, more transitions in this character would be expected and not all characters would thus evolve at the same rate (Fitch and Margoliash 1967; Yang 1994). To consider such heterogeneity in the rate of character changes, we drew values of evolutionary rates of proposed characters from a gamma distribution with four discrete categories (rate categories) (*G4* parameter), with alpha and beta parameters—which are equal to each other because they are parameters of a symmetric distribution—estimated from the dataset. So, we implemented a mixed model called the among-character rate variation (ACRV) model, according to Wright (2019). Because the inclusion of characters in the matrix is considered to be uninformative—invariant characters in the terminology of Wright et al. (2016) (or autapomorphies)—there was a risk that branch lengths would artificially become too long (Lewis 2001; Leaché et al. 2015). Hence, it was necessary to correct this bias ('ascertainment bias') with the use of ASC flag.

For BI, we used the matrix with all datasets combined, including measurements and metric relations. However, we partitioned the dataset into discrete and applied continuous subsets and the appropriate evolutionary models to each partition separately. We used the partitioned matrices in REVBAYES v.1.9 (Höhna et al. 2016) and established six evolutionary scenarios, given the presence of continuous data and its phenotypic variance. Regardless of the process, we always assumed (1) a scheme where

there would have only speciation events (Yule model) (Yule 1925) and (2) a scheme where extinction events would also be considered (birth-death model). So, we adopted two scenarios taking into account Yule and birth-death models separately under a Brownian motion (BM) process (Felsenstein 1985) with only one regime and a constant rate of evolution (Hansen and Martins 1996). We used the same two schemes again and constructed the other two scenarios under a BM process, but now allowing for evolutionary rates of continuous characters to vary over time (BMM model) (Hansen and Martins 1996). Finally, we again assumed the same two schemes mentioned above, but now under the Ornstein–Uhlenbeck (OU) process (Hansen and Martins 1996), with a constant selection force ( $\alpha$ ).

We implemented BI analyses with Markov chain Monte Carlo (MCMC) algorithms, by adopting a more efficient search with the implementation of the Metropolis Coupled MCMC sampling strategy (MC<sup>3</sup>) (Geyer 1991), which generalises the MCMC procedure, as long as it uses multiple chains in parallel (Wheeler 2012). The MC<sup>3</sup> started after the production of a random tree but incorporated priors from either the Yule process or birth-death model of speciation during branch lengths estimation (see above). Subsequently, the implementation rearranged such trees, using NNI, SPR, or both algorithms.

The birth–death–sampling process (Kendall 1948; Rannala and Yang 1996) assumes the emergence of a lineage at some point in time and after a certain amount of time (a sojourn time), the lineage can either speciate or become extinct. To minimise the effects caused by the equal probability of sampling taxa, given the underestimates that occur during procedures of sampling diversification rates from posterior, we relaxed the hypothesis that each species of each supraspecific taxa has been included in the analysis by adopting a diversified sampling of taxa (samplingStrategy = 'diversified') (Höhna et al. 2011; Höhna 2014). Furthermore, we incorporated the information that all or not all species were sampled into analyses by using the  $\rho$  parameter (sampling probability), which is the ratio between the number of sampled species and described species. We adopted the birth–death process with the use of hyperpriors—that is, the speciation ( $\lambda$ ) and extinction ( $\mu$ ) rates—to evaluate the uncertainty in their estimates. Therefore, we reparametrised them because of (1) the usual correlation known to exist among these parameters, becoming difficult to estimate them from posterior, and (2) the remaining possibility of sampling higher values of extinction rate from posterior than those of the speciation rate. We drew values of the diversification rate from a lognormal distribution and values of turnover (number of species being replaced by new ones)—an alternative reparametrisation of the aforementioned hyperpriors—from a gamma distribution. Thus, we defined deterministic nodes in REVBAYES and acquired samples of both speciation (as diversification rate + turnover rate) and extinction (turnover) rates. Finally, we adopted similar procedures to the Yule process (a special case of the birth–death process), considered to be a model where only speciation occurs, so that the extinction rate is null

and all taxa have the same speciation rate (Yule 1925). Despite being unrealistic, this process seems to be more suitable to processes with many events of recent speciation, which seems to be the case. We implemented a pure-birth model, with the diversification rate as hyperprior and free parameter.

We assumed that continuous characteristics used here could evolve according to one of the aforementioned BM processes or according to OU process, a model with the presence of a constant selection force (Hansen and Martins 1996; Harmon 2019). To fit BM processes to the dataset, we assigned priors the evolutionary rate parameter  $\sigma^2$ , the rate of evolution or the scale of change in a trait through time, whose values we drew from a log-uniform distribution (Parins-Fukuchi 2018; Harmon 2019). Moreover, we considered the possibility of changing the rate of phenotypic variation through time—a dynamic BM process (BMM model)—by relaxing BM processes in some BI analyses (relaxed Brownian motion). We thus assumed that outgroups and the ingroup had been submitted to putatively different regimes over time, since some increment in genitalic morphological complexity seems to occur only within the *B. plebejum* group. For this purpose, we stipulated a total of four events of change following the general idea that different rates of change through time could be likely to arise between the ingroup and outgroups, scattered through some outgroups, and in the *B. pygmeum* group. Whatever the processes modelling the evolution of continuous characters, we adopted the restricted maximum likelihood (REML) approach (Garland 1992), producing less-biased rates compared with those estimated with traditional ML optimisations. REML methods are likely to dismiss parameters considered to be less important (nuisance parameters), such as  $\theta_0$  (the starting value of the population mean trait), as long they are usually estimated with much uncertainty; so, they are removed during the likelihood calculation (Harmon 2019). In REVBayes, we adopted a phylogenetic BM using residual ML process (with *dnPhyloBrownian-REML* function), that is, a BM process that uses phylogenetic contrasts (PIC) to integrate all unobserved states in internal nodes (Höhna et al. 2016). In addition, phylogenetic covariance matrices of measurements and metric relations herein used and their inversions are not required, something that also minimises the bias (Harmon 2019).

For the OU process, we fitted the model to the dataset and estimated  $\sigma^2$  by sampling values of selection force  $\alpha$ , the strength of attraction towards a peak, and the optimum trait value, the mean value  $\theta$  of a trait under selection, from exponential and uniform distributions, respectively. The OU model adopted in this study assumed the evolution of continuous characters to be under a constant selection force  $\alpha$  with the constant regime, a homogeneous OU process (Hansen and Martins 1996; Hansen 1997), with this strength being proportional to the difference between ancestral mean trait values and regime states under selection (Harmon 2019). If only models based on BM processes were compared to each other, this could be problematic as if traits had evolved under a homogeneous OU process (with a constant optimum). For example,

they would appear to have evolved faster in newer clades than older clades under BM process with different regimes and different rates of evolution (Revell and Collar 2009; Price et al. 2013). In this way, we used an extension of the BM model implemented in REVBayes (with the *dnPhyloOrnsteinUhlenbeckREML* function: Höhna et al. 2016), now considering the OU process.

Apart from the use of a combined dataset, we also performed BI on the matrix with only discrete characters, in order to compare the topologies obtained using MP with those resulting from model-based methods and to help us understand and quantify the stability and accuracy of obtained relationships with different datasets. The procedure also improved the choice of best hypotheses of phylogenetic relationships (according to the logic presented in Giribet 2003). To model the evolution of the partition consisting of discrete characters, there was a previous selection of models using a Bayesian approach. Because there are much narrower and more precise posterior distributions available for evolution models of discrete binary and multistate characters (see Höhna et al. 2016), it was necessary to increase the chance of sampling these conjugate distributions of parameters by adopting techniques to estimate the marginal likelihood from MCMC with likelihoods weighted by the power of different values of  $\beta$  (thermodynamic integration: Lartillot and Philippe 2006; Ronquist et al. 2012). We implemented stepping-stone sampling (ss) (Xie et al. 2011) and path sampling (ps) (Baele et al. 2012) algorithms simultaneously. Their use not only increased the chance of sampling better but also allowed us to assess whether there was a convergence of each run to the global marginal likelihood optimum (Höhna et al. 2016), by comparing the results of both sampling algorithms. Furthermore, the result obtained by the path sampling algorithm can be used to calculate the Bayes factor (Drummond and Bouckaert 2015), which refers to the ratio of marginal likelihoods of candidate models, indicating which model is the best after integrating all possible parameter values (Harmon 2019). The best model (ss = -543.67 and ps = -543.92, ln Bayes factor = 7.7 [Model 2 vs Model 3]; Table S7), as implemented in ML analysis, assumed heterogeneity in the rate of change of states among discrete characters of the matrix (Mkv model) (Lewis 2001). We drew the evolutionary rate values from a symmetric gamma distribution, discretised into four categories (Mkv +  $\Gamma$  model). We minimised the problem of ascertainment bias, which can inflate branch lengths, with the option ‘coding = “variable”’, available as an argument in *dnPhyloCTMC* function of REVBayes.

Each analysis consisted of four independent runs of 30,000,000 generations, each with four chains. Moreover, we sampled topologies and parameters from every 500 generations in each independent run. We carefully provided slight heating to the cold chains, with the *deltaHeat* option equal to 0.10, to improve the mixing of proposal values among chains. The value provided to *deltaHeat* guaranteed a probability between 20% and 60% of acceptance of sampling values from other chains (Altekar et al. 2004). We calculated Bayesian posterior probabili-

ties by computing both a majority-rule consensus tree and a maximum credibility tree of the total number of trees sampled after discarding 25% of them as burn-in.

Given that differences among topologies in the presence of heterogeneity in evolutionary rates of characters based on morphology could be large, especially when topologies under MP are compared with those of probabilistic methods (Wright and Hillis 2014; Puttick et al. 2019), we used the Robinson and Foulds (1979) symmetric absolute distance measure (RF distance)—mainly focussed on detecting topological differences (Kuhner and Yamato 2015)—as an indication of the differences observed between the topologies obtained under different types of searches and inferences used in this study (Ranala et al. 1998; Kuhner and Yamato 2015).

## 2.5. Male genitalia character optimisation

Apart from non-ambiguous character optimisation in MP trees, we optimised characters 31, 32, 33, 43, 44, 45, 46, 47, and 48, which are discrete characters of male genitalia considered to be important to the taxonomy of *B. plebejum* group, by using both MP and ML procedures over the maximum credibility tree obtained by BI. With this additional procedure of optimisation via ML in the topology obtained by BI, we also sought to explore the veracity of the reconstructed ancestral states via MP (Yang 2014). We chose these characters mainly because they can provide a general panorama about the evolution of some male genitalia components likely to be known to define or even characterise the species of *B. plebejum* group according to Lauck's (1962), Nieser's (1975), and Estévez and Polhemus' (2007) papers. Besides, these characters are considered to be comparable among all taxa in the analysis, therefore not evaluating the presence/absence of the structure and its aspect simultaneously.

The nine discrete characters had previously been modelled using the Mk model (Lewis 2001), assuming that transitions among states follow a Markov process. However, we also tentatively fitted alternative models for multistate characters, taking into account that the transition between all states of these characters did not necessarily occur at the same rate of evolution (extended Mk models according to Harmon 2019). We allowed these characters to evolve according to (1) an equal rate model (ER), the simplest model, which assumes that the rate of transition from one state to another is the same whatever the direction, as well as a stationary distribution of character states (Harmon 2019); (2) a multi-rate symmetric ordered model (ordSYM), similar to ER, but allowing pairs of states to differ from any other pair and assuming no stationary distribution of character states; (3) a model similar to ordSYM, now assuming that the forward transition rate is twice the backward rate (symORD2); (4) a model that assumes transition only occurring between adjacent states (ORD); and (5) an all-rates-different model (Paradis et al. 2004)—the most complex model because it has the highest number of free parameters ( $p = 3$ )—assuming every

possible type of transition has a different rate. In sum, we also aimed to test for differences between the forward and backward rates of character change.

We assumed that each state of the aforementioned characters could occur at the root of phylogeny with equal probability (flat prior according to Harmon 2019), instead of treating it as an estimated nuisance parameter being fitting together with the other free parameters as suggested by Fitzjohn et al. (2009). Besides, we considered the characters to be independent of each other, so those different models could be fitted to the evolution of each of these characters separately. We used ML to find a better estimate of free rate parameters, fitting the statistical models to the dataset. We used the small sample size correction of the Akaike information criterion (AICc) to compare models and to identify those providing the most efficient way to describe patterns in our data with few parameters. Accordingly, we also estimated  $\Delta\text{AICc}$  values, a relative distance between AICc scores of the proposed models, choosing the model with the smallest value, and the relative support for each model using Akaike weights ( $\omega$ ) (Burnham and Anderson 2002).

## 3. Results

### 3.1. Systematics

#### *Belostoma plebejum* group Lauck

**Diagnosis.** Clypeus reaching ocular line; vertex without median longitudinal carina; eyes rounded; sulcus not extending to lorum. Posterior width of pronotum twice its median length; prosternal carina poorly elevated, wide at its base, with its length approximately half its maximum width. Pilosity of connexivum poorly developed, narrow, covering less than half of connexivum, and not extending along genital operculum (Table S1).

#### *Belostoma lanemelo* Ribeiro & Ferrari sp. nov.

<http://zoobank.org/species/FF9C3B4A-FEF3-4D9C-BECE-348DBE418F1C>

Fig. 4A–E

**Type locality.** Pirapora Municipality, Minas Gerais State, Brazil.

**Diagnosis.** Body narrow and long. Knobshaped 'druckknopf' on mesepimeron rounded, as wide as long. Arms of phallosome diverging apically in dorsal view; lateral margin of diverticulum of phallosome strongly curved mesally, in dorsal view; width of diverticulum as long as its length, in ventral view.



**Description. Male.** Body length: 14.4 mm; greatest width of body: 6.3 mm. **Female.** Body length: 13.7 mm; greatest width of body: 6.0 mm. Body whitish brown. **Head.**—Anterior frontogenal suture (anteclypeus-maxillary plate suture) shorter than posterior frontogenal suture (anteclypeus-loral plate suture); anteoculus slightly shorter than interoculus (from 0.89 to 0.94 times); length of visible segment I of rostrum from 0.78 to 0.88 times the length of visible segment II. **Thorax.**—Sclerite 1 of distal median plate becoming broader medially, as wide as long; notch or indentation much more poorly developed along its basilateral margin, which is somewhat concave (Fig. 4A); apical portion with lobed process poorly developed, rounded; posterior portion articulated with veins 1A and 2A bearing process directed posteriorly shorter than half of the longitudinal length of sclerite 1; process not indented apically and approximately straight; knobshaped ‘druckknopf’ on mesepimeron rounded, as wide as long (Fig. 2A); prosternal carina poorly elevated, rectangular, quite pointed anteriorly, with its anterior margin smoothly curved, not projected anterad in lateral view; apex without tubercles (Fig. 4B). **Abdomen.**—Pilosity covering about one third of connexivum, poorly developed on penultimate visible segment. **Male genitalia** (Fig. 4C–E).—Arms of phallosoma diverging apically, shorter than its anterior region (base of phallosoma), not covering sides of ventral diverticulum in dorsal view. Ventral diverticulum curved, rounded apically in lateral view; apical margin smoothly concave, with prominence in dorsal view; lateral margins slender, not thickened, curved at mid-length, directly laterally in dorsal and ventral views. Width of diverticulum as long as its median length in ventral view. Ventroapical protuberance not distinctly produced in lateral view, approximately triangular in ventral view.

**Remarks.** *Belostoma lanemelo* sp. nov. appears closely related to the species of the *B. pygmeum* group, as well as *B. plebejum*. Members of the *B. pygmeum* group share with the new species the posterior portion of the sclerite 1 of the distal median plate, articulated with veins 1A and 2A, bearing a short process directed posteriorly. The new species also shares with *B. nicaeum* and *B. pygmeum* a sclerite 1 that is somewhat straight along its basilateral margin, without any notch or indentation. *Belostoma plebejum* instead shares with *B. lanemelo* sp. nov. a rounded knobshaped ‘druckknopf’. The new species, however, can be distinguished from the species of the *B. pygmeum* group by the body being narrower and longer. Additionally, the new species may be easily mistakenly keyed to *B. plebejum* when using either Estévez and Polhemus’ (2007) or Ribeiro and Alecrim’s (2008) keys to the *B. plebejum* group because of the considerable uniformity in body structures and the smooth genitalic variation among the species of *B. plebejum* group s.l. In fact, *B. lanemelo* sp. nov. shares with *B. plebejum* slender, not thickened, lateral margins of the ventral diverticulum, and a body size that is more than 13.0 mm long. The new species can be distinguished from *B. plebejum* by the arms of the phallosoma diverging in dorsal view and

by the width of diverticulum as long as its median length in ventral view.

**Etymology.** The specific epithet is dedicated in honour of Dr Alan Lane de Melo (Universidade Federal de Minas Gerais, Brazil) for his contributions to the study of Brazilian giant water bug taxonomy.

**Material examined. Holotype.** BRAZIL • 1 ♂; Minas Gerais State, Pirapora Municipality; 17°20.40’S, 44°55.80’W; Nov. 1975; M. Alvarenga leg.; J.R.I. Ribeiro det.; AMNH.—**Paratype.** BRAZIL • 1 ♀; same data as for holotype.

