Cladistics
The Willi Hennig Society

First phylogenetic analysis of the family Neriidae (Diptera), with a study on the issue of scaling continuous characters

Journal:	*Cladistics*
Manuscript ID:	CLA-14-01-0715
Manuscript Type:	Article
Date Submitted by the Author:	09-Jan-2014
Complete List of Authors:	Mongiardino Koch, Nicolás; Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Departamento de Ecología, Genética y Evolución Soto, Ignacio; Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Departamento de Ecología, Genética y Evolución Ramírez, Martín; Museo Argentino de Ciencias Naturales (MACN) – CONICET,
Keywords:	Cladistics, Entomology, Systematics, Phylogeny, Parsimony < Methodology, Evolution

First phylogenetic analysis of the family Neriidae (Diptera), with a study on the issue of scaling continuous characters

Mongiardino Koch, Nicolás [1,*], Ignacio M. Soto [1, 2] and Martín J. Ramírez [3]

[1] Departamento de Ecología, Genética y Evolución. Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires. Ciudad Universitaria, Pabellón II (C1428 EHA). Buenos Aires. Argentina.
[2] Instituto de Ecología, Genética y Evolución de Buenos Aires (IEGEBA) – CONICET.
[3] Museo Argentino de Ciencias Naturales – CONICET, Buenos Aires, Argentina.

* Corresponding author. E-mail: nmongiar@ege.fcen.uba.ar

Running title: Neriidae phylogeny and the scaling of continuous characters

Abstract

Neriidae are a small family of acalyptratae flies, mostly distributed along the tropics. Very little is known about their biology, and the evolutionary relationships among species have never been evaluated. We perform the first comprehensive phylogenetic analysis of the family, including 48 species from all biogeographic regions inhabited, as well as five species of Micropezidae and one Cypselosomatidae as outgroups. We build a morphological data matrix of 194 characters, including 72 continuous characters. We first explore ways to deal with the issue of scaling continuous characters, including rescaling ranges to unity and using implied weighting. We find that both strategies result in very different phylogenetic hypotheses, and that implied weighting reduces only partially the issue of scaling. Furthermore, using implied weighting after rescaling characters improves the congruence between partitions and results in higher values of group support. With respect to the Neriidae, we confirm the monophyly of the family and of most its genera, although we do not obtain any of the currently accepted supra-generic groups. We propose to restrict the *Eoneria* and *Nerius* groups exclusively to the Neotropical fauna, and synonymyze *Glyphidops* subgenus *Oncopsia* Enderlein with *Glyphidops* subgenus *Glyphidops* Enderlein, eliminating the subgenera of the latter. This revised phylogeny presents a striking biogeographic consistency, and shows that previous main divisions of the family were based on events of convergence.

Table of contents

Introduction

Neriidae are a small family of medium to large size acalyptratae flies, with a general morphology characterized by relatively elongated bodies and long and slender legs, generally ornamented with two rows of ventral spines (Aczél, 1951; Carvalho-Filho and Esposito, 2008; Sepúlveda *et al.*, 2013a). Neriid flies can be readily distinguished from all other Acalyptratae by the elongated porrect antennae with an apical to subapical arista, and a pedicel divided into a body and a finger-like projection that connects with the median region of the first flagellomere (Aczél, 1961; Steyskal, 1968; Buck, 2010). Despite being present in all biogeographic regions, most of the family's diversity is concentrated along the tropics (Steyskal, 1968; 1987). So far, 110 species have been described and placed in 19 genera (Sepúlveda *et al.,* 2013a), with almost two-thirds of this diversity occurring in the New World (Eberhard, 1998). Among the American species, three occur in the desertic regions of southwestern United States (Mangan and Baldwin, 1986; Steyskal, 1987), whereas the vast majority dwells throughout Central and South America (Aczél, 1961; Steyskal, 1968). The rest of the species are mainly distributed along the Australian-Oriental (Aczél, 1954a; Steyskal, 1977; Pitkin, 1989) and Afrotropical ecozones (Aczél, 1954b; Steyskal, 1980; Barraclough, 1993a).

Neriidae were historically considered by some as an independent family closely related to the Micropezidae (Hendel, 1922; Cresson, 1930), while others regarded them as a subfamily included within the Micropezidae (Enderlein, 1922; Hennig, 1934; 1936; 1937). Consensus on the placement of this clade at the family level was only reached after the thorough taxonomic work done by Martín L. Aczél, including several monographs on the diversity of neriid flies from all inhabited continents (Aczél, 1951; 1954a; 1954b; 1954c; 1955a; 1955b; 1955c; 1959; 1961). Neriidae are at present recognized as one of the families included within the superfamily Nerioidea, the monophyly of which has been retained in several morphological and molecular higher-level phylogenetic analyses (J. F. McAlpine, 1989; Yeates and Wiegmann, 2005; Yeates *et al.*, 2007; Wiegmann *et al.*, 2011). Several synapomorphies are taken to support the monophyly of the Nerioidea, including the peculiar morphology of the elongated male and female genitalia and the desclerotized lower region of the face (Aczél, 1951; J. F. McAlpine, 1989; D. K. McAlpine, 1996; Yeates *et al.*, 2007; Buck and McAlpine, 2010). However, the internal taxonomic and systematic organization of the clade is still highly controversial. The most conservative view recognizes only three families within the Nerioidea (J. F. McAlpine, 1989): the extremely diverse Micropezidae (often subdivided into several subfamilies, see D. K. McAlpine, 1974; Marshall, 2010), and the less diverse sister taxa Cypselosomatidae and Neriidae. The monophyly of the Cypselosomatidae + Neriidae clade was originally proposed by J. F. McAlpine (1989), and has found subsequent confirmation in a recent molecular phylogeny (Wiegmann *et al.*, 2011). Other authors have proposed to separate Pseudopomyzidae from Cypselosomatidae and to establish this group as a fourth family (D. K. McAlpine, 1966;

Shatalkin, 1994), although many have favored a subfamiliar rank for the two (Griffiths, 1972; Prado, 1984; J. F. McAlpine, 1987; 1989). D. K. McAlpine (1996) rejected this last stance arguing that the similarities between both clades is due to the retention of symplesiomorphies, yet consensus on their position is still lacking. D. K. McAlpine (1996) has also transferred the Megamerinidae from the Diopsoidea into the Nerioidea, but others have not followed this proposal (Buck, 2010).

The modern systematic structure of the Neriidae was mainly erected by Aczél (1954a; 1961), although it was Enderlein (1922) who first defined some of the suprageneric clades that remain currently valid. As a result, the family is divided in two subfamilies: the Neriinae, characterized by the presence of antennal sockets (also called antennal bases), formed by the protruding and more or less inflated frontal region of the upper face (also referred to as mesofacial plate), into which the antennae are inserted; and the basal Telostylinae, which lack these antennal bases. However, much confusion has subsequently arisen from this division, and many authors doubt that they constitute monophyletic groups (Pitkin, 1989; Barraclough, 1993a; Buck, 2010). The entirety of the American fauna belongs to the subfamily Neriinae, which Aczél (1961) further subdivided into two groups: the *Nerius*-group, among which the dorsal region of the antennal bases is polished and shiny; and the *Eoneria*-group, in which the antennal bases are dull, and have at most a faint greasy luster. Furthermore, some authors (Aczél, 1961; Buck and Marshall, 2004; Buck, 2010; Sepúlveda *et al.*, 2013a) have partially discussed morphological similarities between the Neotropical genera *Longina* Wiedemann, *Cerantichir* Enderlein and *Odontoloxozus* Enderlein, although no author has gone as far as to propose these genera as constituting a monophyletic group, and no other phylogenetic hypothesis has been put forward. A summary of all these proposed relationships can be found in Fig. 1.

Ever since the mid 20th century, few taxonomic works have been published on the family. A handful of papers have dealt with the description of some new species (Mangan and Baldwin, 1986; Buck and Marshall, 2004; Sepúlveda *et al.*, 2013b), but the systematic relationships within the family have never been reevaluated nor tested using matrix-based phylogenetic methods. The family does present some peculiarities that have been interpreted by some as impediments towards a morphological evaluation of the relationships among genera. For example, the male genitalia is extremely conserved, and basically useless to determine evolutionary relationships (Buck, 2010). On the other hand, the distribution of setae (chaetotaxy), which has been routinely used with phylogenetic purposes among Diptera (McAlpine, 1987; Simpson *et al.*, 1999; Lambkin *et al.*, 2013) is outstandingly variable within specific limits among neriids (Aczél, 1951; Barraclough, 1993a; Buck, 2010), a feature that has even been experimentally addressed (Bonduriansky, 2009). The confusion that has arisen from such intra-specific variation is most surely epitomized by Steyskal's (1965) synonymization of the genera *Antillonerius* Hennig and *Imrenerius* Aczél due to the presence in the same specimen of "(...) a well-developed

anterior *ntpl* [notopleural] on one side and a barely distinguishable one on the other", a character that not only represented the main difference distinguishing both genera, but that was also considered by Aczél (1961) as having a high taxonomic importance within the family. These and other issues have led some authors to even doubt the value of cladistic enquiry altogether (see D. K. McAlpine, 1996 for such an argumentation concerning the phylogenetic relationships among the Nerioidea).

Nonetheless, the impressive taxonomic legacy of M. L. Aczél (summarized by Mello, 2010), including the description or redescription of almost half of the known species of neriids, may prove to be a treasure vault for phylogenetic analysis. Taxonomic descriptions generally harbor an outstanding amount of morphological data. This information has a quite direct correlation with the one included in morphological matrices used in cladistic analyses (Winston, 1999), making it plausible that a translation of these descriptions into the codified nature of data matrices will result in valid phylogenetic analyses. In particular, Aczél's taxonomic descriptions are extremely rich in anatomical details, including the registry of pigmentation patterns, chaetotaxy, general morphology and more than 40 body measurements (generally expressed employing ranges of variation) for both male and female specimens (see Aczél, 1959; 1961 for a few examples). The key to a new and revised phylogeny of the neriids may be hidden among this precise record.

One type of character that Aczél regularly insisted in using while delimiting groups within the Neriidae are body measurements and proportions. In his revision of American Neriidae (Aczél, 1961) he classified species according to body size, shape of the eyes, length of postcranium, proportions of the legs, and length of the male and female postabdomen. Such continuous characters have been historically neglected from cladistic analysis both due to theoretical and practical issues concerning their implementation. Objections against their use have included concerns regarding the existence of homologies in quantitative characters (Pimentel and Riggins, 1987), their nature as phenetic data (Cranston and Humphries, 1988) and the unavoidable arbitrariness involved in the methods used to discretize them (Archie, 1985; Crisp and Weston, 1987; Felsenstein, 1988; Farris, 1990). Much of the theoretical objections have since been dealt with (Chappill, 1989; Rae, 1998), and many authors have argued that the distinction between discrete and continuous characters is actually just a matter of degree (Stevens, 1991; Gift and Stevens, 1997; Wiens, 2001; MacLeod, 2002). A new reappraisal of quantitative data has followed after Goloboff *et al.* (2006) implemented the treatment of continuous characters as such in the software TNT (Goloboff *et al.*, 2008b). Since then, factual evidence that continuous characters carry useful phylogenetic information has accumulated (Goloboff *et al.*, 2006; Hornung-Leoni and Sosa, 2008; Pereyra and Mound, 2009; de Bivort *et al.*, 2010; Escapa and Catalano, 2013).

Despite proving useful in many cases, some methodological caveats concerning the implementation of continuous characters in parsimony analysis are still poorly explored.

One of such issues is scaling, which has been referred to as "one of the most pervasive problems in the analysis of continuous characters" (Goloboff *et al.*, 2006). As the same authors argued, although it is clear that within a single continuous character a change in state between two species should be proportional to the magnitude of the difference they exhibit, the problem arises when considering the cost of transformations among different characters that potentially vary in widely different magnitudes. When this occurs, characters expressed in higher orders of magnitude will typically dominate the analysis (Thiele and Ladiges, 1988; Wiens, 2001; Goloboff *et al.*, 2006; Baur and Leunberger, 2011), resulting in an unbalanced character influence towards determining the optimal topology. Two methodological approaches have been therefore proposed to reduce this phenomenon. The first one is the practice of rescaling continuous characters, standardizing their ranges of variation to a common magnitude, usually unity. This method has been extensively used, both to standardize discretized (Colles, 1980; Cranston and Humphries, 1988; Thiele and Ladiges, 1988; Vargas *et al.*, 2010) and non-discretized continuous data (Abdala and Juarez Heredia, 2013; Escapa and Catalano, 2013), although some have objected to this practice on different grounds (Mikevich and Farris, 1981; Farris, 1990; Goloboff *et al.*, 2006). The second approach was proposed by Goloboff *et al.* (2006) who argued that the use of implied weighting against homoplasy (Goloboff, 1993) may be a way to reduce the issue of scaling. Since characters measured on larger scales will most likely have a higher amount of homoplasy than characters measured in smaller scales, the former would receive lower implied weights and vice versa, possibly balancing the overall influence of the different characters. This may reduce the effect of the magnitude of continuous characters on the phylogenetic hypothesis while at the same time elegantly eluding the problem of determining an appropriate scaling factor (Goloboff *et al.*, 2006), and has been recently adopted by some (Mannion *et al.*, 2013). Although the argument is logical, no empirical evidence has been presented to support such claim, and it is not clear whether implied weighting successfully deals with the issue of scaling, therefore eliminating the dominance of large characters. Furthermore, the effect that these two strategies have on the resulting phylogenetic hypothesis has never been tested.

In the present work, we pursue a double objective. First, we analyze different strategies for the use of continuous characters under parsimony, including an analysis of the effects of rescaling and implied weighting. Second, we apply such insights to develop the first phylogenetic study of the Neriidae. For this, we built a 194 character matrix, which includes 72 continuous characters. We include in the analysis 48 species of neriids, with representatives from 14 out of 19 valid genera and from all biogeographic regions inhabited. Also included are 5 species of Micropezidae and one species of Cypselosomatidae, used as outgroups. Our results show that rescaling and applying implied weight to continuous characters result in completely different phylogenetic hypothesis, and that implied weights reduces, yet does not eliminate, the issues of scaling. Regarding the phylogenetic relationships among the Neriidae, we do not obtain many of the classical

subdivisions of the family as monophyletic groups, and therefore discuss the evolutionary and biogeographic history of the family in light of a revised phylogeny.

Materials and methods

Taxon sampling

Fifty-four species are included in the present analysis. The ingroup consists of 48 species of Neriidae, representing 44% of the described species. This sampling includes representatives of 14 out of 19 currently valid genera (the remaining five genera are all monotypic), and contains representatives of all major biogeographic regions inhabited. With respect to the outgroup, five species of Micropezidae and one species of Cypselosomatidae were included, therefore testing the monophyly of the Neriidae with respect to the other families conforming the superfamily Nerioidea *sensu* J. F. McAlpine (1989). The five micropezids included were chosen in order to incorporate a wide taxonomic and morphological diversity, and include two representatives of the Taeniapterinae, two Micropezinae and one Eurybatinae. All species incorporated in the analysis can be found in Table 1.

Character sampling

Specimens from 26 of the 54 species included in the study were revised at the collection of the Instituto Superior de Entomología, Fundación Miguel Lillo, Tucumán, Argentina (see Appendix S2 for a list of revised specimens). These include many of the original material deposited by Aczél. For the rest of the species, character states were recorded from taxonomic descriptions. Our sampling resulted in 122 discrete characters that describe the morphology of adult organisms, since immature stages have been described only for a couple of species (Steyskal, 1968). A description of these can be found in Appendix S1. Of these, 77 were coded as binary and 45 as multistate, of which 35 were taken as ordered. Furthermore, a total of 42 measurements were taken for all revised specimens, either by directly measuring them with an Olympus SZ4045 binocular microscope with an ocular micrometer, or from photographs of the specimens taken with an Olympus U-CMAD3 (Infinity 1) digital camera attached to an Olympus SZX7 binocular microscope, and afterwards digitally measuring body features with the software tpsDig2 (Rohlf 2010). For species included in the analysis that were not revised, the same measurements were annotated from available taxonomic descriptions. These were later combined into 72 continuous characters, all of which (except for character 1 and 2: male and female body length, respectively) take the shape of simple ratios. This practice,

although subject to some debate (Atchley *et al.*, 1976; Corrucini, 1977; Albrecht, 1978; Atchley, 1978; Hills, 1978) is extremely common both in morphometric taxonomy and phylogenetic analyses (Baur and Leuenberger, 2011; see de Bivort *et al.*, 2010; Lopardo *et al.*, 2011; Mannion *et al.*, 2013 for recent phylogenetic analysis that incorporate characters expressed as ratios). The objective behind this practice was either to reduce the effect of body size in morphological measurements (e.g.: chars. 3, 35, etc.), or to numerically represent shapes, proportions or relative positions (e.g.: chars. 5, 22, 49, respectively). From the 72 continuous characters, 5 described the male and female genitalia (chars. 68 and henceforth), and 3 were coded using male and female data indistinctively, due to having access to fragmentary data relating to sexual differences (chars. marked with an asterisk in Appendix S1). The remaining ratios were initially built using male and female measurements separately, resulting in 64 (32 pairs) of sex-specific characters.

Analysis of continuous characters

The objective of this section was to study the behavior of quantitative (continuous) characters under different strategies of analysis. Since no phylogenetic conclusions are to be taken from the results, the duplication of phylogenetic information through the use of covarying characters was not considered an issue, and is dealt with latter (see Identification of sexually dimorphic characters). Therefore, the entire set of 72 continuous characters was retained and used without testing their independence. Four strategies for the treatment of continuous characters were explored, involving: the use of continuous data without treatment (strategy A), applying implied weights (strategy B), rescaling (strategy C), and rescaling + applying implied weights (strategy D). When using implied weights, a concavity constant (k) of 6 was always used. When rescaling, the range of all characters was standardized to unity. For this and all subsequent analysis the Willi Hennig Society version of the program TNT (Goloboff *et al.*, 2008b) and the software Statistica (Statsoft 2001) were used for phylogenetic and statistical enquiry, respectively.

On the first place, we aimed at proving Goloboff *et al.*'s (2006) assertion that the fit assigned to a given continuous character when using implied weighting negatively correlates with the scale of such character. Therefore, we searched for the most parsimonious tree (henceforth MPT) supported by the continuous data partition under implied weights, with (strategy D) and without (strategy B) rescaling the data. We used a driven search (Goloboff, 2002; Giribet, 2007), starting with five random addition sequences with SPR and TBR branch swapping followed by ratcheting (Nixon, 1999), sectorial searches, tree drifting and fusing (Goloboff, 1999), until minimum length was found 10 times. Unless otherwise stated, this strategy was maintained in all subsequent tree searches. A single MPT was found for both strategies, and the fit of all characters in both trees was recorded as the complement of the value obtained with the function `fit` (actually the

"distortion" of a character). These values were used in a regression analysis using the range of variation of each character (maximum – minimum, a proxy for a character's magnitude) as the independent variable.

Afterwards, the MPTs for the four strategies were topologically compared in a pair-wise manner, in order to study the impact that different strategies had on the resulting phylogenetic hypothesis. For the comparisons we used the number of taxa in the agreement subtree (Gordon, 1980; Eulenstein *et al.*, 2004), the number of internal nodes in the strict consensus (Mickevich, 1978; Swofford, 1991), and the weighted and unweighted SPR distances (Goloboff, 2007) as measures of topological similarity. The weighted SPR distance has been considered a superior measure of topological similarity (Goloboff, 2007), since it measures not only the number of SPR moves needed to transform one tree into the other, but it also weights each move by its distance (*i.e.*, the number of nodes (n) separating the subtree's location before and after the movement). The weight assigned to each move is then $n / (n + j)$, and a value of 3 was always used for j. For ease of comparison, all measures except for the weighted SPR distance were standardized so that the maximum possible tree similarity obtained using each of them was equal to 1 (since a higher value of SPR distance implies a higher topological difference, rather than a higher resemblance as in the other employed measures, the SPR derived similarity was used instead).

The compatibility between the discrete and continuous character partitions was then explored under the four strategies (when using implied weights, both data partitions were weighted). An incongruence length difference test (ILD, Farris *et al.*, 1995) was used to study the incongruence between partitions, and weighted SPR distances were used to compare the MP trees supported by both partitions. Furthermore, absolute group frequencies under jackknifing (Farris *et al.*, 1996) were estimated for the MPTs of the entire dataset under all strategies, using 500 pseudoreplicate datasets by eliminating characters under $p = 36$ and analyzing each one with 10 runs of TBR + ratcheting with 10 iterations each. Higher mean node support values were taken to represent that a particular strategy of analysis resulted in a higher congruence between characters (as discussed by Goloboff, 1997; Ramírez, 2003; Goloboff *et al.*, 2008a).

We also designed an experiment to test the validity of Goloboff *et al.*'s (2006) second assertion relating to the issue of scaling, *i.e.* that implied weighting balances the overall influence of continuous characters due to the negative correlation between a characters magnitude and its fit. We ordered continuous characters with respect to their magnitude (once again, using their range as an estimate of such variable), and built two subsets of characters by selecting the 18 smallest and 18 largest of them (this number was chosen since it represents the top and bottom quartile). Defined in such a way, the smallest character within the "large" subset had a range 3.2 times larger than that of the largest one in the "small" subset. The MPTs for these two subsets of characters were searched and compared using the weighted SPR distance to the MPT supported by the complete

continuous dataset. A strategy that guarantees a balanced influence of all characters should result in a more or less similar distance from the trees supported by both subsets to the tree of the entire continuous partition. On the contrary, under a strategy where the magnitude of a character determines its influence over the optimal topology, the tree supported by the largest characters should be much more similar to the one supported by the entire partition than the one supported by the smallest characters, resulting in asymmetrical SPR distances. This procedure was repeated for all four strategies. To further explore the implications of the values obtained, we built null distributions of SPR distances by generating 1000 random subsets of 18 continuous character for each strategy, searching for the MPTs for every subset (using 20 RAS + TBR and holding up to 10 trees of the minimum length), and calculating the weighted SPR distance to the MPT of the entire continuous partition (the script used can be found in Appendix S3). Since many subsets supported more than one optimal tree, we calculated the SPR distance using all of them and retained, for each subset, the smallest distance. The resulting value can be interpreted as the "importance" of these subsets in determining the optimal tree of the entire continuous partition, in a similar way as discussed by DeGusta (2004) for individual characters. Therefore, the shortest the SPR distance, the higher the influence on the topology supported by the entire continuous dataset.

Since the calculation of SPR distances is an NP-complete problem (Bordewich and Semple, 2005), the algorithm implemented in TNT is heuristic and depends on two parameters: the number of replications and the number of stratifications (Goloboff, 2007). Given that SPR distances were intensely used, we first performed a tuning analysis to define appropriate values for those two parameters (the script used can be found in Appendix S3). One hundred Wagner trees were created and coupled into 50 pairs, and the SPR distance between them was explored using different combinations of parameters. We first defined a broad and less intense search, calculating the distance between trees for all combinations of number of replications between 1,000 and 20,000 (with steps every 1,000) and number of stratifications between 0 and 75 (with steps every 5). No major change in the resulting distances was found using stratification values above 30. On the other hand, since computational time increments with an increase in the number of replicates, we found that a number of 13,000 replicates guaranteed finding the shortest distance for more than 95% of tree pairs, being a good compromise between computational effort and accuracy. A second, more intense yet restricted search (number of replicates: 1,000 to 13,000, steps of 1,000; number of stratifications: 0 to 30, steps of 1) showed that the shortest path between pairs of trees was, on average, attained using 13,000 replicates and 18 stratifications (Fig. 2). These parameters were therefore used for all SPR distance calculations

Outgroup definition

Historically, the Neriidae were considered to be closely related with the Micropezidae (Aczél, 1951), even being classified as a subfamily included within the later (Hennig, 1936; 1937; Griffith, 1972). After their status as an independent family was recognized, some authors pointed out the existence of morphological similarities between the Neriidae and the Cypselosomatidae (D. K. McAlpine, 1966; 1974), with latter analyses confirming such clade to be monophyletic (J. F. McAlpine, 1989; Wiegmann *et al.*, 2011). Therefore, we rooted our trees with *Taeniaptera annulata*, a micropezid of the subfamily Taeniapterinae, which have been considered to be morphologically very dissimilar with respect to Neriidae by Hennig (1937), Aczél (1951) and D. K. McAlpine (1974).

Identification of sexually dimorphic continuous characters

So far, all continuous characters were *a priori* taken to be independent, including the 32 pairs of male-female characters. The codification of male and female data in separate characters may be useful in many occasions, since by reducing the ranges of variation, this practice results in less overlapping and greater informativeness of each of the resulting characters. Furthermore, it allows a more detailed study of the evolution of sexual dimorphism (Hormiga *et al.*, 2000). However, these advantages are only true if male and female characters are effectively independent, otherwise their codification as separate characters only results in the inclusion of redundant phylogenetic information in the analysis. Therefore, we tested whether our 32 pairs of male-female characters presented differences in their phylogenetic information. This was done following a strategy similar to that of de Bivort *et al.* (2010). Regression analyses were performed using male and female characters corresponding to the same morphological structure. Since some characters were expressed using ranges of variation and others only as single data points, in the first case the mean between the minimum and maximum values was used. In case a species presented missing data for at least one of both characters it was consequently excluded from that particular regression. Two characters were then considered to be dependent when they showed a significant linear regression with a complete absence of outliers. Since the objective of the regression analysis is to study whether a sex-specific character confidently predicts the value of that same character for the opposite sex (therefore showing a lack of independence among all taxa under study), the presence of a single outlier is evidence that, at least for one species, the same attribute is providing different phylogenetic information for males and females (de Bivort *et al.*, 2010). Therefore, for each regression, the residue of each data point was annotated and divided by the corresponding expected value. This resulted in a number that expressed the magnitude of the difference between observed and expected values for each species as a fraction of the expected value. Three nested, progressively stricter confidence intervals were considered, according to which characters were retained as dimorphic only if there was at least a single species for which the observed value differed from the expected one by a 20, 30 or 50% of the latter (Fig. 3). Otherwise, in

the absence of outliers, male and female characters were collapsed into a single character, whose range was defined by using the smallest value of both as a lower limit and the largest one as the upper limit (the script used can be found in Appendix S3). This resulted in the retention as separate characters of approximately 77, 40 and 27% of the 32 original pairs as the criterion became stricter (see Fig. 4c). This analysis included only male-female pairs of identical characters. Different morphological measures were considered independent without further *a priori* confirmation (in a fashion similar as for example Fink and Zelditch, 1995; Strait & Grine, 2004).

Cladistic analysis

Phylogenetic analysis was undertaken under strategy D, that is, rescaling continuous data to unit range and using implied weighting (see below). Trees were searched using constant of concavity *k* ranging from 1 to 10. This resulted in 10 ways to weight characters, which coupled with the 3 male-female collapse criteria, determined 30 different forms to analyze the data. All of these resulted in a single MPT. A majority rule consensus (cut-off 50%) was employed to perform a sensitivity analysis and identify those groups that were recovered under most conditions of analysis. The phylogenetic hypothesis proposed was chosen by calculating the majority consensus frequency for all groups present in each one of the 30 trees, averaging the values, and looking for the tree with higher mean group frequencies. This hypothesis is therefore the one that combines the groups most frequently present in the entire 30 tree set. As measures of support we used absolute frequencies of jackknife (parameters as defined above, but 1000 pseudoreplicates) and Bremer support (Bremer, 1988; 1994). For the last one, a heuristic calculation was done by searching trees that were increasingly suboptimal by 0.05 units of fit (under $k = 6$, the same value employed for the analysis of continuous characters). Tree search was done by TBR swapping, and up to 1000 trees for each round were retained. Search continued until reaching trees suboptimal by 1.35 units of fit, the point at which all nodes had been contradicted. Bremer supports were then calculated from the 27000 existing suboptimal trees, and plotted as units of fit x 100. TBR swapping of trees suboptimal by 0.05 units of fit produced only 933 trees, not overflowing tree space; hence values below 5 are probably exact.

Results

Analysis of continuous characters

We found a significant and negative correlation ($p < 0.0001$) between the fit assigned by implied weights and a character's range (Fig. 4b). The regression analysis determined that in fact, implied weights is almost exclusively weighting against a character's magnitude, with a value of R^2 of 0.76. The strength of this correlation was not dependent on the few extremely large characters, being equally strong when considering only characters with ranges < 1 (Fig. 4a, see regression values at the epigraph). When applying implied weights after rescaling, the correlation was lost ($p = 0.78$), and characters with similar original scales were now assigned widely different fit values (Fig. 4).