### *Belostoma nieseri* Ribeiro & Ferrari, sp. nov.

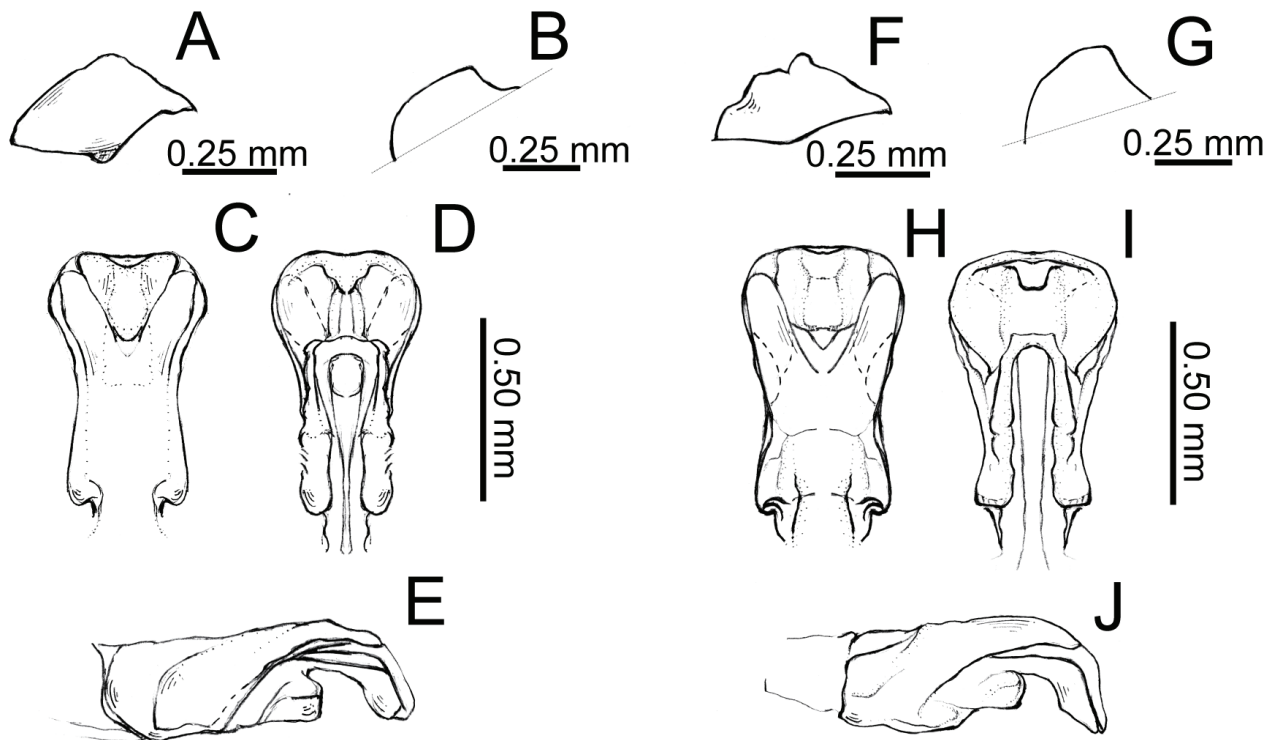
<http://zoobank.org/species/4CEA8935-8295-4380-BC3B-2B5161E1BD4C>

Fig. 4F–J

**Type locality.** Sinop Municipality, Mato Grosso State, Brazil.

**Diagnosis.** Knobshaped ‘druckknopf’ on mesepimeron piriform, longer than wide. Arms of phallosoma divergent apically and broader, in dorsal view, covering entirely the sides of diverticulum; lateral margin of diverticulum of phallosoma weakly curved mesally, in dorsal view; diverticulum of phallosoma narrower than long, with its width less than one time its median length, in ventral view.

**Description. Male.** Body length: 13.5 mm; greatest width of body: 6.2 mm. **Female.** Body length: 13.5 mm; greatest width of body: 6.4 mm. Body whitish brown. — **Head.** Anterior frontogenal suture shorter than posterior frontogenal suture; anteoculus slightly shorter than interoculus (from 0.93 to 0.96 times); length of visible segment I of rostrum from 0.80 to 0.84 times the length of visible segment II. — **Thorax.** Sclerite 1 of distal median plate moderately narrow, becoming somewhat broader medially, narrower than long; notch or indentation not developed along its basilateral margin, which is straight (Fig. 4F); apical portion with distinct lobed process, broadly rounded; posterior portion articulated with veins 1A and 2A bearing a process directed posteriorly as long as half of longitudinal length of sclerite 1; process not indented apically and approximately straight; knobshaped ‘druckknopf’ on mesepimeron piriform, longer than wide (Fig. 2B); prosternal carina somewhat elevated, not rectangular, pointed anteriorly, with its anterior margin smoothly curved, not projected anterad in lateral view (Fig. 4G); apex without tubercles. — **Abdomen.** Pilosity covering about one third of connexivum, poorly developed on penultimate visible segment. — **Male genitalia** (Fig. 4H–J). Arms of phallosoma diverging apically, shorter than its anterior region, covering sides of ventral diverticulum in dorsal view. Ventral diverticulum curved, becoming gradually narrowly rounded apically in lateral



**Figure 4.** *Belostoma lanemelo* Ribeiro & Ferrari, **sp. nov.** and *B. nieseri* Ribeiro & Ferrari, **sp. nov.** A–E Male holotype of *B. lanemelo* **sp. nov.** from Minas Gerais State, Brazil; F–J Male holotype of *B. nieseri* **sp. nov.** from Mato Grosso State. A Sclerite 1 of distal median plate of hindwing; B prosternal carina (lateral view); C–E male genitalia: phallosoma (phallosoma, dorsal arms, and diverticulum); C dorsal view; D ventral view; E lateral view; F sclerite 1 of the distal median plate of the hind wing; G prosternal carina (lateral view); H–J male genitalia: phallosoma (phallosoma, dorsal arms, and diverticulum); H dorsal view; I ventral view; J lateral view.

view; apical margin approximately straight, with small prominence in dorsal view; lateral margins slender, not thickened, curved at mid-length, slightly directly laterally in dorsal and ventral views. Width of diverticulum about 0.70 times its median length in ventral view. Ventroapical protuberance distinctly produced in lateral view, swollen, and approximately quadrangular in ventral view.

**Remarks.** *Belostoma nieseri* **sp. nov.** may be mistakenly keyed to *B. plebejum* when using Nieser and Melo's (1997: 60) key to *Belostoma* species because both species can measure between 13.5 and 15.0 mm. In fact, the body dimensions of members of *B. nieseri* **sp. nov.** resembles that of *B. plebejum*, but the former possesses a slightly narrower shape. *Belostoma nieseri* **sp. nov.** also shares with *B. plebejum* a ventral diverticulum with an apical margin that is approximately straight, with prominence

in dorsal view, the ventroapical protuberance approximately quadrangular, somewhat swollen (especially in *B. plebejum*), and the presence of mesal sulcus on ventral diverticulum in dorsal aspect. However, the arms of the phallosoma in *B. plebejum* are obviously parallel and slender, whereas they diverge apically and are broader in *B. nieseri* **sp. nov.**

**Etymology.** The specific epithet is dedicated in honour of Dr Nico Nieser (the Netherlands) for his contributions to the study of giant water bug taxonomy.

**Material examined.** *Holotype*. BRAZIL • 1 ♂; Mato Grosso State, Sinop Municipality; 11°51.4859'S, 55°30.3359'W; Oct. 1974; M. Alvarenga leg.; J.R.I. Ribeiro det.; AMNH.—*Paratype*. BRAZIL • 1 ♀; Pará State, Jacareacanga Municipality; 7°42'S, 57°36'W; Dec. 1968; M. Alvarenga leg.; J.R.I. Ribeiro det.; AMNH.

### Key to species of the *Belostoma plebejum* group sensu Nieser based on adult specimens (with new species included)

- 1a Posterior pronotal width (usually not the greatest one) twice its median length; prosternal carina usually not elevated and much broader at its base than the apex in lateral view (as in Fig. 7F); phallosoma with diverticulum strongly curved downward, in lateral view (Fig. 5) (*B. plebejum* group *s.l.*) .....2
- 1b Posterior pronotal width (usually not the greatest one) usually less than twice its median length; prosternal carina elevated and with its base as broad as its apex in lateral view, if not, then *B. oxyurum* (from *B. oxyurum* group); phallosoma with diverticulum poorly curved downward, in lateral view (Fig. 5).....*B. oxyurum*, *B. denticolle*, and *B. minor* groups

- 2a (1a) — Phallosoma with diverticulum never deformed, without deep ‘V’-shaped-like cut along apical portion, in dorsal view; mesal portion without deep longitudinal groove (a sulcus), in dorsal view; diverticulum never acute at its apex, in lateral view; margins not thickened, in lateral view (Fig. 5) (*B. plebejum* group s.s.).....3
- 2b Phallosoma with diverticulum deformed, with deep ‘V’-shaped-like cut along apical portion, in dorsal view; mesal portion with deep longitudinal groove, in dorsal view; diverticulum acute at its apex, in lateral view; margins clearly thickened, in lateral view; if not quite evident, then *B. pygmeum* (*B. pygmeum* group) .....10
- 3a (2a) — Knobshaped ‘druckknopf’ on mesepimeron rounded, as wide as long (Fig. 2A); if not, then arms of phallotheca (‘dorsal arms’) broad, covering part of sides of diverticulum in dorsal view .....4
- 3b Knobshaped ‘druckknopf’ on mesepimeron piriform, longer than wide (Fig. 2B) .....7
- 4a (3a) — Length of the first visible segment of rostrum between 0.80 and 0.87 times length of the second visible segment; arms of phallotheca clearly narrow, slightly covering sides of the diverticulum, in dorsal view (Fig. 5) ..... 5
- 4b Length of the first visible segment of rostrum usually 0.87 times or more the length of the second visible segment; arms of phallotheca neither much narrower, nor covering the sides of the diverticulum, in dorsal view .....8
- 5a (4a) — Prosternal carina rectangular, obtuse at apex, in lateral view; apical portion of sclerite 1 of the distal median plate with a lobed process poorly developed, rounded; arms of phallotheca parallel, as long as basal region of phallotheca, in dorsal view (Fig. 5) ..... *B. plebejum* (Stål) (Argentina; Bolivia; Brazil: Bahia, Espírito Santo, Minas Gerais, Pará, Paraíba, Rio de Janeiro, Rio Grande do Sul; Paraguay; Peru; Uruguay)
- 5b Prosternal carina slightly rectangular, acute at apex, in lateral view (as Fig. 7F); apical portion of sclerite 1 of the distal median plate with a distinct lobed process, broadly rounded (as Fig. 6G); arms of phallotheca slightly convergent at apex, shorter than basal region of phallotheca, in dorsal view.....6
- 6a (5b) — Total body length most often more than 10.0 mm; sclerite 1 of the distal median plate as wide as long, robust; posterior portion of sclerite 1 of the distal median plate articulated with veins 1A and 2A without evident narrowing along its lateral margins (as Fig. 6F); diverticulum of phallosoma as wide as long, in ventral view (Fig. 5)..... *B. nessimiani* Ribeiro and Alecrim (Brazil: Amazonas, Pará, Rio Grande do Sul)
- 6b Total body length usually less than 11.0 mm; sclerite 1 of the distal median plate evidently narrow, longer than wide; posterior portion of sclerite 1 of the distal median plate articulated with veins 1A and 2A narrowing along its lateral margins (Fig. 6G); width of diverticulum of phallosoma 1.50 times its median length, in ventral view (Fig. 5)..... *B. parvum* Estévez and Polhemus (Colombia, Guyana, Brazil: Amazonas, Pará; Suriname; Venezuela)
- 7a (3b) — Total body length between 11.0 and 12.0 mm; anterior frontogenal suture as long as posterior frontogenal suture, in dorsal view; apical portion of sclerite 1 of distal median plate with distinct lobed process acute (Fig. 6E); lateral margin of diverticulum of phallosoma strongly curved mesally, in dorsal view (Fig. 5); arms of phallotheca narrow, not entirely covering sides of diverticulum, in dorsal view; diverticulum of phallosoma as wide as long, in dorsal view..... *B. estevezae* Ribeiro and Alecrim (Brazil: Mato Grosso, Rio Grande do Sul)
- 7b Total body length more than 13.0 mm; anterior frontogenal suture shorter than posterior frontogenal suture, in dorsal view; apical portion of sclerite 1 of distal median plate with lobed process broadly rounded (Fig. 4F); lateral margin of diverticulum of phallosoma weakly curved mesally, in dorsal view (Fig. 4I); arms of phallotheca slightly wider, covering entirely the sides of diverticulum, in dorsal view (Fig. 4H); diverticulum of phallosoma narrower than long, with its width less than one time its median length, in ventral view ..... *B. nieseri* sp. nov. (Brazil: Mato Grosso, Pará)
- 8a (4b) — Posterior portion articulated with veins 1A and 2A bearing process directed posteriorly, narrowing abruptly along its basilateral margin (as Fig. 6E); arms of phallotheca as long as basal portion of phallotheca, in dorsal view (Fig. 5); diverticulum of phallosoma usually with medial and longitudinal depression or sulcus weakly developed, in dorsal view..... *B. minusculum* (Uhler) (Brazil: Amazonas, Mato Grosso, Minas Gerais, Pará; Costa Rica; Honduras; Nicaragua; Panama; Trinidad and Tobago; Venezuela)
- 8b Posterior portion articulated with veins 1A and 2A bearing process directed posteriorly, not narrowing along its basilateral margin (as Fig. 4A); arms of phallotheca shorter than basal portion of phallotheca, in dorsal view (Fig. 5 and as Fig. 4C); diverticulum of phallosoma medially depressed, with longitudinal sulcus strongly developed, in dorsal view .....9
- 9a (8b) — Total body length more than 13.5 mm; prosternal carina rectangular, acute at apex, in lateral view; posterior portion articulated with veins 1A and 2A bearing process weakly directed posteriorly, shorter than half of longitudinal length of entire sclerite 1 of distal median plate (Fig. 4A); notch or indentation poorly developed along its basal margin, which is somewhat concave or straight; arms of phallotheca diverging, in dorsal view (Fig. 4C); lateral margin of diverticulum of phallosoma strongly curved mesally, in dorsal view ..... *B. lanemeloï* sp. nov. (Brazil: Minas Gerais)
- 9b Total body length not exceeding 13.5 mm; prosternal carina slightly rectangular, obtuse at apex, in lateral view; posterior portion articulated with veins 1A and 2A bearing process directed posteriorly, subequal or as long as half of longitudinal length of entire sclerite 1 of distal median plate; notch or indentation absent along its basal margin, which is almost straight (Fig. 6F); arms of phallotheca abruptly convergent near to apex, in dorsal view (Fig. 5); lateral margin of diverticulum of phallosoma slightly curved mesally, in dorsal view..... *B. micantulum* (Stål)



(Argentina; Bolivia; Brazil: Amazonas, Espírito Santo, Mato Grosso, Pará, Paraíba, Rio de Janeiro; Colombia; Guyana; Paraguay; Venezuela)

- 10a (2b)** — Posterior portion articulated with veins 1A and 2A bearing a process weakly directed posteriorly, shorter than half of longitudinal length of entire sclerite 1 of distal median plate (as Fig. 6E); apical portion with distinct lobed process, conspicuously acute (as Fig. 6G); arms of phallotheca as long as basal portion of phallotheca, in dorsal view; diverticulum of phallosoma wider than long, with its width between two or three times its median length, in ventral view (Fig. 5)..... **11**
- 10b** Posterior portion articulated with veins 1A and 2A bearing a process directed posteriorly, as long as half of entire sclerite 1 of distal median plate along its longitudinal length; apical portion with distinct lobed process, rounded (as Fig. 6C); arms of phallotheca shorter than basal portion of phallotheca in dorsal view; diverticulum of phallosoma wider than long, with its width more than three times its median length, in ventral view (Fig. 5).....
- B. lariversi** De Carlo (Bolivia; Brazil: Amazonas, Mato Grosso, Minas Gerais, Paraná; Paraguay; Peru; Uruguay)
- 11a (10a)** — Prosternal carina poorly elevated, projected anteriorly, in lateral view (as Fig. 7E); posterior portion articulated with veins 1A and 2A bearing process directed posteriorly, abruptly narrowing along its basilateral margin (as Fig. 6E); diverticulum of phallosoma with its width 2.50 times its median length, in ventral view (Fig. 5)..... **B. nicaeum** Estévez and Polhemus (Brazil: Amazonas, Bahia, Mato Grosso, Mato Grosso do Sul)
- 11b** Prosternal carina poorly elevated, not projected anteriorly, in lateral view (as Fig. 7F); posterior portion articulated with veins 1A and 2A bearing process directed posteriorly, without abrupt narrowing along its basilateral margin (as Fig. 6F); diverticulum of phallosoma with its width two times its median length, in ventral view (Fig. 5)..... **B. pygmeum** (Dufour) (Bolivia; Brazil: Amazonas, Bahia, Mato Grosso, Mato Grosso do Sul, Pará, São Paulo; Paraguay; Peru)

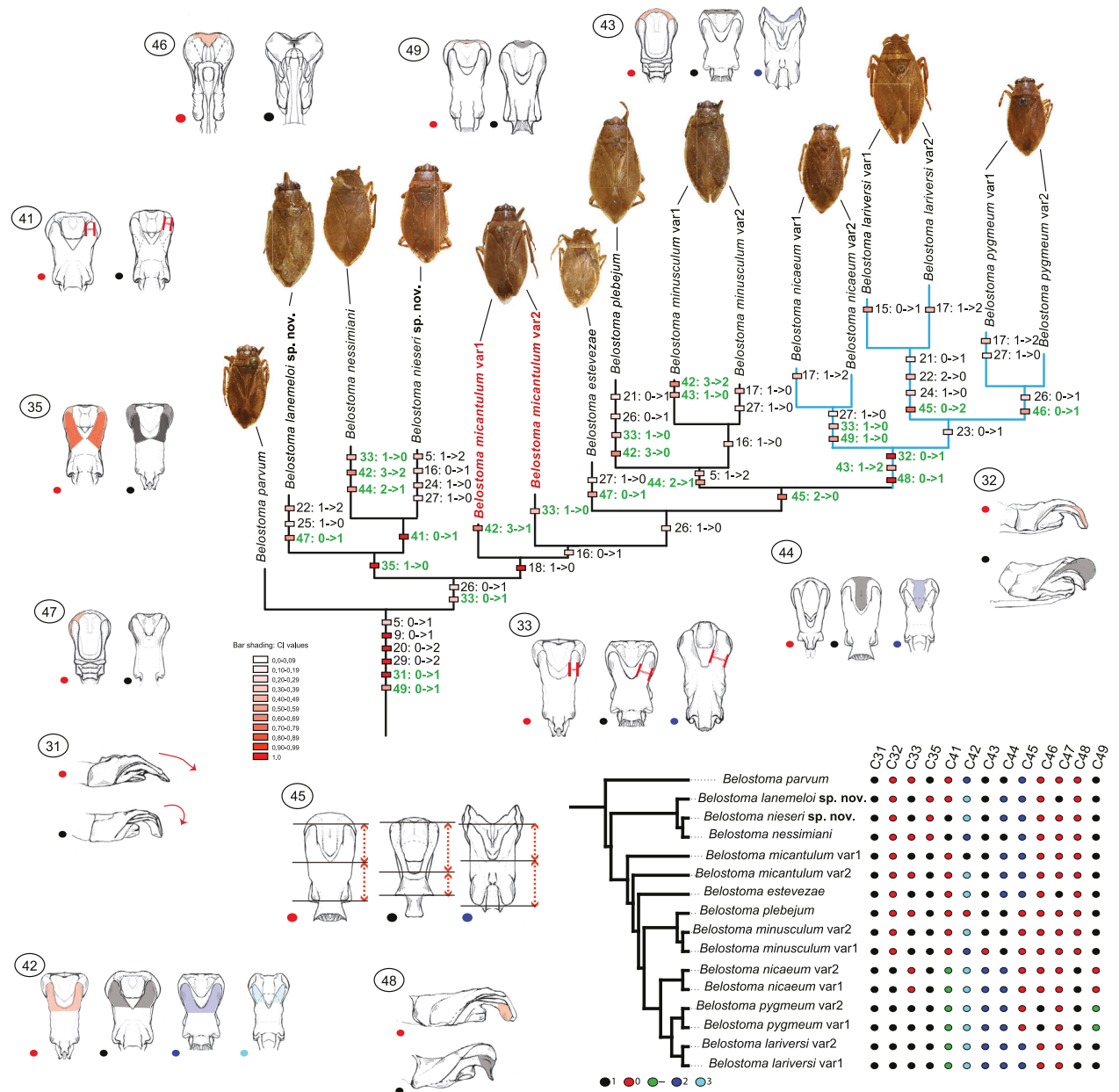
### 3.2. Phylogeny of the *B. plebejum* group *sensu* Nieser (1975)