Therefore, implied weighting is in fact dealing with the magnitude of continuous characters, as originally stated by Goloboff *et al.* (2006). To prove whether the resulting effect was similar to rescaling the data, we performed several tests of topological differences between the MPTs obtained by each of the four strategies under analysis. We found that treating the data (strategies B, C and D) resulted in phylogenetic hypothesis that were all profoundly different from the one obtained by using the continuous dataset without treatment (strategy A), as shown by all measures of topological similarity employed (Fig. 5). When topologically comparing the MPTs of treated vs. untreated data, both the number of nodes in the strict consensus and the number of taxa in the agreement subtree were always between 0.2 and 0.3 of the value corresponding to an absolute congruence, while tree similarity derived from unweighted SPR distances ranged from 0.49 to 0.58. This sharp topological difference could be pointing out that both rescaling and implied weighting are eliminating, or at least strongly reducing, the dominance of large characters on the optimal topology, a phenomenon that otherwise characterizes the use of continuous characters without treatment. In fact, the mean consistency index (Kluge and Farris, 1969) for the two largest characters (male and female body size, chars. 1 and 2) under strategy A was 0.54, while it dropped to values between 0.27 and 0.32 for the remaining strategies. Similarly, the mean retention index (Farris, 1989) for those two characters showed a decrease in value from 0.88 to 0.61 - 0.68 after the data was treated. Fig. 6 shows the MPTs found using exclusively the continuous partition under strategies A (a), B (b), and C (c), with the optimization of the largest character, male body length, superimposed. The topology of strategy A shows a high degree of dependence on this character (as evident also from the values of CI and RI), which decreases both after rescaling and using implied weighting. Evidence supporting a similar effect of rescaling and implied weighting was also found using the ILD test, which revealed that continuous and discrete partitions were significantly incongruent when compared under strategy A ($p = 0.026$), while the three ways to treat data resulted in congruence between partitions (all $p > 0.315$).

However, it can also be seen that the resulting topologies (Fig. 6b and c) are also quite different from each another. In fact, the topological difference of the MPTs resulting from rescaling and using implied weighting on continuous data (comparison B - C) present somewhat intermediate values of SPR distances (Fig. 5) between the highly divergent

values obtained from comparing treated vs. untreated data, and the more congruent values resulting from comparing both trees obtained from data that was rescaled (comparison B - D) or weighted (comparison C - D). The other measures of topological similarity also show this intermediate placement, with the number of shared internal nodes attaining low values and the number of taxa in the agreement subtree higher ones (Fig. 5a). It is therefore evident that, although implied weighting is dealing with the magnitude of continuous characters, applying lower weights to large characters (Fig. 4) and reducing their influence on the final topology (Fig. 6), it is doing so in a different way than the rescaling of data, and the two strategies result in quite different supported phylogenies (Fig. 5).

Finally, a test was designed to study whether both alternatives where equally allowing a balanced influence of all characters on the final topology (see Materials and methods: Analysis of continuous characters), as originally proposed by Goloboff *et al.* (2006). The resulting null distributions can be seen in Fig. 7. Distributions obtained under strategies B, C and D showed a significant fit to a normal distribution (Chi-Square tests < 10.89, $p > 0.2$). Distribution obtained under strategy A showed a clear bimodal pattern, and did not fit a normal distribution (Chi-Square test = 135.11, $p < 0.00001$). The variances of the distributions were all significantly different from each other, with the smallest values corresponding to strategies C and D (3.68 and 3.91 respectively), while the distribution of strategy B showed a variance of 4.37 and that of strategy A of 9.87. Smaller variances can be attributed to a contraction of the distribution towards lower values, with a displacement of the right tail while the location of the left one remains relatively more constant (for example, the position of the value leaving 5% of data to the right varies between distributions 2.5 times as much as the position of the value that leaves 5% of the data to the left). Under all four strategies, the distance obtained using the tree supported by the subset including the 18 smallest characters was always near the median value of the distribution (with 39.4 to 50.8% of obtained distance values being smaller). On the contrary, the position of the distance obtained using the 18 largest characters differed widely among strategies. For strategy A, this subset of characters had an enormous influence in the final topology, showing an SPR distance 1.48 units shorter than the shortest one found by the random subset generation (for the values of SPR distances obtained under each strategy see Fig. 7). For strategy B, the position of the large subset left only 15.2% of values to the left of the distribution, and was therefore considered to be still strongly influencing the overall topology. On the contrary, among the two distributions obtained after rescaling (strategies C and D), the large character subset was placed almost at the mean point of the distribution, leaving 48.9 and 55% of distances to the left. In fact, for these two distributions, the distances separating the position of the small and large subsets not only decreased greatly (Fig. 7) with respect to the ones obtained for strategies A and B, their positions were even inverted, with the large character subset having slightly larger SPR distances than the small character one.

It was therefore concluded that implied weighting was in fact reducing the dominance of large characters on the final topology, although it was doing so in a partial way and was comparatively inefficient to eliminate such influence altogether, as was in fact happening when continuous characters were rescaled (Fig. 7). However, strategies C and D behaved very similarly in all tests developed, and the decision on whether to weight continuous characters against their homoplasy after these were rescaled proved to be the least decisive of all that were explored (SPR distance between nodes C and D in Fig. 5b is the shortest one). However, the MPT for the combined dataset under strategy C proved to be very different than the ones obtained under strategy D for all *k* values tested. Furthermore, strategy D resulted in a higher topological congruence between the trees obtained for the discrete and continuous partitions when analyzed separately (weighted SPR distances: 11.86 vs. 13.16), and in a higher mean group support for the MPT of the combined dataset (average jackknife resampling frequency: 52.51 vs. 44.47; number of nodes above 50%: 26 vs. 19) than strategy C. As a consequence, only strategy D was employed for the phylogenetic analysis, given that it showed evidence of being able to guarantee a balanced character influence and to increase the congruence between different sources of characters.

Cladistic analysis

As already discussed, the phylogenetic analysis was performed under strategy D, exploring 10 concavity functions of implied weighting (*k* = 1 to 10) and 3 progressively stricter criteria for the collapsing of sexually dimorphic continuous characters. This summed up to 30 forms of analysis, and a search using new technologies driven to find the optimum 10 times always resulted in a single most parsimonious tree for each of them (this phenomenon is common when using continuous characters and should not be taken as evidence of a strong phylogenetic signal, see Bardin *et al.*, 2013). All of the obtained trees were very similar, evidencing low levels of sensitivity to the parameters tested. Most major (generic or supra-generic) groups were very stable and almost uncontradicted in all of the obtained topologies, with 42 out of 48 resolved relationships present in the majority-rule consensus showing frequencies higher than 80% (see Fig. 8 for values of frequency). The few unresolved nodes mostly reflected conflict in the exact placement of single species within those major clades.

For each of these 30 trees, we calculated the frequency of occurrence of each of its groups in the complete set of trees. We then retained the phylogenetic hypothesis that had the highest mean group frequency, since it was the tree that showed the most commonly found, and less sensitive clades. This tree, shown in Fig. 8, was found for *k* values from 6 to 10, in congruence with previous studies that found that mild concavity values resulted in higher topological congruence between different morphological and molecular datasets

(Ramírez, 2003; Lopardo, 2005; Goloboff, 2008b); and a criterion to retain sexually dimorphic characters without collapsing only if there was a 30% difference between observed and predicted values in the male-female regression analysis. It is worth noting that this criterion retained without collapsing all characters that have been traditionally reported in taxonomic revisions and experimental works as sexually dimorphic: body size and leg length (Bonduriansky, 2006; 2007), length of the scape (Aczél, 1951; 1961), length of the postcranium (Sepúlveda *et al.*, 2013a); as well as some other features of the antennae and the wings (see Appendix S1 for a list of characters retained as dimorphic).

The following supra-generic groups were obtained with relatively strong support and where insensitive to changes in the parameters of the analysis (jacknifing absolute frequencies/Bremer support/majority rule frequencies): Neriidae, clade A (99/130.75/100); American neriids, clade F (56/34.79/100); Neotropical species of the *Eoneria*-group, clade G (70/16.84/100); and the *Longina* - *Cerantichir* - *Odontoloxozus* group, clade J (79/66.37/100). Likewise, the following genera are confirmed to be monophyletic based on the same evidence: *Chaetonerius* Hendel (60/26.57/100), *Glyphidops* Enderlein (82/15.66/100), *Indonesicesa* Koçak and Kemal (99/51.30/100), *Longina* (94/82.45/100) and *Nerius* Fabricius (91/20.39/80). The synapomorphies supporting these and other groups in Fig. 8 are shown in Appendix 1.

As can be seen, most genera are obtained as monophyletic, with only a few clades contradicting the current taxonomy, although some of these are also shown to be largely insensitive to the parameters explored (Fig. 8). Specifically, the monophyly of the two subgenera contained within *Glyphidops sensu* Aczél (1961) is never recovered, although the genera itself is found to be monophyletic. Furthermore, *Eoloxozus sabroskyi* is placed in all trees as the sister group of *Eoneria maldonadoi* and nested within *Eoneria* Aczél. Likewise, the genus *Cerantichir* is found to be paraphyletic in all of the obtained trees. The chosen phylogenetic hypothesis (Fig. 8) shows *C. peruana* and *Odontoloxozus longicornis* conforming a monophyletic group, in congruence with the original description of the first species as *Od. peruanus* (Hennig, 1937) and contradicting Buck's (2010) reassignment. Finally, the genus *Telostylinus* Enderlein is not retained as a natural group, but subdivided into several different (although closely related) clades. Other than these four examples, the resulting phylogeny is highly congruent with the taxonomy of the family.

Discussion

Continuous characters and the issue of scaling

The general apathy towards the use of continuous characters in phylogenetic analyses has long worn off. Few if any object nowadays to their inclusion in cladistic studies. Both from theoretical and practical points of view, quantitative characters have proven to be both valid and useful data for systematists (Rae, 1998; Wiens, 2001; 2004; Goloboff *et al.*, 2006; de Bivort *et al.*, 2010). Several breakthroughs have contributed to this change in general perception. First, if continuous and discrete characters are in fact not that different from each other, then there is no reason for their *a priori* exclusion from a phylogenetic analysis. Many, if not most, of the phenotypic variability found in nature is quantitative (Wiens, 2001; see also Baum, 1988; Chappill, 1989; Stevens, 1991; Thiele, 1993; Rae, 1998), independent on the decision to discretize them or not. If this is so, both types of characters are actually representing the same attributes of living organisms, with discrete characters being "informally discretized continuous characters" (de Bivort *et al.*, 2010: p. 302; similar arguments in Gift and Stevens, 1997; MacLeod, 2002; Haas, 2003), *i.e.*: continuous variation that is more intuitively assigned to different, non-overlapping categories. It is in fact the overlapping property of continuous data which initially posed a problem for character coding, with much of the subsequent debate stemming from an indiscriminate equation between 'continuous' and 'overlapping' (Thiele, 1993; Rae, 1998). Nonetheless, if in fact "many so-called qualitative characters are based on a quantitative phenomenological base filtered through the reified semantic discontinues of (…) terminology" (Stevens, 1991: p. 553), many of the original theoretical objections to the use of continuous characters (in the sense of Crisp and Weston, 1987; Pimentel and Riggins, 1987; Cranston and Humphries, 1988; Mickevich and Weller, 1990) are unjustified, and discrete characters may even suffer from similar arbitrariness in the circumscription of states.

Another factor leading to the change in perception on the use of quantitative data stems from Goloboff *et al.*'s (2006) incorporation into the software TNT of a set of algorithms to use continuous characters as such. This allowed to entirely eluding all discretization methods, which were considered inappropriate (Reid and Sidwell, 2002) or unavoidably arbitrary (Archie, 1985; Crisp and Weston, 1987; Felsenstein, 1988; Farris, 1990; Gift and Stevens, 1997). Subsequent implementations demonstrated that continuous characters carry useful phylogenetic information and formulate hypothesis congruent with other sources of characters (Goloboff *et al.*, 2006; Hornung-Leoni and Sosa, 2008; Pereyra and Mound, 2009; de Bivort *et al.*, 2010; Escapa and Catalano, 2013). Their dismissal from cladistic analyses is therefore unwarranted.

Despite all this progress, certain profound issues concerning the implementation of continuous characters in phylogenetic analysis have received little attention. One of such issues is that of scaling, that is, the differential influence of characters depending on the scale in which they are measured. Although many have acknowledged the importance of this problem, few have discussed alternatives to deal with it (Thiele, 1993; Wiens, 2001;

Goloboff *et al.*, 2006). So far, two different strategies, rescaling and using implied weights, have been proposed. The efficiency of both towards reducing the dominance of large characters, as well as the consequences of each on the phylogenetic reconstruction, had never been discussed. Our aim in the present study was to address these questions.

If implied weighting was in fact dealing with the issue of scaling, as originally proposed by Goloboff *et al.* (2006), it was necessary to demonstrate that characters were receiving weights according to the range they presented, such that larger character were assigned lower weights and small characters larger ones. To such end, we proved that the fit assigned by implied weights to continuous characters presented a significant and negative linear correlation with the range of a character (Fig. 4). This correlation was strong, explaining 76% of differences in fit, and was equally strong after excluding the largest characters in the matrix. Furthermore, the pattern was lost after characters were rescaled to unit range (Fig. 4), demonstrating that the covariation was in fact a consequence of the character's scale and not any other attribute relating to their phylogenetic signal. It appeared therefore plausible that implied weighting was actually dealing with the issue of scaling. As a matter of fact, different evidences showed that both implied weighting and rescaling reduced the dominance of large characters on the optimal topology: both methods eliminated the otherwise significant incongruence between continuous and discrete datasets when partitions were compared untreated, and resulted in trees that were very different than the one supported by the untreated continuous dataset (Fig. 5) and less dependent on the larger characters (Fig. 6). Nonetheless, the topologies supported after rescaling and implied weighting continuous characters were also quite different from each other, as shown by the measures of topological similarity employed (Fig. 5). It was therefore evident that, although both strategies were dealing with the issue of scaling, resulting in a decrease in the dominance of large characters, they were doing so in different ways. Since the choice of strategy had an important impact on the resulting phylogenetic hypothesis, both of them could not be used interchangeably. Furthermore, it was not clear whether implied weighting was as efficient in dealing with the issue of scaling as was standardizing the ranges of continuous characters.

To further evaluate this, we designed an experiment that aimed at detecting the topological dependence of the optimal phylogenetic hypothesis with respect to the scale of groups of characters, and see whether both methods were efficient in balancing the influence of different characters. This was done by comparing, through weighted SPR distances, the topological difference of trees supported by subsets of characters to the most parsimonious tree of the entire continuous dataset. Characters with a strong influence in determining the optimal tree of the entire partition will, when isolated, support a similar tree. On the contrary, characters with a weaker influence will build hypothesis more dissimilar to the one supported by the entire partition once all other characters are inactivated. The resulting histograms are shown in Fig. 7. When continuous data is used

without treatment (strategy A, Fig. 7a), the variance of the SPR distribution is larger, more than twice that of all other histograms. Since differences in variance where mostly due to changes in the position of the right tail of the distribution, it is evident that higher values represent a decrease in the overall balance of influence of the characters, with certain subsets contributing nothing (or very little) to the resulting phylogeny. In fact, only under strategy A did the distribution's right tail advance to values obtained when comparing random generated trees (weighted SPR distances > 24 units, mean = 28.62). The variance of the obtained distributions was significantly reduced by all other strategies of analysis, reaching the lowest values when characters were rescaled (strategies C and D).

An interesting pattern arose after the SPR distances for the trees supported exclusively by the 18 largest and 18 smallest characters were plotted on top of the distributions. The "small" character subset resulted, for all strategies, in values of SPR distances close to the median of the distribution, showing that their influence was relatively insensitive to the chosen strategy. On the contrary, values for the subset of the largest characters varied widely depending on the way continuous characters were analyzed. For untreated data, these characters resulted in a tree that was by far the most similar to the tree of the entire partition (Fig. 7a), showing that their influence in determining such topology was massive. When characters were analyzed under implied weighting (Fig. 7b), such influence became certainly reduced, with the difference in SPR distances between the "large" and "small" subsets diminishing from 10.85 to 1.52. Despite such reduction in the asymmetry of influences, large characters still retained a strong influence, with 84.8% of randomly generated subsamples having larger SPR distances. The asymmetry in influence is only completely eliminated after characters are rescaled. After this, the SPR distance obtained for "large" and "small" subsets differed only by 0.22 and 0.03 units from one another, depending on whether implied weighting was applied or not, respectively (Fig. 7c-d). Furthermore, both subsets were placed very near the mean value for each distribution, showing that rescaling had eliminated all information on the magnitude of the characters and these had become, despite their original differences in scale, two average subsets of characters with respect to their influence in the supported phylogeny.

All evidence points to the same conclusion: implied weighting reduces the issue of scaling, yet it does so only partially. Large characters still show a more than average influence on the resulting phylogeny, and this is shown to be exclusively due to their larger scales. As Farris (1990, p. 91) stated "(…) domination cannot logically be objectionable in itself, without some independent grounds for objecting to the dominant factor." The scale in which a character is measured is a factor that results in asymmetrical influences among different continuous characters, and yet it is completely arbitrarily determined during character coding and has no relationship whatsoever to phylogenetic signal. Objecting to such factor of dominance is therefore logical (hence the so called "issue of scaling"), and measures should be taken to eliminate it from phylogenetic inference. We found that

standardizing continuous characters to a common range is capable of doing so, while the use of implied weighting with such purposes is only partially successful. Furthermore, the strong correlation between the fit of a character and its scale when the rescaling step is avoided (Fig. 4) interfered with the use of implied weighting for its true purpose, to weight characters according to their homoplasy (Goloboff, 1993). This degree of homoplasy should not, once again, be influenced by a character's scale, since such factor has nothing to do with the degree of discordance between the character and a tree, nor with the lack of hierarchical structure in the character's state transformations. Therefore, characters are only weighted according to their homoplasy after they have been rescaled. When this is done, fit and scale no longer correlate (Fig. 4), and characters with similar original scales receive widely different weights (differences in fit after rescaling were up to 50 times larger than before doing so when comparing pairs of characters with similar scales). Furthermore, we found that such procedure resulted in a higher topological congruence between the discrete and continuous datasets, as well as formulating phylogenetic hypothesis with larger values of group support. Based on all the aforementioned evidence, we advocate that continuous characters should be rescaled before their use in cladistic analysis. We found that this procedure leads to the formulation of phylogenetic hypothesis that are unaffected by the scale in which characters were coded, a factor that is both arbitrary and irrelevant with respect to both the weight a character should receive and its influence on the resulting phylogeny. Rescaling continuous characters also allows for a subsequent use of implied weighting in which weights are assigned exclusively depending on the amount of homoplasy a character presents, a practice that has been shown to improve phylogenetic analyses of morphological datasets and increase the congruence of characters (Goloboff, 1997; Ramírez, 2003; Goloboff *et al.*, 2008a; as well as the present study).

Phylogeny of the Neriidae

After defining the strategy under which characters were used, we performed a sensitivity analysis (Wheeler, 1995; Giribet, 2003) to see which groups were present under different values of *k* and different criteria for the recognition of sexually dimorphic characters. None of these two parameters are likely to have a "correct" value, so instead of looking for one we decided to determine which groups are less dependent on such decision (as defended by Goloboff *et al.*, 2008a). The 30 different combinations of parameters explored resulted in highly similar trees, with only a few groups being resolved differently among cladograms. Moreover, these groups were mostly the result of differences in placement of single species, with higher-level clades remaining stable (see NT values in Fig. 8).

The monophyly of the family Neriidae was always recovered, and proved to be strongly supported by the data. Many of the synapomorphies retained reflect changes in the

morphology of the head and antennae (Appendix 1), among which are an elongated head with a larger postcranial region and a prolonged superior region of the mesofacial plate, porrect and elongated antennae with a longer scape and pedicel, the presence of an inner process of the pedicel, an apical positioning of the arista, which presents long pubescence and thickened and differentially colored basal flagellomeres, convergent postvertical bristles, a reduction/absence of anterior notopleural bristles and yellow colored procoxae.

The first species to branch off are always *Telostylus binotatus* and *Gymnonerius fuscus*. Both genera inhabit exclusively the Australian-Oriental regions, with their ranges of distribution including the Malay Peninsula, Indochina and most of the islands of the Malay Archipelago. The genus *Telostylus* Bigot is composed of 11 species (of which *Tl. binotatus* is the type species) and was already considered by Aczél (1954a; 1955a) to include some of the most basal neriids. On the other hand, the genus *Gymnonerius* Hendel includes only a single species, *Gy. fuscus*. This species shows extensive variability in both coloration and body proportions, which led Hennig (1937) to create numerous subspecies (Steyskal, 1977). Aczél (1955a) did not follow this opinion, arguing that a reexamination of a significant amount of specimens was needed before modifying the taxonomy of the genus. He also considered *Gymnonerius* to be morphologically very different from all other neriids of the Australian and Oriental regions, proposing this was due to its derived nature which brought it closer to the Neotropical neriids. Although in most trees these two genera branch off successively, in two of them they form a monophyletic group, supported by an elongated postcranium, reduced third costal section, shorter preabdomen, a reduction in the number of fronto-orbital bristles, a white antennal arista and long non-apical scutellar bristles. We did not choose this hypothesis given its low frequency in the sensitivity analysis, although the possibility for such basal monophyletic clade remains open for subsequent studies.

The remaining of the Australian-Oriental fauna, represented in the analysis by the genera *Telostylinus*, *Indonesicesa*, and *Paranerius* Bigot, was obtained as a complex group of early and poorly-resolved lineages. Despite the lack of strong support for most of these clades, and the change in position of several species among the obtained trees, some conclusions can be advanced. First of all, evidence is strong with respect to the monophyly of the genus *Indonesicesa*, a clade that was obtained among all trees with high levels of support (for synapomorphies see Appendix 1). All trees also suggested a close relationship between this genus and *Paranerius*, the two of them forming a monophyletic group in most of them. Both genera inhabit the western regions of the island of New Guinea (Aczél, 1954a; Pitkin, 1989), and share several synapomorphies of the female sex (larger size, shorter postcrania, smaller fourth costal section, wider ovipositor, 2-3 dorsal bristles in the procoxae), as well as a brown frontal vitta and short and spiniform bristles. On the other hand, the genus *Telostylinus* is never obtained as a monophyletic group, consisting instead of several clades whose exact number and composition differs among the different trees.

The pattern nonetheless shows a striking biogeographic consistency, with the early branches (*T. papuanus* and *T. spinicoxa*, as well as the two species that group with *Indonesicesa* and *Paranerius*, *T. zonalis* and *T. longipennis*) restricted to the island of New Guinea, while the more derived species form a monophyletic group (clade D of Figure 8, which includes the type species of the genus, *T. lineolatus*) whose distribution is almost entirely Micronesian (except for *T. lineolatus* which is widespread through the entire Southeast Asia and Oceania; Aczél, 1954a; 1955a; Pitkin, 1989). Further study is needed to precisely determine natural groups among this morphologically and geographically heterogeneous group.

The relationships among the remaining groups of neriids were more resolved and many of the groups obtained showed higher levels of stability and support (Fig. 8). We recovered a single origin for all species inhabiting outside the Australian-Oriental regions. This (almost entirely) Afrotropical + Neotropical clade (clade E) was present in 100% of trees, yet it is not taxonomically defined here due to poor values of jackknifing frequencies. Synapomorphies of this group include an enlargement of the frontal and genal regions of the head, longer and taller thorax, several changes in wing venation, a reduction in femoral length, thinner epandria, triangular and thin inner processes of the pedicel, the presence of anterior notopleural bristles and strong rows of spines on the male forecoxae. This group is further subdivided into the genus *Chaetonerius*, including all of the African fauna as well as some species from the Oriental region, of which only *Ch. inermis* is included in the analysis (Steyskal, 1977; Steyskal, 1980; Pitkin, 1989; Barraclough, 1993a); and the entire American fauna (clade F), including all species from both Neotropical and Nearctic regions (Steyskal, 1968; 1987; Buck, 2010). The two clades are strongly supported, and defining synapomorphies are listed in Appendix 1. Aczél's insistence on the condition of *Chaetonerius* as a "phylogenetically homogenous group" (Aczél, 1955b: 9) is therefore confirmed, as well as his subdivision of the genus into a *Ch. apicalis – Ch. collarti – Ch. ghesquierei* group and one including *Ch. brachialis* and *Ch. latifemur*.

Among neriids from the Americas, species are divided into two main groups, both of which include all American genera contained in the *Eoneria* and *Nerius*-groups defined by Aczél. We therefore propose to restrict these names to the Neotropical-Nearctic neriids. Defined in such a way, the *Eoneria*-group (clade G), including the genera *Antillonerius*, *Eoloxozus* and *Eoneria* is supported by an elongated head, a shorter postpedicel, an increase in the number of fronto-orbital bristles, the presence of katepisternal bristles and the continuation of the mesonotal pruinosity into the scutellum. Within this group, the basal position of *Antillonerius* is strongly inferred, while the relationships between *Eoneria* Aczél and *Eoloxozus* Aczél would imply the need to sinonymize both genera. However, a conservative approach is adopted, since the clade conformed by *Eoneria maldonadoi* and *Eoloxozus sabroskyi* shows very low values of Bremer support. On the other hand, the *Nerius*-group (clade H), including the genera *Nerius*, *Longina*, *Odontoloxozus*, *Cerantichir*

and *Glyphidops*, is supported by a strong increase in female size, an enlargement of the region of the thorax before the transverse suture, several changes in wing venation, a polished and shiny dorsal surface of the antennal bases, a reduction in the length of bristles (which become spiniform), the lack of occipital bristles, and a reduction in the number of bristles of both basicosta and male procoxae. The clade is further subdivided into the genus *Glyphidops* on one side, the most diverse of all Neotropical genera (Steyskal, 1968), and the remaining genera on the other (clade I). Aczél (1961) first proposed the close relationship between the genera *Glyphidops* and *Oncopsia* Enderlein (as defined by Enderlein, 1922), including both as subgenera of the former. Although the decision to include both in a single genus is validated by our results, we do not obtain the subgenera as monophyletic clades. In fact, the dense and whitish antennal pubescence that defined the subgenus *Glyphidops* (Aczél, 1961) is shown to be the result of at least two events of convergence. We consequently eliminate the subgeneric divisions within *Glyphidops*, with the subgenus *Oncopsia* becoming a junior synonym of the subgenus *Glyphidops*.

On the other hand, the relationships between the genera *Odontoloxozus*, *Longina* and *Cerantichir* are the only ones that have received recent attention. Buck and Marshall (2004) proposed the existence of a monophyletic clade including *Longina* and *Cerantichir*, citing the bare regions at the base of wings and the strong suprahumeral (first dorsocentral) bristles as possible synapomorphies. Buck (2010) afterwards transferred *Od. peruanus* to the genus *Cerantichir*, given the common apomorphic condition of an elongated antepronotal ridge and scutum which end beyond the level of humeral carina. Nonetheless, as discussed by Sepúlveda *et al.* (2013a), the genus *Longina* also shares this peculiar configuration of the anterior region of the thorax, and the other characteristics used to describe the genus *Cerantichir* by Buck (2010) are either also present in *Longina*, or absent from *C. peruana*. Aczél (1961) on the other hand, favored a close relationship between *Longina* and *Odontoloxozus*. Our results strongly confirm the monophyly of a *Cerantichir* + *Odontoloxozus* + *Longina* group (clade J), as well as the validity and derived status within the group of the genus *Longina* (Fig. 8). On the other hand, the monophyly of the genus *Cerantichir* as currently defined was never obtained. From all the combination of parameters analyzed, 80% of conditions resulted in the groupings shown in Fig. 8, with *C. enderleini* as the first species to branch off followed by a clade that divides into the *Longina* and a *C. peruana* + *Od. longicornis* group. This topology conflicts with Buck's (2010) reassignment. However, 20% of trees supported the following different configuration: (*Od. longicornis*, (*C. peruana*, (*C. enderleini*, (*Longina*))). In such case, placing *C. peruana* in either genus would result in non-natural groups. Once again, a conservative approach is favored, without modifying the taxonomy of the involved genera until their relationships are completely clarified.

The phylogeny resulting from our analysis has little in common with the taxonomic higher-level divisions of the family as proposed by Aczél (1961). The basal split between

the subfamilies Neriinae and Telostylinae is conclusively rejected. In fact, the lack of antennal bases, a character supposedly uniting the Telostylinae (Enderlein, 1922; Aczél, 1961) is determined to be a convergence between the genus *Telostylus*, arising as the most basal split within the family, before the evolutionary origin of the structure, and the genus *Chaetonerius* which have secondarily lost them. Furthermore, the polished and shiny dorsal region of these antennal bases, the character uniting the *Nerius*-group *sensu* Aczél (in contraposition to the dull antennal bases of the *Eoneria*-group) is found to have originated three times independently, in the lineages leading to *Gymnonerius*, *Paranerius* and finally to all Neotropical species presenting such character. Our phylogenetic hypothesis is also more congruent with the biogeography of the neriids. The scheme of relationships derived from Aczél's taxonomy (Fig. 1) proposed a much more complex biogeographic history, uniting in the same groups genera from many different continents. Aczél only referred twice to the question of the geographic origin of the family: first, stating that "We still have no idea where on Earth's surface branched off this family from a common root" (Aczél, 1954a: p. 507), and once again a few months later, when apparently he became convinced that the presence of the entire subfamily Telostylinae in the Oriental region was evidence supporting that region as the family's "primary center of distribution" (Aczél, 1954b: p. 2). Our results confirm this, since all species obtained as early branchings are restricted to Southeast Asia. From that ancestral region, several species of *Telostylinus* colonized the Oceanic islands, while their sister group divided into the mainly African *Chaetonerius* and the American fauna.