A summary of the results obtained by different combinations of weightings and characters is shown in Table 1. The main morphological structures and characters, as well as strict consensus trees, are indicated in Fig. 4C–E and Fig. 4H–J (male genitalia), Fig. 5 (unambiguous optimisation of discrete characters over the tree), Fig. 6 (sclerite 1 of distal median plate), Fig. 7 (prosternal carinae), and Fig. 8 (optimisation of continuous characters over the tree). In all analyses, the *B. plebejum* group is monophyletic (BS = 95%; BRel = 69%), whereas the *B. plebejum* group *s.s.*, as conceived by D. Lauck, was recovered as paraphyletic in relation to the *B. pygmeum* group (Fig. 5). Of the six synapomorphies supporting the *B. plebejum* group, four are non-homoplastic and one of them refers to the phallosoma with the diverticulum strongly curved ventrally (character 31: 1). It is also worth mentioning a general trend in a decreased body length through its members (character 1, Fig. 8). The *B. pygmeum* group *sensu* Lauck (1962) was also recovered as a monophyletic group in almost all analyses, except in the scheme DISC + EW (analysis one), despite its low support (BS = 55%; BRel = 52%). The clade is supported by three synapomorphies—all of them based on the morphology of male genitalia—two of which are non-homoplastic: margins of the diverticulum of the phallosoma clearly thickened (character 32: 1); apical margin of the diverticulum conspicuously acute, in a lateral view (character 48: 1). Considering continuous characters, there was a general enlargement of male genitalia through the members of this group, which is evidenced by the increase in the ratio of the width to the length of the diverticulum (character 30) (Fig. 8). Based on the obtained topologies found by MP analyses of the data matrix consisting of different combinations of matrices of discrete and continuous characters, the clade com-

posed of all species of the *B. pygmeum* group remained consistent throughout them (Fig. 8). Besides, the internal relation between *B. lariversi* and *B. pygmeum* was recovered across almost all schemes combining character matrices with different weights, except in the consensus tree of analysis one (i.e., without continuous characters and under EW; supported only by BS: BS = 81%, BRel = 16%). Our findings support such a clade by the presence of a distinct lobed and robust process, situated along the apical portion of sclerite 1 of the distal median plate (as in Fig. 6A) as a homoplastic synapomorphy.

Conversely, almost all analyses recovered the clade *B. plebejum* + *B. minusculum* with low support (BS < 50%; BRel = 15%), except in analyses one (DISC + EW) and four (DISC + CONT + EW) (Figs 5 and 8). Two homoplastic synapomorphies supported the clade, including the presence of medial and longitudinal depression or sulcus poorly developed on the diverticulum of phallosoma (Fig. 5). Regarding continuous characters, there is either a slight increase in the distance between eyes or a decrease in the size of eyes, which is indicated by these species sharing the increase in the posterior interocular width/width of an eye ratio values (character 7, Fig. 8). Accordingly, such a clade was not consistent across analyses when that condition was combined, in some way, with any of the continuous characters 1, 2, 3, 4, 6, and 30 (Table S8).

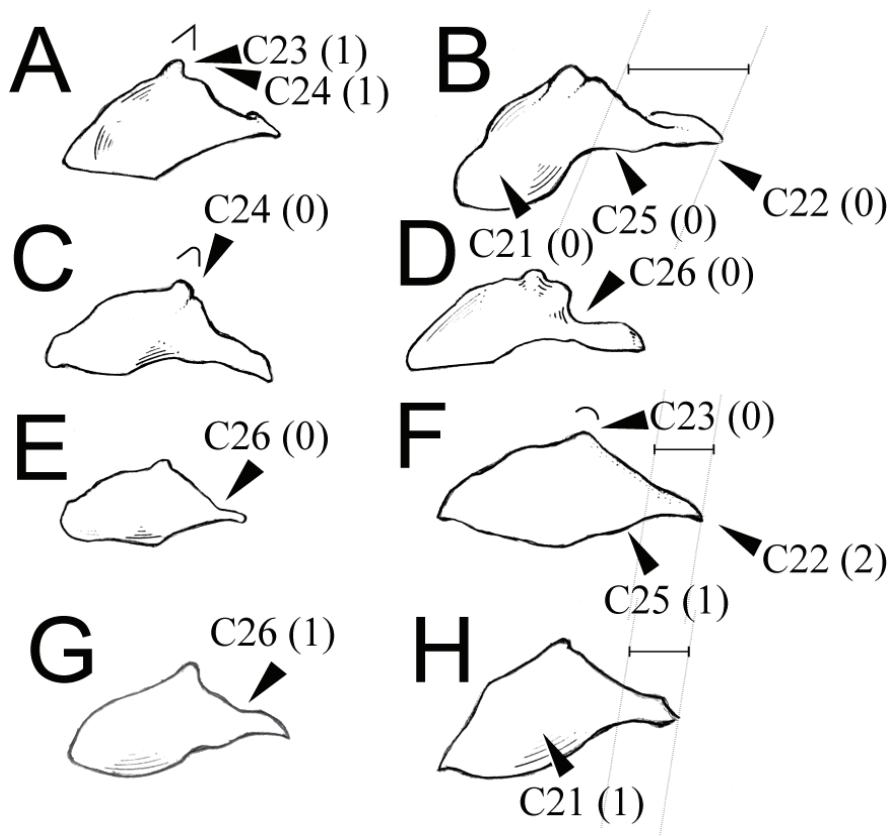
Analyses three (DISC + IW<sub>est</sub>) and six (DISC + CONT + IW<sub>est</sub>) provided low but some support for two topologies (Table S5, Table S6): an overarching clade comprising *B. lanemelo* sp. nov., *B. nessimiani*, and *B. nieseri* sp. nov. (BS < 50%; BRel = 21%), as well as an internal relation between the last two species (BS < 50%; BRel = 21%) (Figs 5 and 8). The first clade is supported by the presence of slightly longer arms of the phallotheca compared with other species, a non-homoplastic synapomorphy, so this condition renders to such arms the possibility of clearly



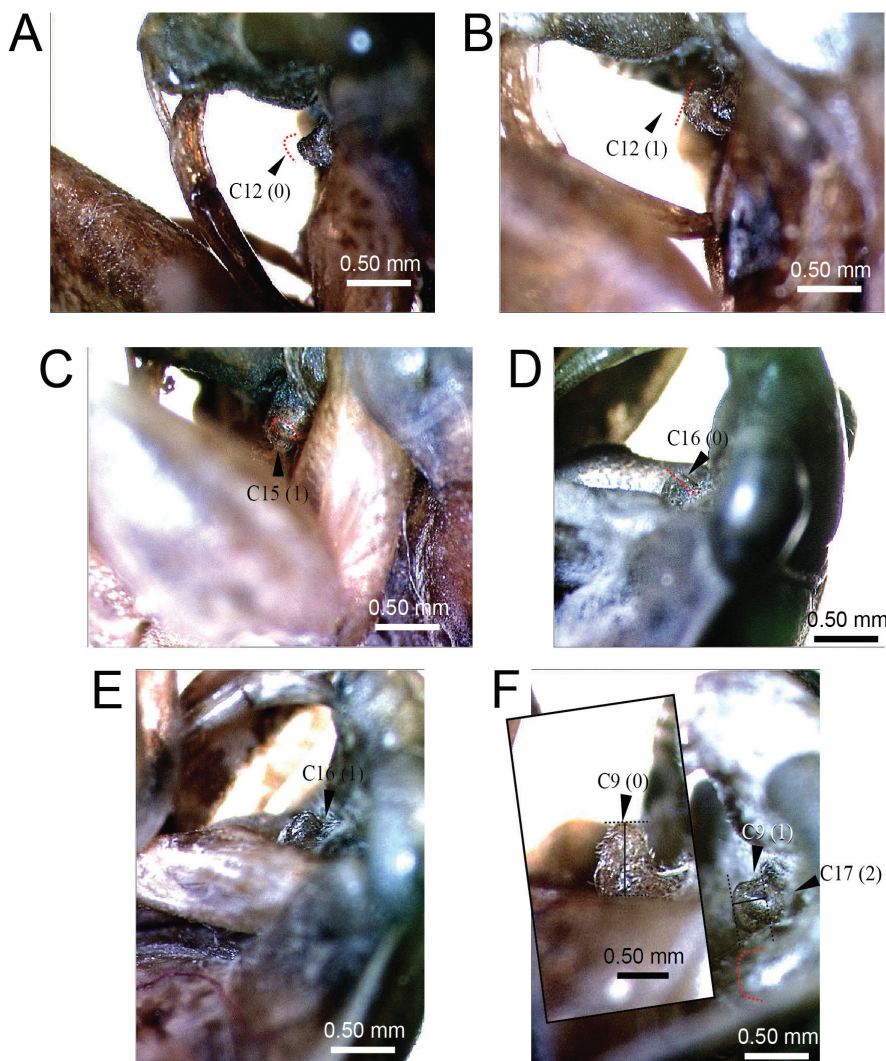
**Figure 5.** Maximum parsimony inference from analysis six (all datasets under implied weights,  $K = 3$ ). Consensus of four most parsimonious trees ( $L = 142.3$ ,  $CI = 0.44$ ,  $RI = 0.74$ ,  $fit = 14.61$ ) with non-ambiguous discrete characters optimised. Apart from non-ambiguous character optimisation, the discrete characters based on male genitalia and their states are indicated on male genitalia sketches and are coloured accordingly. Green character labels and their optimisations refer to the characters based on male genitalia. Numbers inside circles refer to character numbers and coloured circles indicate the states of each aforementioned character. Clades are coloured according to traditional classification: *Belostoma plebejum* group (black clades) and *B. pygmeum* group (light blue clades). The coloured names refer to the taxa whose phylogenetic integrity was not corroborated. Coloured squares refer to consistency index values. Outgroups were omitted. Abbreviation: CI, consistency index.

reaching the posterolateral margins of the diverticulum (character 35: 0). There is also a slight decrease in the ratio of the anteocular to interocular medial lengths (character 41: 1), also a non-homoplastic synapomorphy (Fig. 5). Accordingly, the monophyly of that clade was not recovered when only that metric condition was combined, in some way, with any of the continuous characters 1, 2, and 7 (Table S8), such that the resulting topologies showed differences of six SPR units from the strict consensus tree generated by the whole dataset with  $IW_{est}$  (i.e., analysis six). Likewise, the second clade recovering *B. nessimiani* as the sister of *B. nesi* sp. nov. is supported

by the arms of the phallosome clearly covering the lateral margins of the diverticulum of the phallosoma (character 41: 1), also a non-homoplastic synapomorphy (Fig. 5). There is also both an increase in the posterior interocular width/the anteocular length ratio values and a decrease in the ratio between this width and the width of an eye (characters 6 and 7, respectively). Finally, the continuous characters 6 and 7 provided further support for the aforementioned clade whenever they were combined in the matrices (Table S8).



**Figure 6.** Sclerite 1 of the distal median plate. Different aspects of the sclerite: form, degree of development of process from posterior portion articulated with  $A_1$  and  $A_2$  veins, degree of development and form of apical lobed process, shape of basal margin of process from posterior portion articulated with  $A_1$  and  $A_2$  veins, and form of the basilateral margin of the process from posterior portion articulated with  $A_1$  and  $A_2$  veins. Six characters were obtained based on the study of sclerite 1 of the distal median plate (characters 21–26). **A** *Belostoma denticolle* Montandon var1 **B** *B. minor* (Pasilot de Beauvois) **C** *B. estevezae* Ribeiro and Alecrim **D** *B. oxyurum* (Dufour) **E** *B. sanctulum* Montandon **F** *B. micantulum* (Stål) var1 and var2 **G** *B. parvum* Estévez and Polhemus **H** *B. plebejum* (Stål).



**Figure 7.** Lateral view of part of the head, rostrum, prothorax with the prosternal carina, and forelegs. Characters were obtained based on the different aspects of the prosternal carina in lateral view: form and shape, degree of enlargement taking into account its base, positioning relative to the longitudinal axis of the body, form of the apex and anterior margin, presence of tubercles on the apex. **A** *Belostoma amazonum* Estévez and Polhemus var2 **B** *B. horvathi* Montandon **C** *B. oxyurum* (Dufour) **D** *B. minor* (Pasilot de Beauvois) **E** *B. nicaeum* Estévez and Polhemus **F** *B. pygmeum* (Dufour).



### 3.3. Effect of progressively adding continuous characters

The seven continuous characters herein used allowed us to produce 127 different combinations of continuous characters (from one-to-one to seven-to-seven combinations), each of them having been submitted together with the matrix of discrete characters to MP analyses. Of these, MP analyses using 87 combinations of continuous characters together with discrete characters exclusively found strict consensus topologies that did not differ from that topology found using only the 42 discrete characters with  $K$ -value estimated from the dataset (analysis three: DISC +  $IW_{est}$ ); the SPR distances were equal to zero. On the contrary, 36 combinations produced together with discrete character topologies revealed two units of SPR distance, and only four combinations produced topologies with six units of SPR distance when compared to the strict consensus topology found by MP analyses when using only discrete characters (Table S8).

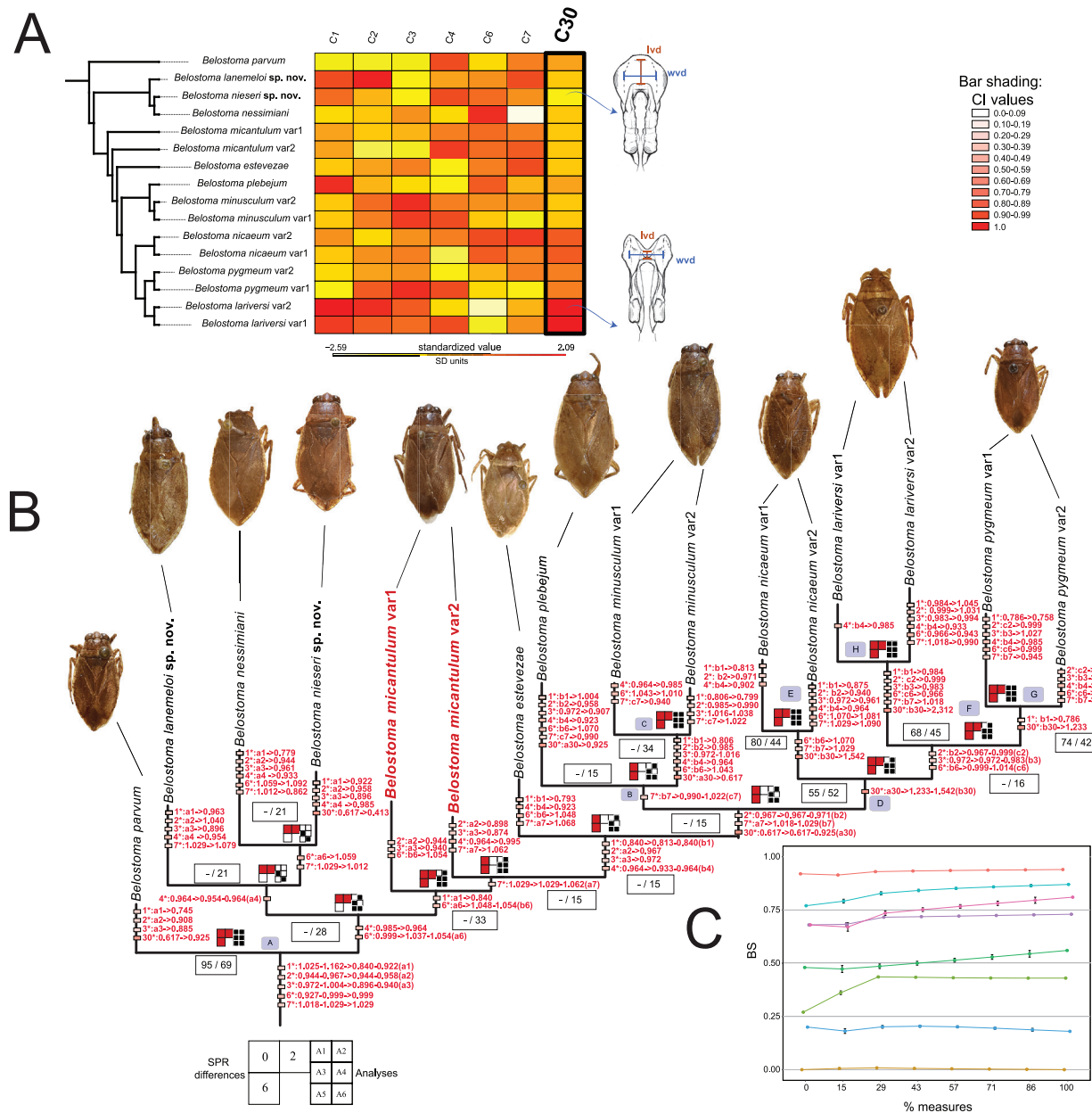
The measurements that somehow compose the four combinations of characters that produced the worst topologies (i.e., topologies with the greatest distances of SPR) are total body length (character 1), the ratio between the total body length and the greatest width of body (character 2), the ratio between the first and second visible segments of the rostrum (character 3), the ratio between the median lengths of the anteoculus and interoculus (character 4), the ratio between the anterior interocular width and the width of an eye (character 6), and the ratio between the width of the diverticulum and its length in ventral view (character 30). The combination with the highest number of continuous characters concomitantly inducing the production of strict consensus topologies with the highest conflict with that topology found only with the partition of discrete characters (i.e., SPR distance equals to six) is the one that uses all the aforementioned measures, except the ratio between the median lengths of the anteoculus and interoculus.

The progressive increment in the number of continuous characters during MP analyses necessarily increased the SPR distances between strict consensus topology found with exclusive use of the partition of discrete characters and the strict consensus topologies found with the use of continuous and discrete partitions (Ordinary Least Squares [OLS] regression:  $b = 0.08$ ;  $p = 0.0187$ ). However, a proxy of node stability—herein measured by nonparametric bootstrap—was not dependent on SPR distance values, when scaling the number of continuous characters added to bootstrap values of the obtained clades through a strict consensus tree. Accordingly, the relationship between these two variables was not significant, so the estimated slope does not suggest a scenario that the increase in SPR distances shows some degree of covariation with the decrease in node supports (OLS regression:  $b = 0.01$ ;  $p = 0.7831$ ). Only the decrease in retention index values showed some significant degree of covariation with the increment in number of continuous characters (OLS regression:  $b = -0.01$ ;  $p < 0.001$ ).

In general, the conflict between the two partitions of characters seems to be weak because the clades found by using character matrices based on measurements and relationships, including their combinations, were consistent with those found by adopting only the partition of discrete characters. More specifically, when comparing the topologies found in analyses one, two, three, and six (under MP), it is clear that the use of continuous characters in this study does not seem to have rescued different relationships among taxa (Fig. 5). This may indicate the presence of some congruence between these partitions of characters, or could have arisen as a kind of sampling behaviour where such a small number of characters—as herein adopted with the use of only seven continuous characters—had no significant effect on the phylogenetic signal present through another partition with an evident higher number of characters (42 discrete characters). Besides, it should be highlighted that the clades found in the strict consensus tree of analysis three are practically the same as found in the strict consensus tree with the use of discrete and continuous characters (Fig. 8). However, the greatest SPR distances between the strict consensus tree of analysis three and the topologies found with combinations of continuous and discrete characters were not obtained with the use of all continuous characters activated, such as topologies found in analysis six, or with these characters each used separately. Furthermore, there was no general increase in overall support with the progressive increment of continuous characters in characters matrices (Fig. 8), but instead—except for the clade comprising *B. plebejum* and *B. minusculum*, as well as the clade comprising *B. lariversi* and *B. pygmeum*—there was a slight increase in node supports, especially those independently formed by the variations of *B. minusculum*, *B. nicaeum*, and *B. lariversi*. Moreover, there seems to be a very fast increment with the progressive increase of about up to 30% of characters based on measurements and relationships, and their combinations, which in general refers to approximately three characters and all their combinations. After that, the increase through the node supports seems to happen very slowly (Fig. 8, plot).

### 3.4. Congruence between different methods and the effect of homology accumulation

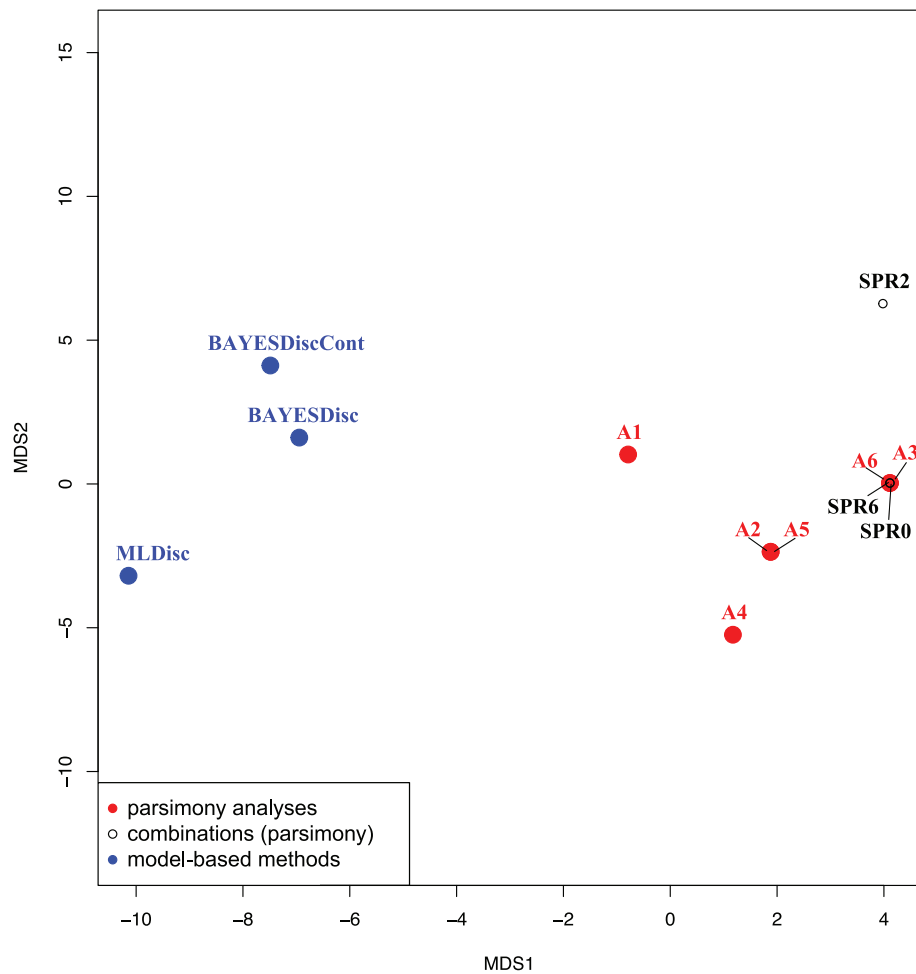
A representation of differences and similarities between the topologies found under different approaches adopted here is summarised in a nonmetric multidimensional scaling (NMDS) biplot (Fig. 9). Based on the calculation of RF distances among these topologies found in different reconstruction methods, after rotation procedures during a preliminary principal component analysis (PCA), we plotted into the relationships among these topologies taking into account their ordering onto such an orthogonal space forced to be presented in only two dimensions. This procedure preserved the estimated RF distances among topologies, and the more different topologies found under



**Figure 8.** Maximum parsimony inference from analysis six (all datasets under implied weights,  $K = 3$ ). Consensus of the four most parsimonious trees ( $L = 142.3$ ,  $CI = 0.44$ ,  $RI = 0.74$ ,  $fit = 14.61$ ) with continuous characters optimised. **A** The continuous characters and their states are indicated by coloured rectangles on the left-top corner and are coloured according to their standardised values. Measurements based on male genitalia are indicated on male genitalia sketches **B** Codes after In-transformed values referring to the ancestral state reconstructions presented on nodes were used in order to reduce space. Numbers inside rectangles above or below branches refer to bootstrap (if  $> 50\%$ ) with a Poisson correction and relative Bremer supports, respectively. Coloured squares refer to consistency index values. Letters inside grey squares refer to nodes that were also recovered in all bootstrap resamples **C** A line graph illustrates trends in bootstrap support of those nodes, as long as combinations of continuous characters were used together with discrete characters. Black and red rectangles indicate the existence of a node recovered by a maximum parsimony analysis and with a combination of continuous characters (indicated by the Subtree Pruning Regrafting distance), respectively. Outgroups were omitted. Abbreviations: BS, bootstrap support; CI, consistency index; C1, continuous character 1; C2, continuous character 2; C3, continuous character 3; C4, continuous character 4; C6, continuous character 6; C7, continuous character 7; C30, continuous character 30; lvd, medial length of the diverticulum; SPR differences, the Subtree Pruning Regrafting distance; wvd, largest width of the diverticulum.

different analyses, the higher the distance among them on the NMDS biplot. When comparing the topologies found in MP, ML, and BI analyses, there is much more discrepancy between the reconstructions based on MP and ML

(RF distances between 16 and 17) than those based on MP and BI (RF distances between 10 and 16) (Table S9). So, the major conflict between the topologies herein observed seems to be related to the adoption of different



**Figure 9.** Nonmetric multidimensional scaling (NMDS) biplot of a Robinson–Foulds symmetric absolute distance matrix. Ordination of topological differences observed between the topologies obtained under different types of searches and inferences used in this study, after rotation procedures using principal component analysis. The relationships among topologies were forced to be presented into only two dimensions, so that the procedure preserved the distances among topologies after iteratively trying to position the objects in the space in such a way to minimise a stress function (stress = 0.0314). The more different topologies found under different analyses, the higher the distance among the points on NMDS biplot. Abbreviations: A1, analysis 1; A2, analysis 2; A3, analysis 3; A4, analysis 4; A5, analysis 5; A6, analysis 6; SPR0, analysis with a combination of continuous characters producing a Subtree Pruning Regrafting distance equal to zero; SPR2, analysis with a combination of continuous characters producing the Subtree Pruning Regrafting distance equal to two; SPR6, analysis with a combination of continuous characters producing the Subtree Pruning Regrafting distance equal to six; MLDisc, maximum likelihood inference performed on the matrix of only discrete characters; BAYESDisc, Bayesian inference performed on the matrix with only discrete characters; BAYESDiscCont, Bayesian inference performed on the matrix with all datasets combined.

types of inferences rather than being related to different weighting schemes and/or different combinations of discrete and continuous characters. Without outgroups, the greatest observed conflict occurs between the topology found under ML, with only discrete characters, and those strict consensus trees found under MP, with the matrix of discrete characters and certain combinations of characters based on measurements and relationships (RF distances between 16 and 17). The consistent clades are those comprising the *B. pygmeum* group (BS = 72%); a monophyletic group composed of *B. minusculum* + *B. plebejum* (BS < 50%); the clades including variations of *B. lariversi* (BS = 53%), *B. minusculum* (BS = 50%), *B. nicaeum* (BS = 85%), and *B. pygmeum* (BS = 56%); and the overarching clade comprising the whole ingroup (BS = 86%) (Fig. S1). For the maximum credibility topologies

found under BI, with the matrix of whole dataset, only the clade composed of the variations of *B. minusculum* + *B. plebejum* is not consistent throughout topologies found using discrete and continuous characters combined. Accordingly, this clade may have enabled a subtle decrease in the distance between the maximum credibility tree found under BI based only on a matrix of discrete characters and the strict consensus tree found in MP under analysis six. In effect, the RF distance changed from 16 to 14. The other aforementioned consistent clades, however, were found with high levels of Bayesian clade probability (BCP > 85%).

In addition, a monophyletic group constituted by the species of *B. plebejum* group sensu Lauck, including *B. estevezae* and *B. nessimiani*, was recovered in the maximum credibility tree under BI, regardless of the presence



of characters based on measurements and relationships ( $BCP_{\text{discrete}} = 55\%$ ;  $BCP_{\text{discrete+continuous}} = 100\%$ ). Using the ACCTRAN criterion when discrete character optimisations were made over BI topology based on the combined discrete and continuous datasets (data not shown), this clade is supported by the anterior portion of the frontogena suture shorter than its posterior portion (character 5: 2)—a homoplastic synapomorphy—as well as the posterior portion of sclerite 1 of distal median plate articulated with veins 1A and 2A without evident narrowing along its lateral margins (Fig. 10; character 26: 1), another homoplastic synapomorphy. BI also recovered all variations of ‘*B. micantulum*’ as a monophyletic group either with low support when only discrete characters were used ( $BCP < 50\%$ ) or with slight moderate support when all dataset was used ( $BCP = 72\%$ ).

### 3.5. Evolution of characters based on male genitalia

We optimised characters based on male genitalic morphology, particularly those for which we studied the aspect of structures other than their presence or absence and their aspect simultaneously (i.e., only comparable characters), over the maximum credibility tree yielded by BI (Fig. 10). Reconstructed ancestral states of nine characters using the MP procedure revealed only one as non-homoplastic synapomorphy being shared unequivocally by the ingroup: a phallosoma with a strongly curved diverticulum downward, in lateral view (character 31:1). In fact, this synapomorphy is supposed to be diagnostic for the ingroup.

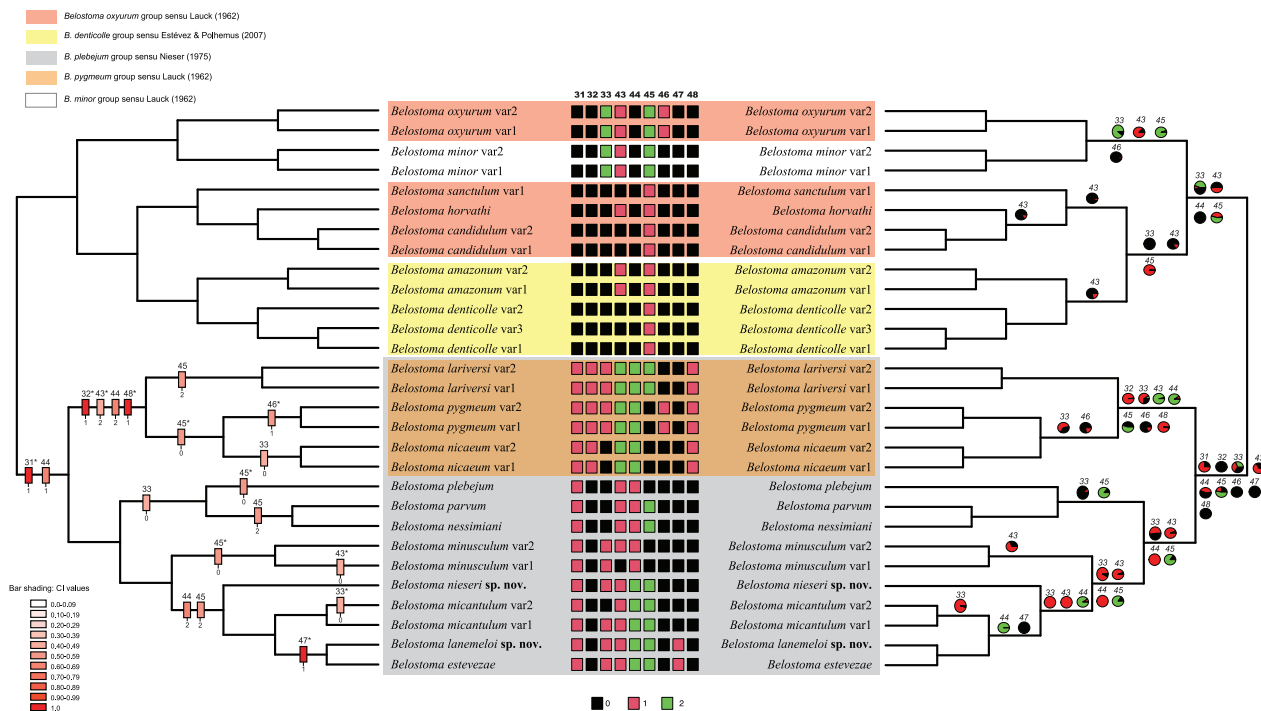
Not only did our ML ancestral state reconstructions determine and reinforce the condition of the aforementioned character, reporting the likelihood of the ancestral conditions at branches where changes were more likely to have happened, but they were also able to report the likelihood of the ancestral condition of other characters based on male genitalia. All species of the *B. plebejum* group share a diverticulum of phallosoma flat, not longitudinally depressed (character 44: 0). The present MP optimisation gives different evidence for this, having reported otherwise the presence of a medial and longitudinal depression or sulcus weakly developed on diverticulum of phallosoma (character 44: 1), as an ambiguous homoplastic synapomorphy shared by all species of the *B. plebejum* group. Our MP-reconstructed ancestral state procedures indicate that some characters, which have been considered somehow diagnostic for the *B. plebejum* group, may have instead occurred more than once within the ingroup, such as the degree of development (character 33: 0) and extension of the arms (character 45: 0). ML ancestral reconstructions over the Bayesian tree otherwise determined arms being shorter than the basal portion of phallosoma (character 45: 2) as a synapomorphy shared by all species of the *B. plebejum* group.

The narrowing of arms seems to have happened, with equal likelihood, from a more robust condition (character 33: 2) to a narrower one (character 33: 0) ( $AICc$

$= 38.095$ ;  $\omega AICc = 0.356$ ) with a higher transition rate (2.28 changes) between moderate and strong narrowing change, or as a maximally connected character (i.e., unordered) with equal transition rates ( $AICc = 38.046$ ;  $\omega AICc = 0.365$ ). A moderate narrowing condition of arms seems to have been the most likely ancestral condition in the clade formed by the species of the *B. plebejum* s.s., as well as in the clade formed by *B. nicaeum* + *B. pygmeum*. Conversely, narrower arms are the most likely condition in the clade formed by *B. nessimiani* + *B. parvum* + *B. plebejum*. MP optimisation provides other evidence, having reported narrower arms as homoplastic synapomorphic condition supporting the clade formed by *B. nicaeum* and another formed by *B. nessimiani* + *B. parvum* + *B. plebejum* only under ACCTRAN. Unequivocally, the same narrowing condition is an autapomorphy of *B. micantulum* var2.