The lack of information regarding the phylogenetic relationships within Neriidae contrasts dramatically with the otherwise thorough knowledge of both higher and lower level phylogeny of many groups of Diptera (see Lambkin *et al.*, 2013 for a summary of phylogenetic studies of dipterans), a pattern that may well be derived from Hennig himself being a dipterologist (Richter and Meier, 1994; Meier, 2005). We provide here the first phylogenetic reconstruction of this family, one that is mainly the legacy of Martín L. Aczél. His thorough morphological descriptions and taxonomic revisions allowed for the construction of a detailed morphological matrix with which to elucidate the evolutionary history of neriid flies. Furthermore, his insistence in the use of continuous characters, as well as his systematic registry of intraspecific variation allowed us to overpass commonly cited obstacles towards the elucidation of a neriid phylogeny (Buck, 2010), resulting in the retention of a large amount of synapomorphies derived from body proportions, chaetotaxy and even male genitalia (Appendix 1). It is fair to say that Aczél has once again proven "the fundamental importance of morphology for entomology" (Aczél, 1951: p. 483).

Acknowledgements

The authors would like to thank E. Pérez, G. Claps and F. Navarro for permission to study the material at Instituto y Fundación Miguel Lillo and for the help provided during the stay; P. Goloboff and S. Rodríguez for contributing some of the scripts used; R. Bonduriansky and T. A. Sepúlveda for providing useful bibliography, S. D. A. Luna for his contribution to early stages of the work; P. Fontanarrosa, J. Padró and J. Hurtado for help capturing specimens of neriids that sparked this study. This work was supported by grants from ANPCyT, Universidad de Buenos Aires and CONICET PIP 2011-01-943. NMK is a student fellow of Universidad de Buenos Aires. IMS and MJR are members of Carrera del Investigador Científico (CONICET).

References

Abdala, C. S., Juárez Heredia, V. I. 2013. Taxonomía y filogenia de un grupo de lagartos amenazados: el grupo de *Liolaemus anomalus* (Iguania: Liolaemidae). Cuad. Herpetol. 27, 109–153.

Aczél, M. L. 1951. Morfología externa y división sistemática de las «Tanypezidiformes» con sinopsis de las especies argentinas de «Tylidae» («Micropezidae») y «Neriidae» (Dipt.). Acta Zool. Lilloana 11, 483–589.

Aczél, M. L. 1954a. Results of the Archbold Expedition: Neriidae von Neuguinea (Diptera). Treubia 22, 505–531.

Aczél, M. L. 1954b. Neriidae of the Belgian Congo (Diptera, Acalyptratae). Bull. Inst. R. Sc. N. B. 30, 1–23.

Aczél, M. L. 1954c. Neriidae in the collection of the Musée royal du Congo belge. Rev. Zool. Bot. Afr. 49, 161–166.

Aczél, M. L. 1955a. Neriidae von Indonesien (Dipt. Acalyptratae). Treubia 23, 19–40.

Aczél, M. L. 1955b. Neriidae (Diptera, Acalyptrata). Exploration du Parc National de l' Upemba. I. Mission G. F. de Witte 38, 85–92.

Aczél, M. L. 1955c. Nerridae in the collections of the Musée Royal du Congo Belge, Tervuren (Supplement). Rev. Zool. Bot. Afr. 51, 1–2.

Aczél, M. L. 1959. Diptera: Neriidae and Micropezidae (Tylidae). Insects of Micronesia 14, 47–90.

Aczél, M. L. 1961. A revision of American Neriidae (Diptera, Acalyptratae). Studia Ent. 4, 257–346.

Albrecht, G. H. 1978. Some comments on the use of ratios. Syst. Zool. 27, 67–71.

Archie, J. 1985. Methods for coding variable morphological features for numerical taxonomic analysis. Syst. Zool. 34, 326–345.

Atchley, W. R. 1978. Ratios, regression intercepts, and the scaling of data. Syst. Zool. 27, 78–83.

Atchley, W. R., Gaskins, C. T., Anderson, D. 1976. Statistical properties of ratios. I. Empirical results. Syst. Zool. 25, 137–148.

Bardin, J., Rouget, I., Yacobucci, M., Cecca, F. 2013. Increasing the number of discrete character states for continuous characters generates well-resolved trees that do not reflect phylogeny. Integr. Zool. http://dx.doi.org/10.1111/1749-4877.12076.

Barraclough, D. A. 1993a. The southern African species of Neriidae (Diptera). Ann. Natal Mus. 34, 1–17.

Barraclough, D. A. 1993b. Review of the type material of African *Chaetonerius* species (Diptera: Neriidae), with lectotype designations and new synonymy. J Afr. Zool. 107, 269–278.

Baum, B. 1988. A simple procedure for establishing discrete characters from measurement data, applicable to cladistics. Taxon 37, 63–70.

Baur, H., Leuenberger, C. 2011. Analysis of ratios in multivariate morphometry. Syst. Biol. 60, 813–825.

Bonduriansky, R. 2006. Convergent evolution of sexual shape dimorphism in Diptera. J. Morphol. 267, 602–611.

Bonduriansky, R. 2007. The evolution of condition-dependent sexual dimorphism. Am. Nat. 169, 9–19.

Bonduriansky, R. 2009. Condition dependence of developmental stability in the sexually dimorphic fly *Telostylinus angusticollis* (Diptera: Neriidae). J. Evol. Biol. 22, 861–872.

Bordewich, M., Semple, C. 2005. On the computational complexity of the rooted subtree prune and regraft distance. Ann. Combinatorics 8, 409–423.

Bremet, K. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. Evolution 42, 795–803.

Bremer, K. 1994. Branch support and tree stability. Cladistics 10, 295–304.

Buck, M. 2010. Neriidae. In: Brown, B.V., Borkent, A., Cumming, J.M., Wood, D.M., Woodley, N.E., Zumbado, M.A. (Eds.), Manual of Central American Diptera, Volume II. NRC Research Press, Ottawa, 56, pp. 815–819.

Buck, M, Marshall, S. A. 2004. A review of the genus *Longina* Wiedemann, with descriptions of two new species (Diptera, Neriidae). Stud. Dipterol. 11, 23–32.

Buck, M., McAlpine, D. K. 2010. Pseudopomyzidae. In: Brown, B.V., Borkent, A., Cumming, J.M., Wood, D.M., Woodley, N.E., Zumbado, M.A. (Eds.), Manual of Central American Diptera, Volume II. NRC Research Press, Ottawa, 57, pp. 821–825.

Carvalho-Filho, F. S., Esposito, M. C. 2008. Neriidae (Diptera: Schizophora) of the Brazilian Amazon: New records of genera and species, and key to species. Neotrop. Entomol. 37, 58–62.

Chappill, J. 1989. Quantitative characters in phylogenetic analysis. Cladistics 5, 217–234.

Colless, D. 1980. Congruence between morphometric and allozyme data for *Menidia* species: A reappraisal. Syst. Zool. 29, 288–299.

Corruccini, R. S. 1977. Correlation properties of morphometric ratios. Syst. Zool. 26, 211–214.

Cox, C. B. 2001. The biogeographic regions reconsidered. J. Biogeogr. 28, 511–523.

Cranston, P., Humphries, C. 1988. Cladistics and computers: a chironomid conundrum? Cladistics 4, 72–92.

Cresson, E. T. 1930. Notes and descriptions of some neotropical Neriidae and Mycropezidae. T. Am. Entomol. Soc. 56, 307–362.

Crisp, M., Weston, P. 1987. Cladistics and legume systematics, with an analysis of the Bossiaeeae, Brongniartieae and Mirbelieae. In: Stirton, C. (Ed.), Advances in Legume Systematics, Part 3. Royal Botanical Gardens, Kew, pp. 65–130.

de Bivort, B., Clouse, R. M., Giribet, G. 2010. A morphometrics-based phylogeny of the temperate Gondwanan mite harvestmen (Opiliones, Cyphophthalmi, Pettalidae). J. Zool. Syst. Evol. Res. 48, 294–309.

DeGusta, D. 2004. A method for estimating the relative importance of characters in cladistic analyses. Syst. Biol. 53, 529–532.

Eberhard, W. G. 1998. Reproductive behavior of *Glyphidops flavifrons* and *Nerius plurivitatus* (Diptera, Neriidae). J. Kansas Entomol. Soc. 71, 89–107.

Enderlein, G. von. 1922. Klassifikation der Mikropeziden. Arch. Naturgesch. 88, 140–229.

Escapa, I. H., Catalano, S. A. 2013. Phylogenetic analysis of Araucariaceae: Integrating molecules, morphology, and fossils. Int. J. Plant Sci. 174, 1153–1170.

Eulenstein, O., Chen, D., Burleigh, J. G., Fernández-Baca, D., Sanderson, M. J. 2004. Performance of flip-supertree construction with a heuristic algorithm. Syst. Biol. 53, 1–10.

Farris, J. S. 1989. The retention index and the rescaled consistency index. Cladistics 5, 417–419.

Farris, J. S. 1990. Phenetics in camouflage. Cladistics 6, 91–100.

Farris, J. S., Albert, V. A., Källersjö, M., Lipscomb, D., Kluge, A. G. 1996. Parsimony jackknifing outperforms neighbor-joining. Cladistics 12, 99–124.

Farris, J. S., Källersjö, M., Kluge, A. G., Bult. C. 1995. Testing significance of incongruence. Cladistics 10, 315–319.

Felsenstein, J. 1988. Phylogenies and quantitative characters. Ann. Rev. Ecol. Syst. 19, 445–471.

Fink, W. L., Zelditch, M. L. 1995. Phylogenetic analysis of ontogenetic shape transformations: a reassessment of the piranha genus *Pygocentrus* (Teleostei). Syst. Biol. 44, 343–360.

Gift, N., Stevens, P. F. 1997. Vagaries in the delimitation of character states in quantitative variation - an experimental study. Syst. Biol. 46, 112–125.

Giribet, G. 2003. Stability in phylogenetic formulations and its relationship to nodal support. Syst. Biol. 52, 554–564.

Giribet, G. 2007. Efficient tree searches with available algorithms. Evol. Bioinform. 3, 1–16.

Goloboff, P. A. 1993. Estimating character weights during tree search. Cladistics 9, 83–91.

Goloboff, P. A. 1997. Self-weighted optimization: tree searches and character state reconstructions under implied transformation costs. Cladistics 13, 225–245.

Goloboff, P. A. 1999. Analyzing large data sets in reasonable times: solutions for composite optima. Cladistics 15, 415–428.

Goloboff, P. A. 2002. Techniques for analyzing large data sets. In: DeSalle, R., Giribet, G., Wheeler, W.C. (Eds.), Techniques in Molecular Systematics and Evolution. Birkhäuser, Basel, pp. 70–79.

Goloboff, P. A. 2007. Calculating SPR distances between trees. Cladistics 23, 1–7.

Goloboff, P. A., Carpenter, J. M., Salvador Arias, J., Miranda Esquivel, D. R. 2008a. Weighting against homoplasy improves phylogenetic analysis of morphological data sets. Cladistics 24, 758–773.

Goloboff, P. A., Farris, J. S., Nixon, K. C. 2008b. TNT, a free program for phylogenetic analysis. Cladistics 24, 774–786.

Goloboff, P. A., Mattoni, C. I., Quinteros, A. S. 2006. Continuous characters analyzed as such. Cladistics 22, 589–601.

Gordon, A. 1980. On the assessment and comparison of classifications. In: Tommassone, R. (Ed.), Analyse de Donnes et Informatique. INRIA, LeChesnay, pp. 149–160.

Griffiths, G. C. D. 1972. The phylogenetic classification of Diptera Cyclorrhapha, with special reference to the structure of the male postabdomen. Series Entom. 8, III.

Haas, A. 2003. Phylogeny of frogs as inferred from primarily larval characters (Amphibia: Anura). Cladistics 19, 23–89.

Hendel, F. 1922. Die paläarktischen Muscidae acalyptratae Girsch. = Haplostomata Frey nach ihren Familien und Gattungen. I. Die Familien (Anm.: 1. Teil). Konowia 1, 145–160.

Hennig, W. 1934. Zur Kenntnis d. Kopulationsorgane de Tyliden. Zool. Anz. 107, 67–76.

Hennig, W. 1936. Beziehungen zwischen geographischer Verbreitung und Systematischer Gliederung bei einigen Dipterenfamilien. Ein Beitrag zum Problem der Gliederung system. Kategorien höherer Ordnung. Zool. Anz. 116, 161–175.

Hennig, W. 1937. Üebersichtüeber die Arten der Neriiden und üeber die Zoogeographie dieser Acalyptraten-Gruppe. Stettin. Ent. Ztg. 98, 240–280.

Hills, M. 1978. On ratios - a response to Atchley, Gaskins, and Anderson. Syst. Zool. 27, 61–62.

Hormiga, G., Scharff, N., Coddington, J. A. 2000. The phylogenetic basis of sexual size dimorphism in orb-weaving spiders (Araneae, Orbiculariae). Syst. Biol. 49, 435–462.

Hornung-Leoni, C. T., Sosa, V. 2008. Morphological phylogenetics of *Puya* subgenus *Puya* (Bromeliaceae). Bot. J. Linn. Soc. 156, 93–110.

Kluge, A. G., Farris, J. S. 1969. Quantitative phyletics and the evolution of anurans. Syst. Zool. 18, 1–32.

Lambkin, C. L., Sinclair, B. J., Pape, T., Courtney, G. W., Skevington, J. H., Meier, R., Yeates, D. K., Blagoderov, V., Wiegmann, B. M. 2013. The phylogenetic relationships

among infraorders and superfamilies of Diptera based on morphological evidence. Syst. Entomol. 38, 164–179.

Lopardo, L. 2005. Phylogenetic revision of the genus *Negayan* (Araneae, Anyphaenidae, Amaurobioidinae). Zool. Scr. 34, 245–277.

Lopardo, L., Giribet, G., Hormiga, G. 2011. Mophology to the rescue: molecular data and the signal of morphological characters in combined phylogenetic analyses – a case study from mysmenid spiders (Araneae, Mysmenidae), with comments on the evolution of web architecture. Cladistics 27, 278–330.

MacLeod, N. 2002. Phylogenetic signals in morphometric data. In: Macleod, N., Forey, P.L. (Eds.), Morphology, Shape and Phylogeny. Taylor and Francis, London, pp. 100–138.

Mangan, R. L., Baldwin, D. 1986. A new cryptic species of *Odontoloxozus* (Neriidae: Diptera) from the Cape Region of Baja California Sur (Mexico). Proc. Ent. Soc. Wash. 88, 110–121.