The shape of the diverticulum seems to have evolved either from not being truncated at the apex until totally deformed, with a deep V-shaped cut, as well as in the opposite direction, or even as a minimally connected character with the possibility of a transition between 0 and 1 and between 0 and 2 (ORD2:  $AICc = 35.814$ ;  $\omega AICc = 0.371$ ; SYM:  $AICc = 36.434$ ;  $\omega AICc = 0.272$ ). Because the ORD2 model provided a slightly better fit, our results suggest that the transition  $1 \rightarrow 0 \rightarrow 1 \rightarrow 2$  is more likely than  $2 \rightarrow 1 \rightarrow 0$ , being the transition rate between a non-truncated apex (character 43: 0) and a horizontally truncated apex (character 43: 1) higher than (5.19 changes) from this condition to that one of deep ‘V’-shaped cut deformation (character 43: 2). The most likely ancestral condition of the *B. plebejum* group s.l., as well as the *B. plebejum* group s.s., seems to be the presence of a diverticulum with a horizontally truncated apex. MP optimisation unequivocally reported the V-shaped deformed condition of the apex of diverticulum as a non-homoplastic synapomorphy shared by the species of *B. pygmeum* group. Besides, a reversal seems to have unequivocally occurred in *B. minusculum* var1, given the untruncated condition of the apex of diverticulum reported to this specimen.

The evolution of the dorsomedial region of the diverticulum seems to have occurred from the absence of a longitudinal groove (flat surface) to the appearance of a very deep groove (a surface strongly depressed) ( $0 \rightarrow 1 \rightarrow 2$ ), according to the most likely model (ORD:  $AICc = 18.580$ ;  $\omega AICc = 0.777$ ). The transition rate between the presence of a poorly developed longitudinal groove and a very deep and evident groove is higher (1.79 changes) than between the non-grooved condition and the slightly grooved diverticulum condition (0.84 changes). The most likely ancestral condition of the ingroup seems to be the absence of a groove in the dorsomedial region of the diverticulum (character 44: 0), a plesiomorphic condition given by MP optimisation. However, the slight development of a groove onto the diverticulum (character 44: 1) seems to be the most likely ancestral condition of the clade formed by the species of *B. plebejum* s.s. Besides, a very deep groove is the most likely ancestral condition in the clade formed by the *B.*



**Figure 10.** Maximum likelihood and maximum parsimony ancestral states of characters based on male genitalia. Key genitalic characters of the species of the *Belostoma plebejum* group, particularly those whose aspect of structures was studied rather than their presence or absence and their aspect simultaneously (i.e., only comparable characters), were optimised over the ultrametric maximum credibility tree yielded by Bayesian inference performed on the matrix with all datasets combined. Likelihood of ancestral states are shown as pie charts only in the nodes where evolutionary changes are likely to have happened, and changes in terminal branches were omitted. Characters optimised under maximum parsimony indicated by an asterisk (on the left side) are non-ambiguous. The scoring of the states of each character is indicated by coloured rectangles.

*pygmeum* group and in the clade formed by *B. estevezeae* + *B. lanemeloï* **sp. nov.** + '*B. micantulum*' + *B. nieseri* **sp. nov.** MP optimisation reported rather similar results: the slight development of a groove onto the diverticulum is a homoplastic synapomorphy shared by the species of *B. plebejum* s.l. only under ACCTRAN, and the strongly grooved condition of diverticulum (character 44: 2) is a homoplastic synapomorphy (a parallelism) in the clade formed by the *B. pygmeum* group, as well as in the clade formed by *B. estevezeae* + *B. lanemeloï* **sp. nov.** + '*B. micantulum*' + *B. nieseri* **sp. nov.**, also only under ACCTRAN.

Considering their length when compared with the longitudinal length of the anterior portion of the phallosoma, the arms seem to have evolved with equal transition rates (0.68 changes) between maximally connected available conditions. Accordingly, arms shorter than the base of the phallosoma (character 45: 2) seem to be the most likely ancestral condition of the clade formed by the ingroup (ER: AICc = 31.044;  $\omega$ AICc = 0.498). This condition becomes increasingly most likely as the clades become less internal and slightly more likely in the clade formed by the species of *B. pygmeum* group. MP optimisation unequivocally reported that the length of arms equals to that of the base of the phallosoma (character 45: 0) is a homoplastic synapomorphy (parallelism) that defines both the clade formed by *B. nicaeum* + *B. pygmeum* and the clade formed by *B. minusculum*. Moreover, the condition is unequivocally an autapomorphy

of *B. plebejum*. The condition of the arms being shorter than the base of phallosoma (character 45: 2) is a homoplastic synapomorphy (reversal) in the clade formed by *B. lariversi*, in the clade formed by *B. nessimiani* + *B. parvum*, and in the clade formed by *B. estevezeae* + *B. lanemeloï* **sp. nov.** + '*B. micantulum*' + *B. nieseri* **sp. nov.** only under ACCTRAN.

The unique binary character proposed as diagnostic that appeared independently in some instances in the groups of species studied is the absence of a protuberance on the apicoventral portion of the diverticulum (character 46: 1) because it occurs both in *B. pygmeum* and in the clade formed by *B. oxyurum*. Besides, the aspect of the apicoventral portion of diverticulum appears to have evolved with equal rates of transition from either 0 to 1 or vice versa, with a rate of 0.59 changes. Despite being more likely, this model (ER: AICc = 17.139;  $\omega$ AICc = 0.425) is only a little more likely than the others evaluated (ORD<sub>01</sub>: AICc = 17.693; ARD: AICc = 18.178). According to the ER, the protruding aspect of the apicoventral portion is the most likely condition of the clade formed by the ingroup, which only changes in *B. pygmeum*.

## 4. Discussion

### 4.1. The monophyly of the *B. plebejum* group s.l. is corroborated and there is also some evidence of the monophyly of the *B. plebejum* s.s. and *B. pygmeum* groups

This is the first phylogenetic study to infer the relationship among members of a group of small species in *Belostoma*. The present analyses based on morphological characters (seven continuous and 42 discrete) (see supplementary file 3: List of characters) seem to provide support for the *B. plebejum* group as monophyletic, as proposed by Nieser (1975), now with two new species included. The results, therefore, corroborate the ideas of monophyly previously suggested by Estévez and Polhemus (2007) and Ribeiro and Alecrim (2008). Accordingly, in the strict consensus tree of analysis six (Fig. 5) with combined datasets and Mirande's (2009) *K* estimated from data, the following unambiguous non-homoplastic synapomorphies provide evidence for the monophyly of the ingroup: (1) prosternal carina usually not elevated and much broader at its base than the apex, in lateral view (character 9: 1); (2) posterior pronotal width (usually not the greatest one) twice its median length (character 20: 2); (3) pilosity covering about one third of the connexivum (character 29: 2); and (4) phallosoma with the diverticulum strongly curved downward, in lateral view (character 31: 1).

The enlargement of the base of the prosternal carina is probably independent of its elongation, which seems to be evident through the analyses. Previous clues provided by some essays about the development of prosternal carina and body size in Belostomatidae (Ribeiro et al. 2014, 2017, 2018) also support the idea that these conditions have to be considered separate components. Estévez and Polhemus (2007) established the condition of the prosternal carina being both wide and projected to the anterior part of body as diagnostic for *B. nicaeum* and *B. plebejum*. One can still highlight that the poorly elevated condition of the prosternal carina, in lateral view, is shared by the members of *B. plebejum* group s.l. and has been thought to be characteristic of its members (Nieser 1975; Nieser and Melo 1997; Ribeiro 2007; Ribeiro and Alecrim 2008; Stefanello 2021). Besides, based on the sketches prepared by Nieser (1975), the prosternal carina of the studied specimens of *B. amazonum*—a species we used as an outgroup—apparently shows most of the plesiomorphic conditions highlighted by our results. In their taxonomic revision of the *B. denticolle* group, Estévez and Polhemus (2001) suggested that the condition of the prosternal carina projected anterad is diagnostic for *B. amazonum* and *B. denticolle*. However, it is in fact homoplastic because that condition is also shared with other species of the *B. plebejum* group. Likewise, the presence of tubercles on the apex of the prosternal carina with a broader base (character 15: 1) is diagnostic for

*B. oxyurum*, and this has easily allowed its recognition. However, some exemplars of *B. lariversi* also show this condition on the prosternal carina, and our results provide further evidence that the aforementioned condition is also homoplastic (parallelism), favouring the hypothesis of independent origins of this condition in *B. pygmeum* group and in some outgroups.

Conversely, our results based on MP inference do not corroborate the restriction of the *B. plebejum* group to two other more inclusive groups, but rather suggest the *B. plebejum* group s.s. to be paraphyletic with respect to the *B. pygmeum* group. The existence of an apicoventral protuberance on the diverticulum of phallus (character 46: 0) was considered by Lauck (1962) as a diagnostic feature of *B. plebejum* group. One could have plausibly accepted that the species that the author used to create that group of species had indeed borne such a characteristic. However, the aforementioned diagnostic condition is not exclusive (Ribeiro and Alecrim 2008). Besides, the absence of this protuberance (character 46: 1) is unequivocally a homoplastic synapomorphy shared by the *B. pygmeum* clade, whereas the former condition proposed by Lauck (1962) is plesiomorphic in the present study. Our results still provide evidence that the apomorphic condition arose independently in both the *B. pygmeum* and *B. oxyurum* clades (parallelism).

Bayesian inference of combined datasets instead indicates *B. plebejum* group s.s. is a monophyletic group ( $BCP_{\text{combined datasets}} = 100\%$ ), which justifies a future study of this group, comprising not only morphological but also molecular datasets, in order to verify its monophyly as proposed by Lauck (1962). Lastly, the putative synapomorphies that support the monophyly of that clade, reported by MP optimisation over the maximum credibility tree, are ambiguous and have not involved any diagnostic characteristic regarding male genitalic morphology (Fig. 10). Accordingly, an anterior portion of the frontogenal suture that is shorter than the posterior portion (character 5: 2) seems to provide some phylogenetic signal to such a group, which had been highlighted by Estévez and Polhemus (2007) and Ribeiro and Alecrim (2008). However, we have suggested for the first time that the posterior portion of sclerite 1 of the distal median plate articulated with veins 1A and 2A without evident narrowing along its lateral margins (character 26: 1) is a putative synapomorphy. ML and BI also recovered the clade formed by *B. estevezae* + *B. lanemelo* **sp. nov.** as a monophyletic group, despite having low support (Fig. S1). Accordingly, this clade is unequivocally supported by the lateral margin of the diverticulum of the phallosoma that is strongly curved mesally, in dorsal view (character 47: 1), a non-homoplastic synapomorphy. Such a condition is instead a homoplastic autapomorphy (parallelism) of the aforementioned species under MP inference. We have indicated this feature for the first time as important in the taxonomy of these species, and this may be an outcome of the use of more adequate treatment of measures and relations during the analyses, as well as the achievement of fewer errors. Indeed, scenarios with high rates of homoplastic or superimposed changes make the MP inconsistent. It



is plausible to suggest that the discrepancy observed between the existence or absence of these topologies, which seem to be dependent on the method used, could be due to the presence of different evolutionary rates through the constructed characters, because we clearly obtained the phylogenetic signal of this clade and the increased accuracy of the resulting topologies. Therefore, that scenario better handles parametric models (Wright 2019), especially when matrices with a low number of characters are implemented (Wright and Hillis 2014).

Our results also provide evidence to support another hypothesis of monophyly: the clade formed by all species of the *B. pygmeum* group sensu Lauck (1962), considered by some authors to be a group of specialised species (Nieser 1975; Estévez and Polhemus 2007; Ribeiro and Alecrim 2008). Four synapomorphies support the monophyly of the clade, two of which are unequivocally non-homoplastic and based on characters of male genitalia: (1) margins of the diverticulum of the phallosoma clearly thickened, in lateral view (character 32: 1), and (2) an acute apex of the diverticulum, in lateral view (character 48:1) (Fig. 10). Furthermore, the wider condition of the diverticulum, evidenced by a width that is much greater than its length, is a continuous character that helps to support the monophyly of this group. Nieser and Melo (1997) used this relation for the first time in their book about the taxonomy of *Belostoma* species, allowing the differentiation of some species from other groups of species, such as *B. anurum* (Herrich-Schäffer, 1848) and *B. dallasi* De Carlo, 1932. This relation has been used efficiently by some other authors (e.g., Ribeiro 2000, 2007; Ribeiro et al. 2017) and has allowed a more precise understanding of the taxonomy of other species, given its common use throughout some identification keys (e.g., Ribeiro and Alecrim 2008).

The thickened condition of the laterodorsal margins of the diverticulum also allows easy recognition of the representatives of the *B. pygmeum* group, despite not being clearly visible in *B. pygmeum*. As mentioned before, Nieser (1975) established such a peculiar condition of the diverticulum of the phallosoma as an extreme case of specialisation occurring throughout the species of the *B. plebejum* group. The V-shaped deformed condition of the apex of the diverticulum (character 43: 2) is also something quite conspicuous in *B. lariversi*, and it was recovered as an unambiguous homoplastic synapomorphy for this group in the present study either under MP or BI. Furthermore, MP inference under the implicit weighting scheme recovered such a clade with low support, but model-based inferences corroborate the monophyly of that clade with high posterior probability ( $BCP_{discrete} = 98\%$ ;  $BCP_{discrete + continuous} = 100\%$ ) and with slight moderate bootstrap support under ML inference ( $BS = 72\%$ ) (Fig. S1). In addition to the already mentioned putative synapomorphies that support that clade, optimised over the maximum credibility tree, the presence of a longitudinal sulcus strongly developed in the diverticulum of the phallosoma, much evident in dorsal view (character 44: 2), supports the clade only under ACCTRAN. However, our results suggest the hypothesis of independent origins

(parallelism) of that condition in this clade and in the clade formed by *B. estevezae*, *B. lanemelo* **sp. nov.**, '*B. micantulum*', and *B. nieseri* **sp. nov.** ( $BPS_{discrete} = 52\%$ ;  $BPS_{discrete + continuous} = 86\%$ ) also only under ACCTRAN. Curiously, many authors have considered the aforementioned condition of the mesal portion of the diverticulum to be diagnostic in *B. micantulum* (e.g., Nieser and Melo 1997; Estévez and Polhemus 2007; Ribeiro 2007) and, thus, such a characteristic of this species must be re-evaluated in the future.

In the present study, ML and BI assigned *B. lariversi* as a sister-group of *B. nicaeum* + *B. pygmeum* ( $BPS_{discrete} = 65\%$ ;  $BPS_{discrete + continuous} = 98\%$ ). However, analysis six of MP inference supported a clade composed of (*B. nicaeum*, (*B. lariversi* + *B. pygmeum*)) in parenthetical form (Fig. S1). According to the MP optimisation of this character over the maximum credibility tree, the presence of arms shorter than the length of the base of the phallosoma (character 45: 2) is an ambiguous synapomorphy of the clade formed by *B. estevezae*, *B. lanemelo* **sp. nov.**, '*B. micantulum*', and *B. nieseri* **sp. nov.** under the ACCTRAN criterion, whereas the arms becoming as long as the base of the phallosoma (character 45: 0) unambiguously supports the clade formed by *B. nicaeum* + *B. pygmeum* and also appears independently in *B. plebejum* and *B. minusculum* (parallelism). Conversely, arms shorter than the length of the phallobasis is plesiomorphic in the strict consensus tree of analysis six recovered by MP inference, whereas arms becoming as long as the base of the phallosoma unambiguously supports the clade formed by the species of *B. pygmeum* group + *B. plebejum* + *B. minusculum*, with a reversal to state 2 in *B. lariversi* (Fig. 5). The different dispositions of arms, as well as their degrees of development, have long been used to recognise many *Belostoma* species, such as representatives of *B. denticolle*, *B. oxyurum*, and the *B. plebejum* group, over the history of the taxonomy of small species of *Belostoma* (Lauck 1962; Ribeiro and Estévez 2007; Estévez and Polhemus 2007; Ribeiro and Alecrim 2008). Moreover, Ribeiro and Estévez (2009) established the condition of smooth convergence towards the apex of them, taking into account their position over the diverticulum (character 42: 2), as a diagnostic characteristic of members of *B. sanctulum* (an outgroup), one species of *B. oxyurum* group. Likewise, within the ingroup, arms of the phallosoma diverging towards their apexes (character 42: 3) is a plesiomorphic condition in the strict consensus tree recovered by analysis six of MP inference; it is also shared by the clade formed by the species of the *B. pygmeum* group. Accordingly, arms converge abruptly at their apex (character 42: 1) in *B. micantulum* var1, converge smoothly at their apex in *B. nessimiani*, and are parallelly disposed (character 42: 0) in *B. plebejum*.

Based on the previous paragraphs, some increment in morphological complexity—especially through male genitalia—seems to occur within the *B. plebejum* group, and they could somewhat be associated with a process of body size decreasing in *Belostoma*. Although *B. parvum* is the smallest species of genus, this process seems to culminate in the species of *B. pygmeum* group. The

most likely process of evolution in body dimensions, according to the models evaluated in this study, is the one supported by a small force of selection channelling the characteristics based on measures and relation to an optimal global value under constant evolution rate (OU model:  $ss = -378.03$ ;  $ps = -377.49$ ;  $\ln$  Bayes factor = 3.3 - Yule + homogeneous OU model vs. Yule + relaxed BM model, or  $\ln$  Bayes factor = 5.2 - Yule + homogeneous OU model vs. Yule + homogeneous BM model) (Hansen and Martins 1996) (Table S10). Many aforementioned novelties seem to support the hypothesis—especially when male genitalic characteristics are optimised over the maximum credibility tree, obtained under BI with combined datasets—because they showed a certain tendency to arise in groups of species with smaller body sizes. Shorter arms arising of phallosoma (or phallobasis) and not reaching the lateral margins of diverticulum, for example, has been suggested as diagnostic for the *B. pygmeum* group (Lauck 1962). Besides, Nieser (1975) established that the plesiomorphic condition of this character (i.e., arms reaching the lateral margins of the diverticulum) would also represent a case of specialisation found in the *B. plebejum* group, consequently allowing the emergence of the lineage that originated the species of the *B. pygmeum* group. As suggested by our results, however, the clade formed by *B. lanemelo* **sp. nov.**, *B. nessimiani*, and *B. nieseri* **sp. nov.** under MP inference is that one that was supported unambiguously by the non-homoplastic synapomorphic condition of this character. Unfortunately, this clade was not recovered under BI, which may be a result of the presence of many homoplasies underlying the evolution of those species, the mutual presence of characters with slower or faster evolutionary rates (Wright and Hillis 2014), somehow highlighted by our investigation about evolutionary rates of some male genitalic characters, or simply an increase in accuracy that occurs as long as the size of character matrices also increases (Parins-Fukuchi 2018).

#### 4.2. *Belostoma micantulum* is not supposed to have phylogenetic integrity

We tested the phylogenetic integrity (de Pinna 1999) of four species of the *B. plebejum* group. *Belostoma lariversi* (BS = 68%; BRel = 45%), *B. minusculum* (BS < 50%; BRel = 34%), *B. nicaeum* (BS = 80%; BRel = 44%), and *B. pygmeum* (BS = 74%; BRel = 42%) with their exemplars formed their respective monophyletic clades, except for '*B. micantulum*', in all MP analyses, when combining or not combining continuous with discrete characters (Figs 5 and 8).

The variations of *B. lariversi* and *B. nicaeum* formed their respective monophyletic groups, suggesting phylogenetic integrity of these species either with MP or BI. Four unambiguous homoplastic synapomorphies supported the clade formed by the variations of *B. lariversi*, but only one referred to the degree of development of the arms of the phallosoma (Fig. 5, character 45: 2), a unique

recovered character based on the male genitalia, while three other unambiguous homoplastic synapomorphies supported *B. nicaeum*. In this case, the key characteristics based on male genitalia are as follows: narrower arms of the phallosoma (character 33: 0) and the diverticulum of the phallosoma with a weak and poorly conspicuous dilated portion along its apex, in dorsal view (character 49: 0). Besides, both species were supported by a conspicuous enlargement of diverticulum in ventral view compared with its length (character 30), especially in *B. lariversi* (Fig. 8). Therefore, despite the variations of the prosternal carina pattern morphology, it is not possible to refute the hypothesis that the terminals assigned to *B. lariversi* and *B. nicaeum* form respective single species. Such a variation in the prosternal carina has been documented by some papers (Ribeiro 2007; Ribeiro et al. 2017), which corroborates the idea that there is much intraspecific variation in this structure.

The variations of *B. minusculum* also formed a monophyletic group under the MP criterion (and also under BI), suggesting the phylogenetic integrity of the species despite the lowest support recovered—probably due to the presence of a broad prosternal carina with its anterior margin straight (character 16: 0)—the unique unambiguous homoplastic synapomorphy supporting that clade. The variations are among the 'Druckknopf' system and male genitalia, especially in the disposition of arms (characters 17, 27, 42, and 43), suggesting some degree of polymorphism among these features. Against our expectation, the complementary structure of the apparatus, located in the thorax, also seems to be polymorphic, although Ribeiro and Alecrim (2008) and A. L. Estévez (pers. comm.) indicated such a characteristic is important to identify species of *B. plebejum* group.

Under the MP criterion (and also under the Bayesian criterion), *B. pygmeum* is supported by the absence of a median apicoventral protuberance on the diverticulum, a characteristic based on the morphology of male genitalia (Fig. 10; Fig. S1), which Lauck (1962) considered being diagnostic to the species of *B. plebejum* group. Again, the complementary structure of 'Druckknopf' system apparatus, as well as the shape of the prosternal carina, are herein indicated as polymorphisms.

On the other hand, the two variations of '*B. micantulum*' did not form a monophyletic group, suggesting no phylogenetic integrity of this species with the use of the MP criterion (Fig. 5). The exemplars were instead paraphyletic in all analyses, except for analysis one. Variation 1 of *B. micantulum*, occurring in the states of Pará, Amazonas, Paraíba, Mato Grosso, Espírito Santo, and Rio de Janeiro, formed a monophyletic group with all species of the *B. pygmeum* group + *B. micantulum* var2 + *B. estevezae* + *B. plebejum* + *B. minusculum*, despite having low support (BS < 50%; BRel = 33%). The group was supported by the presence of a rectangular prosternal carina, obtuse at the apex, an unambiguous non-homoplastic synapomorphy (character 18: 0). Although all combinations of continuous characters corroborate the clade (Fig. 8), variation 2 of *B. micantulum* constituted a more internal clade, a finding that indicates segregation

of characters and differentiation (de Pinna 1999). Such an internal clade (BS < 50%; BRel = 15%) was not supported by some combinations of continuous characters together with discrete ones, despite being supported by the presence of a wide prosternal carina with anterior margin curved (character 16: 1), an unambiguous homoplastic synapomorphy. Although the variations co-occur partially in Brazil, because both are found in Pará, Amazonas, and Espírito Santo States, some differences in arms of the phallosome seem not to be just intraspecific variations as long as arms are abruptly convergent at their apexes in representatives of variation 1, which authors have reported to be more widely distributed. However, in representatives of variation 2 arms are evidently narrower, which only occurs in Pará, Amazonas, and Espírito Santo States. Lastly, variation 1 of *B. micantulum* is slightly narrower than variation 2 (Fig. 8; Table S3).

Under BI, variations of '*B. micantulum*' formed a monophyletic group with moderate support (BCP = 84%) (Fig. S1). As mentioned before, analysis one of the MP criterion also recovered these variations as a monophyletic group. It might be that the difference between topologies found under the MP criterion and model-based methods is an outcome of the interaction of discrete and continuous characters, as long as most likely evolutionary models were fitted to different evolutionary rates likely to exist among discrete and continuous datasets (Felsenstein 1985; Parins-Fukuchi 2018). Parametric methods, such as BI, are also thought to produce better results than MP in situations with small datasets and the presence of missing data (Wright and Hillis 2014; O'Reilly et al. 2018; Puttick et al. 2019). Moreover, BI usually generates results that are slightly more similar to MP with equal weighing schemes (Silveira et al. 2021). Thus, taking into account these discrepant results in the present study, we are much inclined to not deliberately refute the hypothesis that the variations assigned to '*B. micantulum*' form a single species. Nevertheless, historically the specific delimitation of *B. micantulum* has been a problem. Much has been said about the existence of putative species complexes of small *Belostoma* species. Accordingly, some important small cryptic species, in terms of general body appearance, are as follows: *B. micantulum*, *B. plebejum*, *B. pygmeum* (by some authors), and *B. denticolle* (A. Bachmann pers. comm.; Nieser 1975; Nieser and Melo 1997; Ribeiro 2007), some of them being differentiated from one another only by body size (Nieser and Melo 1997). Furthermore, the use of a diversified diet to monitor the life cycle of *B. micantulum*, under laboratory conditions, has suggested the existence of remarkable phenotypic plasticity regarding its size (Pereira et al. 1991, see Pereira 1992 for *B. plebejum*). Therefore, the idea of these variations found in *B. micantulum* being instead an outcome from differentiated interactions between specimens and environment through their lives (Armbruster et al. 2004 with examples in plants; Klingenberg 2015) is not implausible.

Stefanello et al. (2018) highlighted that taxonomy strongly or exclusively based on a study of male genital morphology in *Belostoma* may not always be recom-

mended because some intraspecific variation in male genitalia has been found inside some *Belostoma* species with the use of geometric morphometric tools. The monophyly of '*B. micantulum*' we have obtained with the use of model-based methods should also be interpreted with caution. It might still be coherent because the combination of what is supposed to be either uninformative or informative in MP provides only an overall idea of a plethora of different evolutionary rates likely to underlie their respective datasets (Wright and Hillis 2014). Besides, combined with such monophyly recovered, the existence of differences found among the disposition of arms of the phallosome described in the previous paragraph reinforces the suggestion of the existence of polymorphism in some *Belostoma* species, including '*B. micantulum*'. Recent results with the use of molecular dataset recovered some *Belostoma* lineages inside two or three small species with some level of polymorphism or phenotypic plasticity in male genitalia (Gauterio et al. 2021). This would also explain the moderate level of homoplasies found in those characters based on male genitalia that we have studied in more detail.

### 4.3. The effects of using different methods, in general, seem to be much more visible than using different combinations of continuous characters

Contrarily to Thiele's (1993) approach, we treated continuous characters as continuous series that we did not divide into an arbitrary set of states (Goloboff et al. 2006). In addition, we have used different implicit character weighting schemes and different amounts and combinations of continuous characters to investigate the effect of these practices on the character matrix, considering the evenness in the production of repeated relationships and character congruence. So, we have developed a new procedure for combining continuous characters with a discrete dataset, under MP specifically for this study. Besides, we compared recovered topologies of these analyses under the MP criterion in the light of those obtained with explicitly statistical methods (i.e., ML and BI). In the present study, comparing the topologies obtained by different approaches clearly showed the presence of different scenarios in terms of heterogeneity of evolutionary rates among characters, because we noted a certain disparity between the results obtained with different reconstructions in the biplot produced from the NMDS approach. Based on a triangular matrix of RF distances (Robinson and Foulds 1981) estimated among topologies (Table S9), the disposition of points on the biplot (Fig. 9) indicated that the discrepancy between the topologies reconstructed both under explicitly statistical methods and MP, in general, seems to be much more important than the supposed discrepancy related to the use of different weighing schemes and/or different combinations of discrete and continuous characters. However, it is worth



mentioning that the disproportionate number of discrete characters compared with the continuous dataset could be responsible for the almost nonexistence of differences among our topologies recovered with different combinations of measures and relations.

Issues related to the controversial use and treatment of continuous characters have received more attention in parsimony analyses (see Hauser and Presch 1991; Slowinski 1993; Kitching et al. 1998; Wiens 2001; Goloboff et al. 2006; Baur and Leuenberger 2011; Kock et al. 2014). In analyses carried out under different weighing schemes, the influence in using and weighting characters based on the implicit weighting approach (Goloboff 1993) onto strict consensus topologies was similar, whether or not we used continuous characters. However, the influence of adopting any weighting was greater in those topologies where weight equal to one was assigned to all characters than in those where characters received different implicit weightings, according to the number of homoplasies that a referred character could present and topological stability (as explored in Barão et al. 2020). Only when there was a weight equal to one assigned to all discrete characters (analyses one and four) did we find a greater number of alternative hypotheses, which resulted in a greater number of taxa integrating polytomies in strict consensus trees. Nevertheless, some of the clades are still in agreement with the clades recovered in analyses two, three, five, and six, whose matrices were submitted to implicit weighting procedures. Similar clades recovered through different weighting and ordering schemes constitute consistent and stable hypotheses of phylogenetic relationships (Gutberlet 1998; Barão et al. 2020). The *B. plebejum* group *sensu* Nieser (1975) and the phylogenetic integrity of some species are clear examples of this condition.

An overlooked fact instead is that a multistate character (which could be ordered or not ordered) exerts greater influence in discriminating among alternative tree topologies than a binary character (Donoghue and Ree 2000; Wiens 2001). A character with four states will have three times as much influence as a binary character, while one with 10 states will have three times as much influence as one with four states and nine times as much influence as a binary character, which is particularly relevant for continuous characters broken into an arbitrary number of discrete states (Swofford and Begle 1993). Furthermore, discrete binary characters have smaller state space and develop much more slowly at low evolutionary rates than multistate characters (Donoghue and Ree 2000; Wright and Hillis 2014; Parins-Fukuchi 2018). Therefore, treating measures and relations as such is also an advantage in scenarios with high evolutionary rates because these characters are less subject to the saturation phenomenon, as they present a high magnitude of variation (Goloboff et al. 2006) and, in turn, the loss of phylogenetic signal is lower (Wright and Hillis 2014).

Continuous characters, therefore, tend to retain more information regarding transformation and segregation, especially when submitted to explicitly statistical methods of tree reconstruction, such as BI (Wright and Hil-

lis 2014; Parins-Fukuchi 2018; Puttick et al. 2019). This seems to be the case in the present study because there was a slight increase in the stability of some nodes, despite a slight decrease in retention index values through MP analyses with the use of different combinations of continuous characters.

Nevertheless, there are known negative effects of high magnitude of variation imposed by characters based on measures and relations under implicit weighting procedures during MP reconstructions ('an asymmetric influence' according to Barão et al. 2020). The effects are usually reduced when those characters are standardised or normalised so that they can be weighted more accurately according to their degree of homoplasy (Koch et al. 2014). In the present study, we adopted weighing schemes associated with scaling or standardisation of metric and meristic data during MP procedures. However, due to the large amount of homoplasy normally attributed to the continuous characters (Wiens 1995) and a reduction likely due to the dominance of characters with a large range of values (Goloboff et al. 2006; Baur and Leuenberger 2011; Koch et al. 2014), the phylogenetic signal of discrete character partition still produced some discrepancies. These issues become evident with the extent of conflict or concordance found in those reconstructions produced by probabilistic methods. Hence, representatives of the *B. plebejum* group underwent a scenario with heterogeneity among the evolutionary rates of their studied morphological traits, more specifically when we estimated the evolutionary rates of some key characters based on male genitalia.

The procedure we have adopted involves the gradual combination of continuous characters to the matrix of discrete characters during the MP procedure. It showed that the progressive increase in the number of continuous characters necessarily increased SPR distances between the consensus tree produced only by the partition of discrete characters with implicit weighting (analysis three) and consensus trees recovered with different combinations of the two partitions, although accuracy and stability were not dependent on the SPR distance values. However, there was a slight overall progressive increase in stability with the inclusion of up to approximately 30% of traits based on measures and relations, as well as their combinations (Fig. 8). Some authors have mentioned that errors during phylogenetic reconstructions are generally smaller when continuous characters are used in conjunction with discrete character matrices (Mounce et al. 2016; Parins-Fukuchi 2018), a strategy that agrees with a 'total evidence' reasoning (de Queiroz and Donoghue 1995). Continuous characters are assigned more objectively, while they bring much accuracy because they considerably increase the size of character matrices (Gatesy et al. 1999). On the other hand, accuracy might be reduced if such included characters lead to greater covariance among them, even though they behave relatively well in terms of consistency because of lower levels of saturation so that their phylogenetic signal is not compromised (Donoghue and Ree 2000; de Bivort et al. 2010; Escapa and Catalano 2013; Wright and Hillis 2014). This phenomenon would cause a decrease in their informational

content, probably caused by their residual covariance, even under an explicitly statistical model of continuous data evolution (Parins-Fukuchi 2018). Another possibility of incongruence might be related to the existence of different selective pressures upon the modules being represented by these partitions (Clarke and Middleton 2008). As already mentioned, under MP the procedures adopted in the present study showed that an increase in node supports is not so evident after the inclusion of about 30% of continuous characters. We suggest that this seems to be caused by the existence of some covariance component that was gradually added into datasets.

Using model-based methods, our recovered phylogenetic reconstructions indicated that accuracies provided by the two dataset partitions appear to be the same, even though they behave slightly differently when used in different reconstruction methods. IQ-TREE, for example, seems to perform worse when compared with other ML software, especially when character matrices have missing data and are small (Vernygora et al. 2020). In addition, topologies recovered with BI are generally more accurate than those recovered with ML inference when both are evaluated by estimating RF distances (Brown et al. 2017). Our continuous dataset is probably under selective constraint force that seems to act similarly upon each of the seven characters and in a constant way over time. The use of a homogeneous OU model (Hansen and Martins 1996) as the most likely process of evolution to our continuous dataset highlighted such a constraint, therefore, improving the performance of our inferences. In turn, the accuracy of our reconstructions was also improved, regardless of the covariance likely to be found among these characters. This indicates that the evolutionary process of those continuous traits slowly reached stationarity towards an optimum, due to the low estimated overall evolutionary rate ( $\alpha = 0.202$ ) (Parins-Fukuchi 2018; Harmon 2019).

Polytomies may consistently show low RF distances with any other topology in a given tree space (Rannala et al. 1998). Accordingly, this could also explain why the consensus trees recovered in analyses one and four of MP inference were closer to other topologies recovered by probabilistic methods in the NMDS biplot. Metrics based on distances focussed on topologies, such as RF distance, usually evaluate differences and similarities among trees well (Kuhner and Yamato 2015). However, such a metric has been known to benefit methods of tree reconstruction—and datasets—that recover consensus trees with polytomies (Vernygora et al. 2020), therefore estimating low distances among them when compared with fully resolved trees. Likewise, trees without polytomies but with only a few very stable nodes—that is, with high support values—tend to be a problem. Unlike trees without polytomies but with few stable nodes, polytomies can still be treated as partially correct and, unfortunately, there is currently no metric that avoids such a bias created by its lacking of resolution (Rannala et al. 1998; Holder et al. 2008; O'Reilly and Donoghue 2018). In the present study, this reasoning allowed us to adopt RF distance—a metric that strongly penalises fully resolved trees with low-supported nodes—thus reducing high estimates of

RF distance among two trees when they are not quite different from each other (i.e., type-I error) (Vernygora et al. 2020). Regarding MP inferences, the NMDS biplot shows that the additional use of continuous dataset did not improve the phylogenetic information supposed to be present in our discrete dataset, probably because of the disproportional number of discrete characters compared with continuous characters in our matrix.

## 5. Conclusion

Our study is the first phylogenetic study to infer the relationship among members of a group of small species in *Belostoma*. In all analyses, the *B. plebejum* group was monophyletic, but the *B. plebejum* group *s.s.*, as conceived by Lauck, was recovered as paraphyletic in relation to the *B. pygmeum* group under MP. The following four non-homoplastic synapomorphies support the *B. plebejum* group: (1) prosternal carina usually not elevated and much broader at its base than the apex, in lateral view; (2) posterior pronotal width (usually not the greatest one) twice its median length; (3) pilosity covering about one third of the connexivum; and (4) phallosoma with a strongly curved downward diverticulum, in lateral view. Conversely, the *B. plebejum* group *s.s.* was recovered as monophyletic by BI, which justifies future studies comprising phylogenetic inferences with molecular and morphological datasets together.