Mannion, P. D., Upchurch, P., Barnes, N., Mateus, O. 2013. Osteology of the Late Jurassic Portuguese sauropod dinosaur *Lusotitan atalaiensis* (Macronaria) and the evolutionary history of basal titanosauriforms. Zool. J. Linn. Soc. 168, 98–206.

Marshall, S. A. 2010. Micropezidae. In: Brown, B.V., Borkent, A., Cumming, J.M., Wood, D.M., Woodley, N.E., Zumbado, M.A. (Eds.), Manual of Central American Diptera, Volume II. NRC Research Press, Ottawa, 55, pp. 805–813.

McAlpine, D. K. 1966. Description and biology of an Australian species of Cypselosomatidae (Diptera), with a discussion of family relationships. Aust. J. Zool. 14, 673–685.

McAlpine, D. K. 1974. The subfamily classification of the Mycropezidae and the genera of Eurybatinae (Diptera: Schizophora). J. Entomol. Ser. B 43, 231–245.

McAlpine, D. K. 1996. Relationships and classification of the Pseudopomyzidae (Diptera: Nerioidea). Proc. Linn. Soc. N. S. W. 116, 223–232.

McAlpine, J. F. 1987. Cypselosomatidae. In: McAlpine, J.F., Peterson, B.V., Shewell, G.E., Teskey, H.J., Vockroth, J.R., Wood, D.M. (Eds.), Manual of Nearctic Diptera, Volume II. Research Branch, Agriculture Canada, Ottawa, Ontario, 55, pp. 757–760.

McAlpine, J. F. 1989. Phylogeny and classification of the Muscomorpha. In: McAlpine, J.F., Wood, D.M. (Eds.), Manual of Nearctic Diptera, Volume III. Research Branch, Agriculture Canada, Ottawa, Ontario, pp. 1397–1518.

Meier, R. 2005. Role of Dipterology in phylogenetic systematics: the insight of Willi Hennig. In: Yeates, D.K., Wiegmann, B.M. (Eds.), The Evolutionary Biology of Flies. Columbia University Press, New York, pp. 45–62.

Mello, R. L. 2010. The Diptera described by Martín L. Aczél. Stud. Dipterol. 17, 223–236.

Mickevich, M. F. 1978. Taxonomic congruence. Syst. Zool. 27**,** 143–158.

Mickevich, M. F., Farris, J. S. 1981. The implications of congruence in *Menidia*. Syst. Zool. 30, 351–370.

Mickevich, M. F., Weller, S. J. 1990. Evolutionary character analysis: tracing character change on a cladogram. Cladistics 6, 137–170.

Nixon, K. C. 1999. The parsimony ratchet, a new method for rapid parsimony analysis. Cladistics 15, 407–414.

Pereyra, V., Mound, L. A. 2009. Phylogenetic relationships within the genus *Cranothrips* (Thysanoptera, Melanthripidae) with consideration of host associations and disjunct distributions within the family. Syst. Entomol. 34, 151–161.

Pimentel, R., Riggins, R. 1987. The nature of cladistic data. Cladistics 3, 201–209.

Pitkin, B. R. 1989. Family Neriidae. In: Evenhuis, N.L. (Ed.), Catalog of the Diptera of the Australasian and Oceanian Regions. Bishop Museum & E. J. Brill, Honolulu, pp. 468–469.

Prado, A. P. 1984. Family Cypselosomatidae. In: Papavero, N. (Ed.), A catalogue of the Diptera of the Americas South of the United States. Departamento de Zoologia, Secretaria de Agricultura, São Paulo, pp. 1–2.

Rae, T. 1998. The logical basis for the use of continuous characters in phylogenetic systematics. Cladistics 14, 221–228.

Ramírez, M. J. 2003. The spider subfamily Amaurobioidinae (Araneae, Anyphaenidae): a phylogenetic revision at the generic level. Bull. Am. Mus. Nat. Hist. 277, 1–262.

Reid, G., Sidwell, K. 2002. Overlapping variables in botanical systematics. In: Macleod, N., Forey, P.L. (Eds.), Morphology, Shape and Phylogeny. Taylor and Francis, London, pp. 53–66.

Richter, S., Meier, R. 1994. The development of phylogenetic concepts in Hennig's early theoretical publications (1947–1966). Syst. Biol. 43, 212–221.

Rohlf, F. J. 2010. tpsDig version 2.16. http://life.bio.sunysb.edu/morph/index.html.

Sepúlveda, T. A., Pereira-Colavite, A., de Carvalho, C. J. B. 2013a. Revision of the Neotropical genus *Cerantichir* (Diptera: Neriidae) with new records and a key to species. Rev. Colomb. Entomol. 39, 125–131.

Sepúlveda, T. A., Wolf, M. I., de Carvalho, C. J. B. 2013b. Revision of the Neotropical genus *Eoneria* (Diptera: Neriidae) with description of a new species from Colombia. Zootaxa 3636, 245–256.

Shatalkin, A. 1994. Palearctic species of Pseudopomyzidae (Diptera). Russ. Entomol. J. 3, 129–145.

Simpson, P., Woehl, R., Usui, K. 1999. The development and evolution of bristle patterns in Diptera. Development 126, 1349–1364.

StatSoft, Inc. 2001. STATISTICA (data analysis software system), version 6. www.statsoft.com

Stevens, P. F. 1991. Character states, morphological variation, and phylogenetic analysis: A review. Syst. Bot. 16, 553–583.

Steyskal, G. C. 1965. Synonymy of the genera *Antillonerius* and *Imrenerius* (Diptera: Neriidae). Proc. Ent. Soc. Wash. 67, 60.

Steyskal, G. C. 1968. Family Neriidae. In: Papavero, N. (Ed.), A Catalogue of Diptera of the Americas South of the United States. Departamento de Zoologia, Secretaria da Agricultura, São Paulo, pp. 1–7.

Steyskal, G. C. 1977. Family Neriidae. In: Delfinado, M.D., Hardy, D.E. (Eds.), A Catalogue of the Diptera of the Oriental Region, Volume III. Suborder Cyclorrapha (excluding Division Aschiza). University Press of Hawaii, Honolulu, pp. 8–11.

Steyskal, G. C. 1980. Family Neriidae. In: Crosskey, R.W. (Ed.), Catalogue of the Diptera of the Afrotropical Region. British Museum (Natural History), London, p. 578.

Steyskal, G. C. 1987a. Neriidae. In: McAlpine, J.F., Peterson, B.V., Shewell, G.E., Teskey, H.J., Vockroth, J.R., Wood, D.M. (Eds.), Manual of Nearctic Diptera, Volume II. Research Branch, Agriculture Canada, Ottawa, Ontario, 57, pp. 769–771.

Steyskal, G. C. 1987b. Mycropezidae. In: McAlpine, J.F., Peterson, B.V., Shewell, G.E., Teskey, H.J., Vockroth, J.R., Wood, D.M. (Eds.), Manual of Nearctic Diptera, Volume II. Research Branch, Agriculture Canada, Ottawa, Ontario, 56, pp. 761–768.

Strait, D. S, Grine, F. E. 2004. Inferring hominoid and early hominid phylogeny using craniodental characters: the role of fossil taxa. J. Hum. Evol. 47, 399–452.

Swofford, D. L. 1991. When are phylogeny estimates of molecular and morphological data incongruent? In: Miyamoto, M.M., Cracraft, K. (Eds.), Phylogenetic Analysis of DNA Sequences. Oxford Univ. Press, New York, pp. 295–333.

Thiele, K. 1993. The Holy Grail of the perfect character: The cladistic treatment of morphometric data. Cladistics 9, 275–304.

Thiele, K., Ladiges, P. Y. 1988. A cladistic analysis of *Angophora* Cav. (Myrtaceae). Cladistics 4, 23–42.

Vargas, S., Breedy, O., Guzman, H. M. 2010. The phylogeny of *Pacifigorgia* (Coelenterata, Octocorallia, Gorgoniidae): a case study of the use of continuous characters in the systematics of hte Octocorallia. Zoosystema 32, 5–18.

Wheeler, D. C. 1995. Sequence alignment, parameter sensitivity, and the phylogenetic analysis of molecular data. Syst. Biol. 44, 321–331.

Wiegmann, B. M., Trautwein, M., Winkler, I., Barr, N., Kim, J.-W., *et al*. 2011. Episodic radiations in the fly tree of life. Proc. Natl. Acad. Sci. USA 108, 5690–5695.

Wiens, J. J. 2001. Character analysis in morphological phylogenetics: problems and solutions. Syst. Biol. 50, 689–699.

Wiens, J. J. 2004. The role of morphological data in phylogeny reconstruction. Syst. Biol. 53, 653–661.

Winston, J. E. 1999. Describing species: Practical taxonomic procedure for biologists. Columbia University Press, New York.

Yeates, D. K., Wiegmann, B. M. 2005. Phylogeny and evolution of Diptera: recent insights and new perspectives. In: Yeates, D.K., Wiegmann, B.M. (Eds.), The Evolutionary Biology of Flies. Columbia University Press, New York, pp. 14–44.

Yeates, D. K., Wiegmann, B. M., Courtney, G. W., Meier, R., Lambkin, C. & Pape, T. 2007. Phylogeny and systematics of Diptera: two decades of progress and prospects. Zootaxa 1668, 565–590.

Appendix 1: Synapomorphies

The following table (Table 2) shows the synapomorphic features of selected genera and supra-generic groups retained in the phylogenetic hypothesis shown in Fig. 8. Character number is stated between parentheses. For synapomorphies derived from uncollapsed (that is, dimorphic) continuous characters, the intervening sex is made explicit. Otherwise, no particular sex is mentioned, and both character numbers are separated by a slash. Continuous characters were mapped with their original scales. See Appendix S1 for character description and abbreviations.

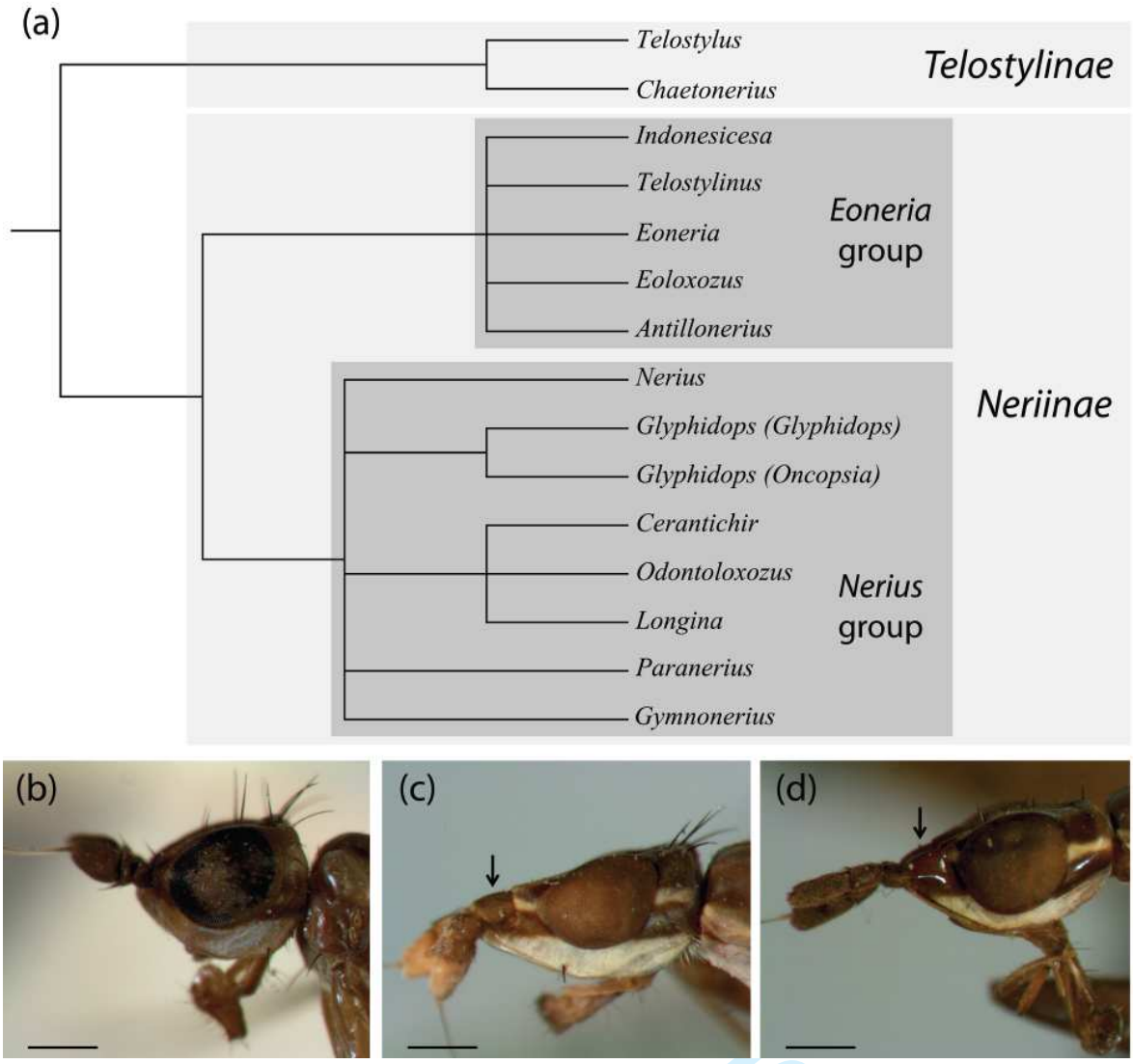

Fig. 1: (a) Current phylogenetic hypothesis within the family Neriidae (only genera included in the present study are shown). **(b)** *Chaetonerius apicalis* (female), a member of the Telostylinae which lack antennal bases. **(c)** *Eoneria maldonadio* (female), a member of the *Eoneria*-group that present antennal bases with a dull dorsal region. **(d)** *Nerius pilifer* (male), a member of the Nerius-group that present antennal bases with a shiny dorsal region. Black arrows show the position of the antennal bases. Scale bars = 0.5 mm.