We submitted some specimens of *B. lariversi*, *B. micantulum*, *B. minusculum*, *B. nicaeum*, and *B. pygmeum*—considered to show intraspecific variations in some studied structures—to the phylogenetic integrity test, under MP. Only the exemplars of '*B. micantulum*' did not formed a monophyletic clade. Their exemplars were instead, in general, paraphyletic. Accordingly, variation 1 of *B. micantulum* formed a monophyletic group with all species of the *B. pygmeum* group + *B. micantulum* var2 + *B. estevezae* + *B. plebejum* + *B. minusculum*, whereas variation 2 of *B. micantulum* constituted a more internal clade, a finding that indicates segregation of characters and differentiation. Although the variations co-occur partially, because both are found in Pará, Amazonas, and Espírito Santo Brazilian States, some differences in the arms of the phallosoma seem to be more than just intraspecific variations.

In the present study, we have developed a procedure to verify the effect of progressively adding measures and metric relationships to discrete character matrices of MP analyses. Our procedure allowed us to produce 127 different combinations of continuous characters (from one-to-one to seven-to-seven combination), each of them having been submitted together with the matrix of discrete characters to MP analyses. The progressive increment in the number of continuous characters necessarily increased the SPR distances between strict consensus topology found with exclusive use of the partition of discrete characters. Moreover, the strict consensus topologies found with the use of continuous and discrete partitions, but

nonparametric bootstrap supports—a proxy of node stability—was not dependent on SPR distance values. In general, the conflict between the two partitions of characters in the present study seemed to be weak because the clades found by both datasets were consistent with each other. This might indicate the presence of some congruence between these partitions of characters, or could have arisen as a sort of sampling behaviour where such a small number of characters had no significant effect on the phylogenetic signal present through another partition with a higher number of characters.

Even though we adopted weighing schemes associated with scaling or standardisation of metric and meristic data during MP procedures, the phylogenetic signal of discrete character partition used in this study still produced some discrepancies. These are evident with the extent of conflict or concordance found in those reconstructions produced by probabilistic methods. It might be a result of the presence of many homoplasies underlying the evolution of those species, the mutual presence of characters with slower or faster evolutionary rates—somehow highlighted by our investigation about evolutionary rates of some male genitalic characters—or simply an increase in accuracy that occurs as long as the size of character matrices also increases. Therefore, representatives of the *B. plebejum* group have undergone a scenario with heterogeneity among evolutionary rates of their studied morphological traits, more specifically when we estimated the evolutionary rates of some key characters based on male genitalia. All together could justify a future study of this group of species, comprising morphological as well as molecular datasets.

## 6. Authors' contributions

J.R.I.R. designed in part the study, analysed data, performed the study, and wrote the paper. A.F. designed the study, analysed in part the data, and wrote the paper.

## 7. Funding

The authors have no funding to report.

## 8. Competing interests

The authors have declared that no competing interests exist.

## 9. Acknowledgements

The first author gratefully acknowledges his adviser, Jorge L. Nessimian (Instituto de Biologia, UFRJ), and Alan L. Melo (Instituto de Ciências Biológicas, UFMG) for their immense support, suggestions and discussion on this manuscript and other tedious details concerning its preparation. He is still much more indebted to A. L. Melo for receiving him in Belo Horizonte Municipality, Minas Gerais State, Brazil in 1999 and allowing him to access his important collection of Nepomor-

pha from Minas Gerais State. He is also indebted to Alcimar do Lago Carvalho (MN/UFRJ) for their special support, important suggestions and discussion. Thanks to J. Becker (in memorian) for assistance with securing crucial references. This manuscript benefited from the useful comments of E. R. da Silva (UNIRIO), G. Mejdalani (MN/UFRJ), and N. Ferreira-Jr. (Instituto de Biologia, UFRJ). Thanks to Ana L. Estévez (in memorian) for her important suggestions about some character statements and the identification of *Belostoma* species of *B. plebejum* group. Thanks to Axel O. Bachmann (in memorian), C. Magalhães (INPA), D. Pluot-Sigwalt (MNHN), G. Mejdalani (MN/UFRJ), Jorge L. Nessimian (DZRJ), P. Grootaert (ISNB), R. T. Schuh (AMNH), and R. W. Brooks (SEMC) for the loan of specimens and access to collections. We thank M. Felix for assisting in the preparation of this work. Fellowships from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). This work was sponsored in part by Conselho Nacional de Pesquisa (CNPq).

## 10. References

- Altekar G, Dwarkadas S, Huelsenbeck J, Ronquist F (2004) Parallel Metropolis coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. *Bioinformatics* 20(3): 407–415. <https://doi.org/10.1093/bioinformatics/btg427>
- Amorim DS (1997) Elementos básicos de Sistemática Filogenética (2a edição). Holos Editora e Sociedade Brasileira de Entomologia, Ribeirão Preto, 276 pp.
- Armbruster WS, Pélabon C, Hansen TF, Mulder CPH (2004) Floral integration, modularity, and accuracy. Distinguishing complex adaptations from genetic constraints. In: Pigliucci M, Preston K (Eds) *Phenotypic integration: studying the ecology and evolution of complex phenotypes*. Oxford University Press, New York, 23–49.
- Arnett-Jr. RH, Samuelson GA, Nishida GM (1993) *The insect and spider collections of the world* (2nd ed.). Sandhill Crane Press, Inc, Gainesville, 316 pp.
- Baele G, Lemey P, Bedford T, Rambaut A, Suchard MA, Alekseyenko AV (2012) Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Molecular Biology and Evolution* 29: 2157–2167. <https://doi.org/10.1093/molbev/mss084>
- Barão KR, Ferrari A, Grazia J (2020) Phylogenetic analysis of the *Euschistus* group (Hemiptera: Pentatomidae) suggests polyphyly of *Dichelops* Spinola, 1837 with the erection of *Diceraeus* Dallas, 1851, stat. rev. *Austral Entomology* 59(4): 770–783. <https://doi.org/10.1111/aen.12489>
- Baur H, Leuenberger C (2011) Analysis of ratios in multivariate morphometry. *Systematic Biology* 60(6): 813–825. <https://doi.org/10.1093/sysbio/syr061>
- Bergsten J (2005) A review of long-branch attraction. *Cladistics* 21(2): 163–193. <https://doi.org/10.1111/j.1096-0031.2005.00059.x>
- Betts CR (1986) The comparative morphology of the wings and axillae of selected Heteroptera. *Journal of Zoology* 1: 255–282. <https://doi.org/10.1111/j.1096-3642.1986.tb00639.x>
- Bremer K (1994) Branch support and tree stability. *Cladistics* 10: 295–304. <https://doi.org/10.1111/j.1096-0031.1994.tb00179.x>
- Brower AVZ, Schawaroch V (1996) Three steps of homology assessment. *Cladistics* 12: 265–272. <https://doi.org/10.1111/j.1096-0031.1996.tb00014.x>
- Brown JW, Parins-Fukuchi C, Stull GW, Vargas OM, Smith SA (2017) Bayesian and likelihood phylogenetic reconstructions of morpho-



- logical traits are not discordant when taking uncertainty into consideration: a comment on Puttick *et al.* Proceedings of the Royal Society B: Biological Sciences 284: 20170986. <https://doi.org/10.1098/rspb.2017.0986>
- Burnham KP, Anderson DR (2002) Model selection and multimodel inference: a practical information theoretic approach. Springer, New York, NY, 488 pp. <https://doi.org/10.1007/b97636>
- Clarke JA, Middleton KM (2008) Mosaicism, modules, and the evolution of birds: results from a Bayesian approach to the study of morphological evolution using discrete character data. Systematic Biology 57(2): 185–201. <https://doi.org/10.1080/10635150802022231>
- Cobben RH (1957) Beitrag zur Kenntnis der Uferwanzen (Hem. Het. Fam. Saldidae). Entomologische Berichten 17: 245–257.
- de Bivort BL, Clouse RM, 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. <https://doi.org/10.1111/j.1439-0469.2009.00562.x>
- de Pinna MCC (1991) Concepts and tests of homology in the cladistic paradigm. Cladistics 7: 367–394. <https://doi.org/10.1111/j.1096-0031.1991.tb00045.x>
- de Pinna MCC (1999) Species concepts and phylogenetics. Reviews in Fish Biology and Fisheries 9: 353–373. <https://www.doi.org/10.1023/A:1008911414399>
- de Queiroz A, Donoghue MJ, Kim J (1995) Separate versus combined analysis of phylogenetic evidence. Annual Review of Ecology, Evolution, and Systematics 26: 657–681. <https://doi.org/10.1146/annurev.es.26.110195.003301>
- Donoghue MJ, Ree RH (2000) Homoplasy and developmental constraints: a model and example from plants. American Zoologist 40: 759–769. <https://doi.org/10.1093/icb/40.5.759>
- Drummond AJ, Bouckaert RR (2015) Bayesian evolutionary analysis with BEAST. Cambridge University Press, Cambridge, 249 pp. <https://doi.org/10.1017/CBO9781139095112>
- Dupuis C (1955) Les génitalia des hémiptères hétéroptères (génitalia externes des deux sexes; voies ectodermiques femelles). Revue de la morphologie. Lexique de la nomenclature. Index bibliographique analytique. Mémoires du Muséum National d'Histoire Naturelle (série A, Zoologie) 6: 183–278.
- Escapa IH, Catalano SA (2013) Phylogenetic analysis of Araucariaceae: integrating molecules, morphology, and fossils. International Journal of Plant Sciences 174(8): 1153–1170. <https://doi.org/10.1086/672369>
- Estévez AL, Polhemus JT (2001) The small species of *Belostoma* (Heteroptera, Belostomatidae). 1. Key to species groups and a revision of the *denticolle* group. Iheringia 91: 151–158. <https://doi.org/10.1590/S0073-47212001000200021>
- Estévez AL, Polhemus JT (2007) The small species of *Belostoma* (Heteroptera, Belostomatidae): revision of the *plebejum* group. Revista de Biología Tropical 55: 147–155.
- Farris J (1970) Methods for computing Wagner trees. Systematic Zoology 19: 83–92. <https://doi.org/10.2307/2412028>
- Farris JS (1990) Phenetics in camouflage. Cladistics 6: 91–100. <https://doi.org/10.1111/j.1096-0031.1990.tb00528.x>
- Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39(4): 783–791. <https://doi.org/10.1111/j.1558-5646.1985.tb00420.x>
- Felsenstein J (2004) Inferring phylogenies. Sinauer Associates, Inc., Sunderland, Massachusetts, 664 pp.
- Fitch WM (1971) Towards defining the course of evolution: minimum change for a specific tree topology. Systematic Zoology 20: 406–416. <https://doi.org/10.2307/2412116>
- Fitch WM, Margoliash E (1967) Construction of phylogenetic trees. Science 155: 279–284. <https://doi.org/10.1126/science.155.3760.279>
- Fitzjohn RG, Maddison WP, Otto SP (2009) Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. Systematic Biology 58(6): 595–611. <https://doi.org/10.1093/sysbio/syp067>
- Forey PL, Kitching IJ (2000) Experiments in coding multistate characters. In: Scotland R, Pennington RT (Eds) Homology and Systematics, Special Volume Series 58. Systematics Association, London, 54–80.
- Garland T Jr. (1992) Rate tests for phenotypic evolution using phylogenetically independent contrasts. American Naturalist 140: 509–519. <https://doi.org/10.1086/285424>
- Gatesy J, O'Grady P, Baker RH (1999) Corroboration among data sets in simultaneous analysis: hidden support for phylogenetic relationships among higher level artiodactyl taxa. Cladistics 15: 271–313. <https://doi.org/10.1111/j.1096-0031.1999.tb00268.x>
- Gauterio TB, Ribeiro JRI, Robe LJ, Ferrari A (2021) Disentangling the taxonomy of *Belostoma* Latreille (Hemiptera, Heteroptera, Belostomatidae) with the aid of DNA barcoding approaches. Austral Entomology 2021: 1–11. <https://doi.org/10.1111/aen.12538>
- Geyer CJ (1991) Markov chain Monte Carlo maximum likelihood. Interface Foundation of North America. Computing Science and Statistics, Proceedings of the 23<sup>rd</sup> Symposium on the Interface: 156–163. <http://hdl.handle.net/11299/58440>.
- Giribet G (2003) Stability in phylogenetic formulations and its relationship to nodal support. Systematic Biology 52(4): 554–564. <https://doi.org/10.1080/10635150390223730>
- Goloboff PA, Carpenter JM, Arias JS, Esquivel DRM (2008) Weighting against homoplasy improves phylogenetic analysis of morphological datasets. Cladistics 24: 758–773. <https://doi.org/10.1111/j.1096-0031.2008.00209.x>
- Goloboff PA (1993) Estimating characters weights during tree search. Cladistics 9: 83–91. <https://doi.org/10.1111/j.1096-0031.1993.tb00209.x>
- Goloboff PA (2007) Calculating SPR distances between trees. Cladistics 23: 1–7. <https://doi.org/10.1111/j.1096-0031.2007.00189.x>
- Goloboff PA, Catalano SA (2016) TNT version 1.5, including a full implementation of phylogenetic morphometrics. Cladistics 32: 221–238. <https://doi.org/10.1111/cla.12160>
- Goloboff PA, Farris JS, Källersjö M, Oxelman B, Ramírez MJ, Szumik CA (2003) Improvements to resampling measures of group support. Cladistics 19: 324–332. <https://doi.org/10.1111/j.1096-0031.2003.tb00376.x>
- Goloboff PA, Mattoni CI, Quinteros AS (2006) Continuous characters analyzed as such. Cladistics 22: 589–601. <https://doi.org/10.1111/j.1096-0031.2006.00122.x>
- Goloboff P, Farris JS (2002) Methods for quick consensus estimation. Cladistics 17: S26–S34. <https://doi.org/10.1111/j.1096-0031.2001.tb00102.x>
- Gorb SN, Perez-Goodwyn PJ (2003) Wing-locking mechanisms in aquatic Heteroptera. Journal of Morphology 257: 127–146. <https://doi.org/10.1002/jmor.10070>
- Gutberlet RL Jr. (1998) The phylogenetic position of the Mexican Black-tailed Pitviper (Squamata: Viperidae: Crotalinae). Herpetologica 54(2): 184–206.

- Hamilton KGA (1981) Morphology and evolution of the rhynchotan head (Insecta: Hemiptera, Homoptera). The Canadian Entomologist 113(11): 953–974. <https://doi.org/10.4039/Ent113953-11>
- Hansen TF (1997) Stabilizing selection and the comparative analysis of adaptation. Evolution 51: 1341–1351. <https://doi.org/10.1111/j.1558-5646.1997.tb01457.x>
- Hansen TF, Martins EP (1996) Translating between microevolutionary process and macroevolutionary patterns: the correlation structure of interspecific data. Evolution 50: 1404–1417. <https://doi.org/10.1111/j.1558-5646.1996.tb03914.x>
- Harmon L (2019) Phylogenetic comparative methods: learning from trees. EcoEvoRxiv Preprints, Chicago, 234 pp. <https://doi.org/10.32942/osf.io/e3xnr>
- Hauser DL, Presch W (1991) The effect of ordered characters on phylogenetic reconstruction. Cladistics 7: 243–265. <https://doi.org/10.1111/j.1096-0031.1991.tb00037.x>
- Hennig W (1968) Elementos de una Sistemática Filogenética. EUDEBA, Buenos Aires, 224 pp.
- Höhna S (2014) Likelihood inference of non-constant diversification rates with incomplete taxon sampling. PLoS One 9: e84184. <https://doi.org/10.1371/journal.pone.0084184>
- Höhna S, Landis MJ, Heath TA, Boussau B, Lartillot N, Moore BR, Huelsenbeck JP, Ronquist F (2016) RevBayes: Bayesian phylogenetic inference using graphical models and an interactive model-specification language. Systematic Biology 65(4): 726–736. <https://doi.org/10.1093/sysbio/syw021>
- Höhna S, Stadler T, Ronquist F, Britton T (2011) Inferring speciation and extinction rates under different species sampling schemes. Molecular Biology and Evolution 28: 2577–2589. <https://doi.org/10.1093/molbev/msr095>
- Holder MT, Sukumaran J, Lewis PO (2008) A justification for reporting the majority-rule consensus tree in Bayesian phylogenetics. Systematic Biology 57: 814–821. <https://doi.org/10.1080/10635150802422308>
- Jeffreys H (1961) The theory of probability. Third Edition. Clarendon Press, Oxford, 447 pp.
- Kendall DG (1948) On the generalized “birth-and-death” process. Annals of Mathematical Statistics 19: 1–15. <https://doi.org/10.1214/aoms/1177730285>
- Kitching IJ, Forey PL, Humphries CJ, Williams DM (1998) The theory and practice of parsimony analysis. Second Edition. Oxford University Press, Oxford, 228 pp.
- Klingenberg CP (2015) Analyzing fluctuating asymmetry with geometric morphometrics: concepts, methods, and applications. Symmetry 7: 843–934. <https://doi.org/10.3390/sym7020843>
- Kluge AG, Farris JS (1969) Quantitative phyletic and the evolution of anurans. Systematic Zoology 18: 1–32. <https://doi.org/10.2307/2412407>
- Koch NM, Soto IM, Ramírez MJ (2014) First phylogenetic analysis of the family Neriidae (Diptera), with a study on issue of scaling continuous characters. Cladistics 31: 142–165. <https://doi.org/10.1111/cla.12084>
- Kühner MK, Yamato J (2015) Practical performance of tree comparison metrics. Systematic Biology 64(2): 205–214. <https://doi.org/10.1093/sysbio/syu085>
- Larsén O (1945) Das thorakale Skelettmuskelsystem der Heteropteren. Lunds Universitets Årsskrift 41: 1–96.
- Lartillot N, Philippe H (2006) Computing bayes factor using thermodynamic integration. Systematic Biology 55(2): 195–207. <https://doi.org/10.1080/10635150500433722>
- Lauck DR (1962) A monograph of the genus *Belostoma* (Hemiptera) Part I. Introduction to *B. dentatum* and *subspinosum* groups. Bulletin of the Chicago Academy of Sciences 11(3): 34–81.
- Lauck DR (1963) A monograph of the genus *Belostoma* (Hemiptera), Part II. Bulletin of the Chicago Academy of Sciences 11(4): 82–101.
- Lauck DR (1964) A monograph of the genus *Belostoma* (Hemiptera) Part III. *B. triangulum*, *bergi*, *minor*, *bifoveolatum*, and *flumineum* groups. Bulletin of the Chicago Academy of Sciences 11(5): 102–154.
- Lawing AM, Meik JM, Schargel WE (2008) Coding meristic characters for phylogenetic analysis: a comparison of step-matrix gap-weighting and generalized frequency coding. Systematic Biology 57(1): 167–173. <https://doi.org/10.1080/10635150801898938>
- Leaché AD, Banbury BL, Felsenstein J, Nieto-Montes de Oca A, Stamatakis A (2015) Short tree, long tree, right tree, wrong tree: new acquisition bias corrections for inferring SNP phylogenies. Systematic Biology 64(6): 1032–1047. <https://doi.org/10.1093/sysbio/syv053>
- Lewis PO (2001) A likelihood approach to estimating phylogeny from discrete morphological character data. Systematic Biology 50: 913–925. <https://doi.org/10.1080/106351501753462876>
- Maddison DR, Maddison WP (2005) MacClade 4: Analysis of phylogeny and character evolution. Version 4.08a. <http://macclade.org>
- Maddison WP, Maddison DR (2019) Mesquite: a modular system for evolutionary analysis. Version 3.61. <http://www.mesquiteproject.org>
- Mirande JM (2009) Weighted parsimony phylogeny of the family Characidae (Teleostei: Characiformes). Cladistics 25: 574–613. <https://doi.org/10.1111/j.1096-0031.2009.00262.x>
- Moreira FFF, Barbosa JF, Ribeiro JRI, Alecrim VP (2011) Checklist and distribution of semiaquatic and aquatic Heteroptera (Gerromorpha and Nepomorpha) occurring in Brazil. Zootaxa 2958: 1–74. <https://doi.org/10.11646/zootaxa.2958.1.1>
- Mounce RCP, Sansom R, Wills MA (2016) Sampling diverse characters improves phylogenies: craniodental and postcranial characters of vertebrates often imply different trees. Evolution 70(3): 666–686. <https://doi.org/10.1111/evo.12884>
- Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ (2015) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Molecular Biology and Evolution 32: 268–274. <https://doi.org/10.1093/molbev/msu300>
- Nieser N (1975) The water bugs (Heteroptera: Nepomorpha) of the Guyana Region. Studies on the Fauna of Suriname and Other Guyanas 16: 88–128. [https://doi.org/10.1007/978-94-017-7118-4\\_1](https://doi.org/10.1007/978-94-017-7118-4_1)
- Nieser N, Melo AL (1997) Os heterópteros aquáticos de Minas Gerais. Guia introdutório com chave de identificação para as espécies de Nepomorpha e Gerromorpha. Ed. UFMG, Belo Horizonte, 180 pp.
- Nixon KC, Carpenter JM (1993) On outgroups. Cladistics 9(4): 413–426. <https://doi.org/10.1111/j.1096-0031.1993.tb00234.x>
- O’Reilly JE, Donoghue PCJ (2018) The efficacy of consensus tree methods for summarizing phylogenetic relationships from a posterior sample of trees estimated from morphological data. Systematic Biology 67: 354–362. <https://doi.org/10.1093/sysbio/syx086>
- O’Reilly JE, Puttick MN, Pisani D, Donoghue PCJ (2018) Probabilistic methods surpass parsimony when assessing clade support in phylogenetic analyses of discrete morphological data. Palaeontology 61: 105–118. <https://doi.org/10.1111/pala.12330>
- Paradis E, Claude J, Strimmer K (2004) APE: analysis of phylogenetics and evolution in R language. Bioinformatics 20: 289–290. <https://doi.org/10.1093/bioinformatics/btg412>
- Parins-Fukuchi C (2018) Use of continuous traits can improve morphological phylogenetics. Systematic Biology 67(2): 328–339. <https://doi.org/10.1093/sysbio/syx072>

- Parsons MC (1964) The origin and development of the Hemipteran cranium. *Canadian Journal of Zoology* 42: 409–432. <https://doi.org/10.1139/z64-036>
- Parsons MC (1972) Respiratory significance of the thoracic and abdominal morphology of *Belostoma* and *Ranatra* (Insecta, Heteroptera). *Zeitschrift für Morphologie der Tiere* 73: 163–194. <https://doi.org/10.1007/BF00280774>
- Parsons MC (1974) The morphology and possible origin of the Hemipteran loral lobes. *Canadian Journal of Zoology* 52: 189–202. <https://doi.org/10.1139/z74-023>
- Pereira MH (1992) Avaliação da capacidade predatória e aspectos da biologia de *Belostoma anurum* e *B. plebejum* (Hemiptera, Belostomatidae), em laboratório. PhD Thesis, Universidade Federal de Minas Gerais, Minas Gerais, Brazil.
- Pereira MH, Melo AL, Pereira LH (1991) Laboratory rearing of *Belostoma micantulum* (Stål, 1858) (Hemiptera, Belostomatidae). *Revista Brasileira de Biologia* 51(3): 603–606.
- Perez-Goodwyn PJ, Gorb SN (2003) Attachment forces of the hemelytra-locking mechanisms in aquatic bugs (Heteroptera: Belostomatidae). *Journal of Insect Physiology* 49: 753–764. [https://doi.org/10.1016/S0022-1910\(03\)00112-4](https://doi.org/10.1016/S0022-1910(03)00112-4)
- Pimentel RA, Riggins R (1987) The nature of cladistic data. *Cladistics* 3: 201–209. <https://doi.org/10.1111/j.1096-0031.1987.tb00508.x>
- Polhemus JT, Polhemus DA (2008) Global diversity of true water bugs (Heteroptera; Insecta) in freshwater. *Hydrobiologia* 595: 379–91. <https://doi.org/10.1007/s10750-007-9033-1>
- Pollock DD, Bergman A, Feldman MW, Goldstein DB (1998) Microsatellite behavior with range constraints: parameter estimation and improved distance estimation for use in phylogenetic reconstruction. *Theoretical Population Biology* 53: 256–271. <https://doi.org/10.1006/tpbi.1998.1363>
- Price SA, Tavera JJ, Near TJ, Wainwright PC (2013) Elevated rates of morphological and functional diversification in reef-dwelling haemulid fishes. *Evolution* 67: 417–428. <https://doi.org/10.1111/j.1558-5646.2012.01773.x>
- Puttick MN, O'Reilly JE, Pisani D, Donoghue PCJ (2019) Probabilistic methods outperform parsimony in the phylogenetic analysis of data simulated without a probabilistic model. *Paleontology* 62 (part 1): 1–17. <https://doi.org/10.1111/pala.12388>
- Rae TC (1998) The logical basis for the use of continuous characters in phylogenetic systematics. *Cladistics* 14: 221–228. <https://doi.org/10.1111/j.1096-0031.1998.tb00335.x>
- Rannala B, Yang Z (1996) Probability distribution of molecular evolutionary trees: a new method of phylogenetic inference. *Journal of Molecular Evolution* 43: 304–311. <https://doi.org/10.1007/BF02-338839>
- Rannala B, Huelsenbeck JP, Yang Z, Nielsen R (1998) Taxon sampling and the accuracy of large phylogenies. *Systematic Biology* 47(4): 702–710. <https://doi.org/10.1080/106351598260680>
- Revell LJ, Collar DC (2009) Phylogenetic analysis of the evolutionary correlation using likelihood. *Evolution* 63(4): 1090–1100. <https://doi.org/10.1111/j.1558-5646.2009.00616.x>
- Ribeiro JRI (2000) Description of the male of *Belostoma foveolatum* and new records of *B. costalimai* and *B. stollii* (Heteroptera: Belostomatidae). *Entomological News* 111(3): 159–170.
- Ribeiro JRI (2007) A review of the species of *Belostoma* Latreille, 1807 (Hemiptera: Heteroptera: Belostomatidae) from the four southeastern Brazilian states. *Zootaxa* 1477(1): 1–70. <https://doi.org/10.11646/zootaxa.1497.1.1>
- Ribeiro JRI, Alecrim VP (2008) Duas novas espécies de *Belostoma* Latreille, 1807 (Hemiptera, Heteroptera: Belostomatidae) do grupo *plebejum* sensu Nieser, 1975. *Acta Amazonica* 38: 179–188. <https://doi.org/10.1590/S0044-59672008000100021>
- Ribeiro JRI, Estévez AL (2009) The small species of *Belostoma* Latreille (Heteroptera, Belostomatidae). III. A revision of the *oxyurum* group, with a new species from Brazil and description of the male of *B. noualhieri* Montandon. *Revista Brasileira de Entomologia* 53: 207–215. <https://doi.org/10.1590/S0085-56262009000200004>
- Ribeiro JRI, Ebong SEM, Le-Gall P, Guilbert E (2014) A taxonomic synopsis of *Limnogeton* Mayr, 1853 (Insecta: Hemiptera: Heteroptera: Belostomatidae). *Zootaxa* 3779(5): 573–584. <https://doi.org/10.11646/zootaxa.3779.5.7>
- Ribeiro JRI, Estévez AL, Moreira FFF, Guilbert E (2017) Revision of the *Belostoma dentatum* group sensu Nieser (Insecta, Heteroptera, Belostomatidae). *Zootaxa* 4276(2): 177–203. <https://doi.org/10.11646/zootaxa.4276.2.2>
- Ribeiro JRI, Ohba S, Pluot-Sigwalt D, Stefanello F, Bu W, Ebong SEM, Guilbert E (2018) Phylogenetic analysis and revision of subfamily classification of Belostomatidae genera (Insecta: Heteroptera: Nepomorpha). *Zoological Journal of the Linnean Society* 182(2): 319–359. <https://doi.org/10.1093/zoolinnean/zlx041>
- Robinson DF, Foulds LR (1979) Comparison of weighted labeled trees. *Lecture Notes in Mathematics* 748: 119–126. <https://doi.org/10.1007/BFb0102690>
- Robinson DF, Foulds LR (1981) Comparison of phylogenetic trees. *Mathematical Biosciences* 53: 131–147. [https://doi.org/10.1016/00-25-5564\(81\)90043-2](https://doi.org/10.1016/00-25-5564(81)90043-2)
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61(3): 539–542. <https://doi.org/10.1093/sysbio/sys029>
- Schuh RT (1979) Evolutionary trends in Heteroptera. Part II. Mouthpart-structures and feeding strategies. *Systematic Zoology* 28(4): 653–656. <https://doi.org/10.2307/sysbio/28.4.653>
- Schuh RT, Slater JA (1995) True bugs of the world (Hemiptera: Heteroptera): classification and natural history. Cornell University Press, New York, xii + 336 pp.
- Sereno PC (2007) Logical basis for morphological characters in phylogenetics. *Cladistics* 23: 565–587. <https://doi.org/10.1111/j.1096-0031.2007.00161.x>
- Silveira LFL, Roza AS, Vaz S, Mermudes JRM (2021) Description and phylogenetic analysis of a new firefly genus from the Atlantic Rainforest, with five new species and new combinations (Coleoptera: Lampyridae: Lampyrinae). *Arthropod Systematics & Phylogeny* 79: 115–150. <https://doi.org/10.3897/asp.79.e67185>
- Slowinski JB (1993) “Unordered” versus “ordered” characters. *Systematic Biology* 42: 155–165. <https://doi.org/10.1093/sysbio/42.2.155>
- Stål C (1860) Bidrag till Rio Janeiro - Traktens Hemipter-Fauna. *Kungliga Svenska Vetenskapsakademiens handlingar, Estocolmo* 2(7): 1–84.
- Stefanello F (2021) The new *Belostoma fittkaui* species group with supplemental descriptions of *B. fittkaui* De Carlo and *B. sayagoi* De Carlo. *Zootaxa* 4942(4): 583–591. <https://doi.org/10.11646/zootaxa.4942.4.6>
- Stefanello F, Bugs C, Stenert C, Maltchik L, Guilbert E, Ribeiro JRI (2018) Integration and modularity in the male genitalia and parameters of *Belostoma* species of *bifoveolatum* group sensu Lauck, 1962



- (Insecta, Heteroptera, Belostomatidae). *Zoologischer Anzeiger* 272: 45–64. <https://doi.org/10.1016/j.jcz.2017.11.013>
- Stefanello F, Ribeiro JRI, Estévez AL (2021) Revision of the *Belostoma triangulum* group (Insecta: Heteroptera: Belostomatidae). *Zootaxa* 4958(1): 95–102. <https://doi.org/10.11646/zootaxa.4958.1.8>
- Swofford DL, Begle DP (1993) User's manual for PAUP, Phylogenetic analysis using parsimony, version 3.1. Illinois Natural History Survey, Champaign, 280 pp.
- Thiele K (1993) The holy grail of the perfect character: the cladistic treatment of morphometric data. *Cladistics* 9: 275–304. <https://doi.org/10.1111/j.1096-0031.1993.tb00226.x>
- Vernygora OV, Simões TR, Campbell EO (2020) Evaluating the performance of probabilistic algorithms for phylogenetic analysis of big morphological datasets: a simulation study. *Systematic Biology* 69(6): 1088–1105. <https://doi.org/10.1093/sysbio/syaa020>
- Weber H (1930) *Biologie der Hemipteren. Eine Naturgeschichte der Schnabelkerfe*. Julius Springer, Berlin, 543 pp.
- Weiler LM, Ferrari A, Grazia J (2016) Phylogeny and biogeography of the South American subgenus *Euschistus* (*Lycipta*) Stål (Heteroptera: Pentatomidae: Carpororini). *Insect Systematics and Evolution* 47: 313–346. <https://doi.org/10.1163/1876312X-47032145>
- Wheeler WC (2012) *Systematics: a course of lectures*. Wiley-Blackwell, Oxford, 426 pp.
- Wiens JJ (1995) Polymorphic characters in phylogenetic systematics. *Systematic Biology* 44: 482–500. <https://doi.org/10.2307/2413656>
- Wiens JJ (1999) Polymorphism in systematics and comparative biology. *Annual Review of Ecology, Evolution, and Systematics* 30: 327–362. <https://doi.org/10.1146/annurev.ecolsys.30.1.327>
- Wiens JJ (2001) Character analysis in morphological phylogenetics: problems and solutions. *Systematic Biology* 50(5): 689–699. <https://doi.org/10.1080/106351501753328811>
- Wiens JJ, Hollingsworth BD (2000) War of the Iguanas: conflicting molecular and morphological phylogenies and long-branch attraction in Iguanid Lizards. *Systematic Biology* 49: 143–159. <https://doi.org/10.1080/10635150050207447>
- Wiley EO (1981) *Phylogenetics. The theory and practice of phylogenetic systematics*. John Wiley and Sons, New York, 439 pp.
- Wright AM (2019) A systematist's guide to estimating Bayesian phylogenies from morphological data. *Insect Systematics and Diversity* 3(3): 1–14. <https://doi.org/10.1093/isd/ixz006>
- Wright AM, Hillis DM (2014) Bayesian analysis using a simple likelihood model outperforms parsimony for estimation of phylogeny from discrete morphological data. *Plos One* 9: e109210. <https://doi.org/10.1371/journal.pone.0109210>
- Wright AM, Lloyd GT, Hillis DM (2016) Modeling character change heterogeneity in phylogenetic analyses of morphology through the use of priors. *Systematic Biology* 65: 602–611. <https://doi.org/10.1093/sysbio/syv122>
- Xie W, Lewis PO, Fan Y, Kuo L, Chen MH (2011) Improving marginal likelihood estimation for Bayesian phylogenetic model selection. *Systematic Biology* 60: 150–160. <https://doi.org/10.1093/sysbio/syq085>
- Yang Z (1994) Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *Journal of Molecular Evolution* 39(3): 306–314. <https://doi.org/10.1007/BF00160154>
- Yeates D (1992). Why remove autapomorphies? *Cladistics* 8: 387–389. <https://doi.org/10.1111/j.1096-0031.1992.tb00080.x>
- Yoshizawa K, Saigusa T (2001) Phylogenetic analysis of paraneopteran orders (Insecta: Neoptera) based on forewing base structure, with comments on monophyly of Auchenorrhyncha (Hemiptera). *Systematic Entomology* 26: 1–13. <https://doi.org/10.1046/j.1365-3113.2001.00133.x>
- Yule GU (1925) A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F. R. S. *Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character* 213: 21–87. <https://doi.org/10.1098/rstb.1925.0002>

## Supplementary material 1

### Tables S1–S10

**Authors:** Ribeiro JRI, Ferrari A (2023)

**Data type:** .xlsx

**Explanation note:** **Table S1.** Taxonomic overview of the representatives of *Belostoma plebejum* group. – **Table S2.** Studied species for the phylogenetic analysis of the *Belostoma plebejum* group sensu Nieser (1975). – **Table S3.** Mean values of measurements (in mm) and relations used as such (continuous dataset). – **Table S4.** Distribution of discrete character states (42 characters) among terminal taxa. – **Table S5.** Kref and fit (adjusted homoplasy) of the trees under IW (analysis three). – **Table S6.** Kref and fit (adjusted homoplasy) of the trees under IW (analysis six). – **Table S7.** Estimated marginal likelihoods for different models for the morphological dataset of the *Belostoma plebejum* phylogeny. – **Table S8.** Summary of 127 different combinations of continuous characters (from one-to-one combinations to seven-to-seven combinations). – **Table S9.** Similarity of the trees as measured using Robinson-Foulds absolute symmetric distance. – **Table S10.** Estimated marginal likelihoods for different combinations of continuous character evolution.

**Copyright notice:** This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

**Link:** <https://doi.org/10.3897/asp.81.e87378.suppl1>

## Supplementary material 2

### Figure S1

**Authors:** Ribeiro JRI, Ferrari A (2023)

**Data type:** .pdf

**Explanation note:** Phylogenetic trees based on maximum parsimony, maximum likelihood, and Bayesian inferences.

**Copyright notice:** This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

**Link:** <https://doi.org/10.3897/asp.81.e87378.suppl2>

## Supplementary material 3

### List of characters

**Authors:** Ribeiro JRI, Ferrari A (2023)

**Data type:** .docx

**Explanation note:** Morphological characters and states for the phylogenetic analysis of the *Belostoma plebejum* group sensu Nieser (1975).

**Copyright notice:** This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

**Link:** <https://doi.org/10.3897/asp.81.e87378.suppl3>