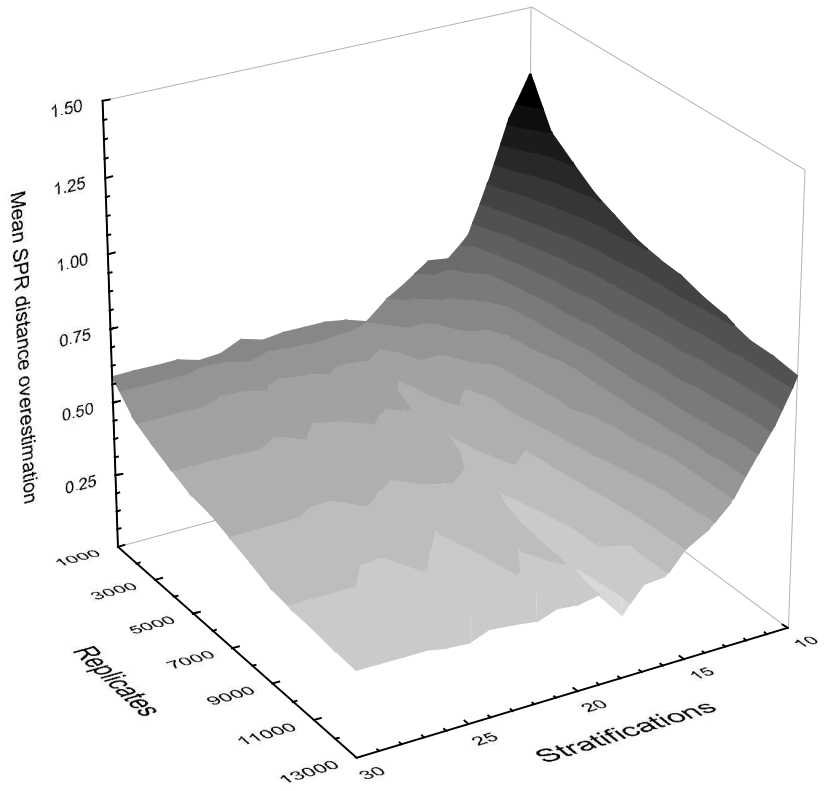

Fig. 2: Efficiency of the weighted SPR distance algorithm (j = 3) in finding the minimum distance for several combinations of parameters. The plotted surface shows the mean overestimation of 50 different SPR distances. This was calculated by finding the shortest distance for each one of the 50 pairs of trees, and subtracting that amount to all values obtained for that particular pair (the shortest path for each pair therefore becomes 0, and all other distances are expressed as the positive amount by which they overestimated that same distance). Minimum bias was obtained using 18 stratifications and 13,000 replicates.

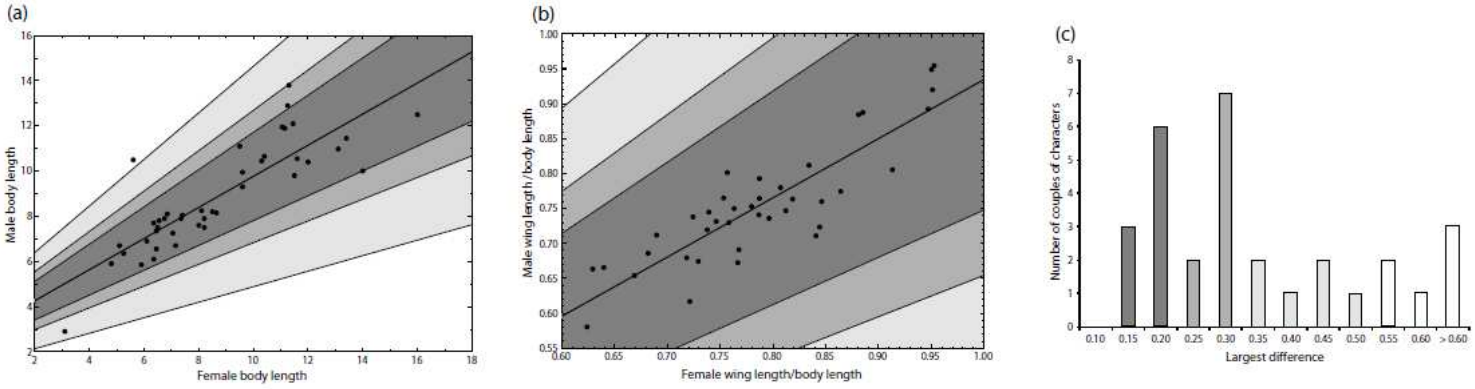

Fig. 3: Independence analysis for male-female couples of characters. **(a, b)** Two examples of the regression analysis performed for all couples of male-female characters. The three stricter criteria for the recognition of sexual dimorphism are shown as more inclusive gray regions. Characters shown in **(a)** (char. numbers 1, 2) are retained as sexually dimorphic for all criteria, while those shown on **(b)** (char. numbers 35, 36) are always collapsed into a single character. **(c)** Histogram grouping all male-female couples of characters into categories (range 0.05) according to the largest difference between observed and expected values, as fraction of the latter, found in the regression analyses. Colors are as in figures **(a)** and **(b)**, and show the number of characters collapsing for each criterion.

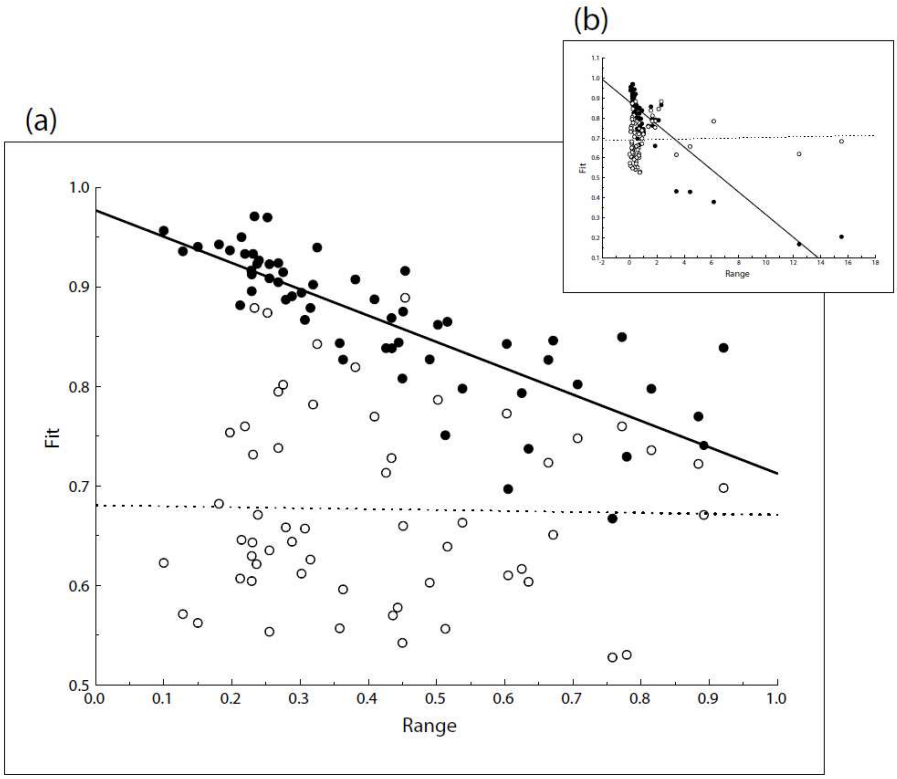

Fig. 4: Regression of the fit assigned by implied weights on the character's range, corresponding to the MPT obtained with (empty circles, dotted line) and without (black circles, solid line) rescaling continuous data. Results are the same whether using the entire continuous dataset (figure **(b)**, statistical values in text), or only characters with ranges < 1 (80% of total characters, figure **(a)**). The significant and strong correlation between both variables ($p < 0.0001$, $R^2 = 0.66$) is lost if implied weighting is applied after rescaling characters ($p = 0.87$, $R^2 < 0.001$).

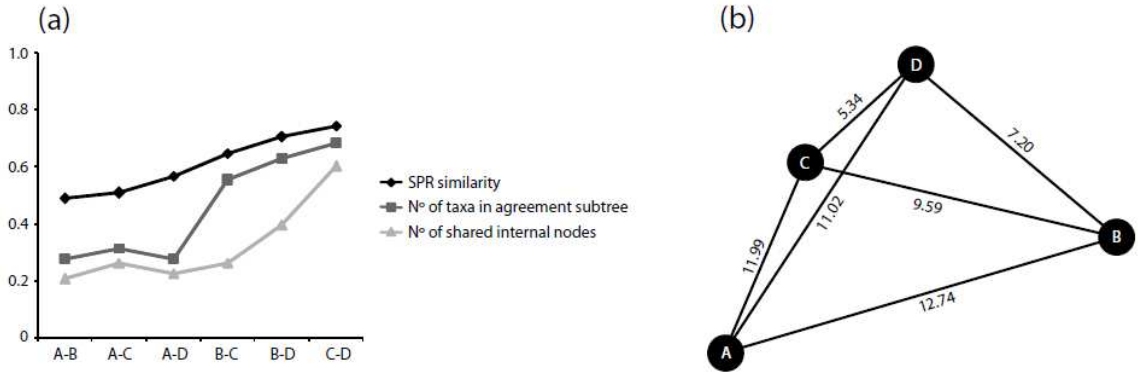

Fig. 5: Topological effect of different strategies of analysis, resulting from the comparison of the resulting MPTs. **(a)** All pair-wise comparisons of MPTs using unweighted SPR derived similarity (black), number of taxa in agreement subtree (dark gray) and number of shared internal nodes (light gray). **(b)** In-scale tetrahedron, with each side representing the weighted SPR distances (j = 3) between MPTs (vertex labels according to the strategy employed).

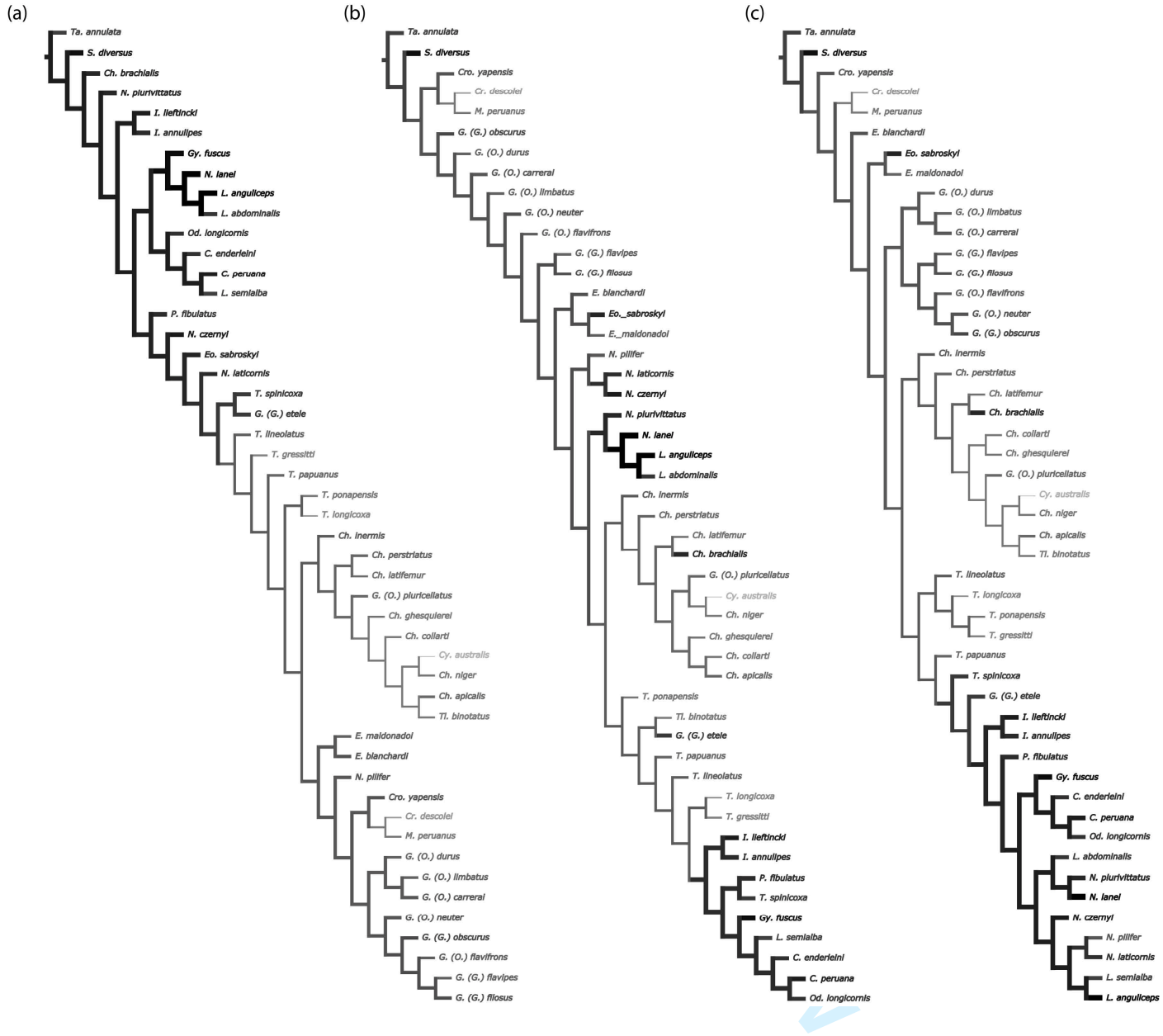

Fig. 6: Single MPTs obtained using the continuous character partition under strategies A **(a)**, B **(b)** and C **(c)**. The largest character (male body length) has been optimized on each topology, with broader and darker branches representing higher character values. All three topologies differ considerably among each other. Both implied weighting and rescaling reduce the MPT's dependence on the extremely large characters, as seen by the higher homoplasy present in trees **(b)** and **(c)**.

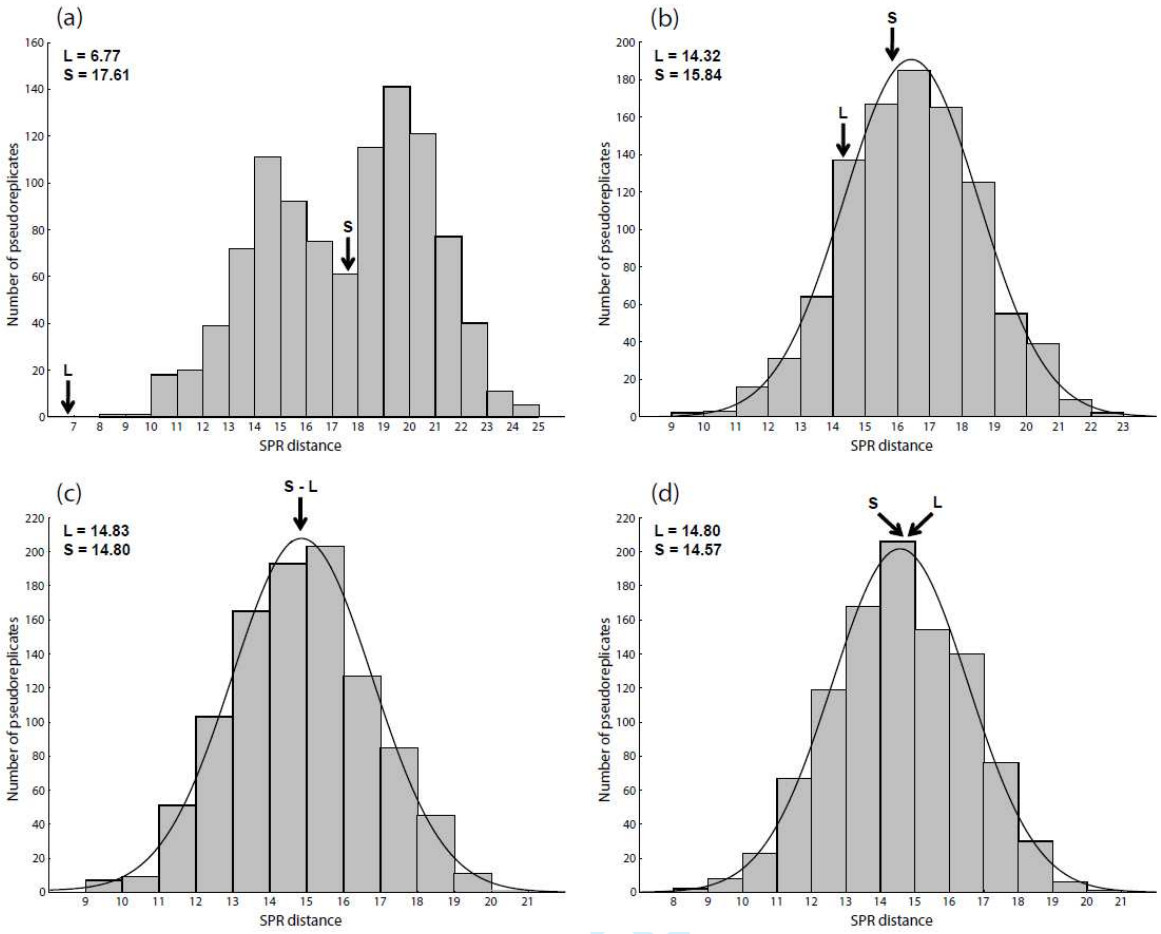

Fig. 7: Distribution of weighted SPR distances comparing 1000 trees obtained from randomly generated subsets of 18 continuous characters, with the MPT from the entire continuous dataset. The letter naming each distribution coincides with the strategy used to obtain it. Distance values are grouped into discrete categories of unit range. A line delineating the inferred normal distribution is shown for those distributions that showed significant fitting. The SPR distances for the large (L) and small (S) character subsets can be seen at the top left corner of each graph, and black arrows show their location on the histogram.

Fig. 8: Preferred topology. The tree was found after rescaling continuous characters, collapsing male-female pairs in the absence of residues > 30% of the expected value, and using implied weights with *k* from 6 to 10. Fit = 60.72176 (for *k* = 6), length = 847.175, CI = 0.283, RI = 0.623. Numbers above branches represent Bremer support (BS) in units of fit x 100, and numbers below branches are, to the left, jackknifing absolute frequencies (JK, only shown are values > 50), and, to the right, number of trees containing the group (NT, maximum value = 30). Capital letters denote supra-generic clades that are discussed in the text (see Appendix 1 for a list of synapomorphies).

Table 1. List of species included in the analysis, with their taxonomic position and authority (according to Steyskal, 1968; 1980; 1987; Pitkin, 1989; Barraclough, 1993b; Buck, 2010), and geographical distributions (limits following Cox, 2001).

Family	Subfamily	Genus	Species	Author	Distribution
Micropezidae	Taeniapterinae	*Taeniaptera*	*Ta. annulata*	Fabricius, 1787	Neotropical
		Scipopus	*S. diversus*	Hendel, 1936	Neotropical
	Micropezinae	*Micropeza*	*M. (Micropeza) peruanus*	Hennig, 1936	Neotropical
		Cryogonus	*Cr. (Cressonius) descolei*	Aczél, 1949	Neotropical
	Eurybatinae	*Crosa*	*Cro. yapensis*	Steyskal, 1952	Australian
Cypselosomatidae		*Cypselosoma*	*Cy. australis*	McAlpine, 1966	Australian
Neriidae	Telostylinae	*Telostylus*	*Tl. binotatus*	Bigot, 1859	Australian/Oriental
		Chaetonerius	*Ch. apicalis*	Walker, 1849	Afrotropical
			Ch. brachialis	Enderlein, 1922	Afrotropical
			Ch. collarti	Aczél, 1954	Afrotropical
			Ch. ghesquièrei	Aczél, 1954	Afrotropical
			Ch. inermis	Schiner, 1868	Oriental
			Ch. latifemur	Enderlein, 1922	Afrotropical
			Ch. niger	Czerny, 1932	Afrotropical
			Ch. perstriatus	Speiser, 1910	Afrotropical
	Neriinae	*Nerius*	*N. czernyi*	Aczél, 1961	Neotropical
			N. lanei	Aczél, 1961	Neotropical
			N. laticornis	Hennig, 1937	Neotropical
			N. plurivittatus	Bigot, 1886	Neotropical
			N. pilifer	Fabricius, 1805	Neotropical
		Glyphidops	*G. (Glyphidops) etele*	Aczél, 1961	Neotropical
			G. (G.) filosus	Fabricius, 1805	Neotropical
			G. (G.) flavipes	Wiedemann, 1830	Neotropical
			G. (G.) obscurus	Hennig, 1937	Neotropical
			G. (G.) ochreus	Hennig, 1937	Neotropical
			G. (Oncopsia) carrerai	Aczél, 1961	Neotropical
			G. (O.) durus	Cresson, 1926	Neotropical
			G. (O.) flavifrons	Bigot, 1886	Neotropical/Nearctical
			G. (O.) limbatus	Enderlein, 1922	Neotropical
			G. (O.) neuter	Hennig, 1937	Neotropical
			G. (O.) pluricellatus	Schiner, 1868	Neotropical
		Longina	*L. abdominalis*	Wiedemann, 1830	Neotropical
			L. anguliceps	Buck & Marshall, 2004	Neotropical
			L. semialba	Buck & Marshall, 2004	Neotropical
		Odontoloxozus	*Od. longicornis*	Coquillet, 1904	Neotropical/Nearctical
		Cerantichir	*Ce. enderleini*	Hennig, 1937	Neotropical
			Ce. peruana	Hennig, 1937	Neotropical
		Gymnonerius	*Gy. fuscus*	Wiedemann, 1984	Oriental
		Paranerius	*P. fibulatus*	Enderlein, 1922	Australian
		Eoneria	*E. blanchardi*	Aczél, 1951	Neotropical
			E. maldonadoi	Aczél, 1961	Neotropical
		Eoloxozus	*Eo. sabroskyi*	Aczél, 1961	Neotropical
		Antillonerius	*A. cinereus*	Röder, 1885	Neotropical
		Indonesicesa	*I. annulipes*	Doleschall, 1858	Australian
			I. lieftincki	Aczél, 1954	Australian
		Telostylinus	*T. gressitti*	Aczél, 1959	Australian
			T. lineolatus	Wiedemann, 1930	Australian/Oriental
			T. longicoxa	Thomson, 1869	Australian
			T. longipennis	Aczél, 1954	Australian
			T. papuanus	Meijere, 1915	Australian
			T. ponapensis	Aczél, 1959	Australian
			T. spinicoxa	Aczél, 1954	Australian
			T. yapensis	Aczél, 1959	Australian
			T. zonalis	Aczél, 1954	Australian

Table 2: List of sinapomorphies.

Group	Synapomorphies
Neriidae (clade A)	Head height/length (7/8): 0.830-0.952 → 0.667
	Male eye length/height (12): 0.930-0.936 → 0.971
	Male postcranium length/head length (16): 0.163-0.173 → 0.202-0.225
	Male length of scape/head length (18): 0.036-0.060 → 0.092
	Length of pedicel/head length (20/21): 0.050-0.051 → 0.267
	Male third costal section/wing length (43): 0.161 → 0.098-0.113
	Ultimate section of M_{1+2}/third section of costal vein (47/48): 2.928 → 3.897-3.917
	Superior region of mesofacial plate (84): short → prolonged
	Postvertical bristles (93): divergent → convergent
	Antenna (97): porrect and short → porrect and elongated
	Shape of scape (98): linear → subglobose to obconical
	Inner process of pedicel (99): absent → present
	Position of antennal arista (104): dorsal → apical
	Basal flagellomeres of arista (105): slender → thickened
	Pubescence of antennal arista (106): absent → long
	Differentially pigmented region of antennal arista (109): absent → encompassing basally enlarged region
	Anterior pair of notopleural bristle (121): equal/subequal to posterior pair → absent, or hair-like
	Color of procoxa (171): brown → yellow
Clade B	Female body length (2): 6.800-8.200 → 8.500-8.700
	Anterior/posterior region of frons (11): 0.974-1.000 → 0.706-0.833
	Wing length/body length (35/35): 0.755-0.772 → 0.796-0.897
	Female second costal section/wing length (42): 0.570-0.579 → 0.605
	Female third costal section/wing length (44): 0.132-0.136 → 0.108-0.118
	Upper margin of dark lateral vitta of occiput (78): running straight → moving downwards
	Occipital bristles (96): 2-3 → absent
	Proepisternal bristles (120): one strong and spine-like pair → one short and inconspicuous pair
	Preabdominal marginal longitudinal vittae (189): absent → present
Clade C	Female body length (2): 6.800-8.200 → 6.400
	Genae height/head height (14/15): 0.108-0.115 → 0.132-0.147
	Female length of pedicel/head length (21): 0.277-0.286 → 0.246-0.247
	Length of postpedicel/head length (24/25): 0.312-0.316 → 0.300
	Male width/length of postpedicel (26): 0.522-0.544 → 0.571-0.581
	Wing width/length (37/38): 0.263-0.264 → 0.277-0.282
	Male first costal section/wing length (39): 0.027-0.028 → 0.030-0.035
	Male third costal section/wing length (43): 0.111-0.113 → 0.117-0.123

	Male fourth costal section/wing width (45): 0.069-0.072 → 0.104-0.114
	Ultimate section of M_{1+2}/third section of costal vein (47/48): 3.897-3.917 → 3.647-3.750
	Female mid femur length/thorax length (55): 1.633 → 1.624-1.629
	Hind femur length/thorax length (56/57): 1.833-1.871 → 1.760-1.786
	Length of epandrium/body length (69): 0.092-0.096 → 0.107-0.111
	Clear vitta in occiput (76): absent → complete
Clade D	Male body length (1): 7.500-7.900 → 7.200-7.400
	Female frons width/head width (10): 0.404-0.412 → 0.350-0.364
	Male length of scape/head length (18): 0.115-0.116 → 0.135-0.147
	Hind tibia length/hind femur length (62/63): 0.829-0.831 → 0.880-0.882
	Color of meso and metacoxae (172): brown → yellow
Clade E	Male frons width/head width (9): 0.387-0.458 → 0.473-0.475
	Genae height/head height (14): 0.132-0.147 → 0.161-0.163
	Thorax length/body length (28/29): 0.311-0.315 → 0.322-0.325
	Thorax height/length (33/34): 0.838-0.839 → 0.841-0.852
	Female first costal section/wing length (40): 0.030-0.031 → 0.035-0.045
	Female second costal section/wing length (42): 0.570-0.576 → 0.547-0.554
	Male third costal section/wing length (43): 0.117-0.123 → 0.135
	Ultimate section of M_{1+2}/third section of costal vein (47/48): 3.647-3.750 → 3.110-3.300
	Length of A_1+CuA_2/length of CuA_2 (51): 3.455 → 3.369-3.439
	Fore femur length/thorax length (52/53): 1.491-1.544 → 1.226-1.239
	Female mid femur length/thorax length (55): 1.624-1.629 → 1.451-1.550
	Female preabdomen width/length (67): 0.450-0.464 --> 0.477-0.481
	Width/length of epandrium (70): 0.320-0.325 → 0.263-0.271
	Shape of inner process of pedicel (100): finger-like → triangular and thin
	Anterior pair of notopleural bristles (121): absent → hair-like, or equal/subequal to posterior pair
	Antero and posteroventral row of spines on fore femur (162): absent, or reduced and hair-like → present
	Apical dark stripe in hind femur (184): absent → present
Clade F	Male postcranium length/head length (16): 0.167-0.169 → 0.176-0.177
	Thorax length before suture/behind suture (30): 0.568-0.571 → 0.670
	Thorax width/length (31/32): 0.583-0.593 → 0.559-0.560
	Female first costal section/wing length (40): 0.035-0.045 → 0.048-0.049
	Female second costal section/wing length (42): 0.547-0.554 → 0.516-0.517
	Length of A_1+CuA_2/length of CuA_2 (51): 3.369-3.439 → 3.186-3.355
	Fore femur length/thorax length (52/53): 1.226-1.239 → 0.882-1.079
	Male mid femur length/thorax length (54): 1.651-1.786 → 1.159-1.636
	Female mid femur length/thorax length (55): 1.451-1.550 → 1.029-1.446
	Hind femur length/thorax length (56/57): 1.600-1.786 → 1.233-1.365
	Female preabdomen width/length (67): 0.477-0.481 → 0.485-0.500

	Length of syntergite 7+8/body length (68): 0.080-0.081 → 0.061-0.063
	Color of frontal vitta (74): apically yellow → completely yellow
	Position of ocellar plate (80): before posterior eye margin → on posterior eye margin
	Shape of posterior margin of head (83): curved → straight
	Relative lengths of *vti* and *vte* bristles (92): *vti* shorter → equal
	Position of antennal arista (104): apical → subapical
	Pubescence of antennal arista (106): long → microscopically short
	Proepisternal bristles (120): one strong and spine-like pair → one short and inconspicuous pair
	Relative lengths of *sa* and *pa* bristles (131): approximately equal → *pa* larger
	Color of pruinosity of mesonotum (137): yellow → gray
	Longitudinal dustless medial line in mesonotum (138): absent → present
Clade G	Head length/body length (3/4): 0.208-0.215 → 0.222-0.252
	Female eye length/height (13): 1.212-1.220 → 1.225-1.229
	Length of postpedicel/head length (24/25): 0.292-0.300 → 0.266-0.279
	Number of *orsa* bristles (88): 2 → 3
	Katepisternal bristle (128): absent → present
	Lateral pruinosity in scutellum (144): absent → present
Clade H	Female body length (2): 6.400 → 8.000-8.200
	Thorax length before suture/behind suture (30): 0.670 → 0.914-0.950
	Male second costal section/wing length (41): 0.554-0.563 → 0.485-0.489
	Length of A_1+CuA_2/length of CuA_2 (51): 3.186-3.355 → 2.262-2.315
	Fore tibia length/fore femur length (58/59): 0.985-1.016 → 0.983
	Texture of dorsal surface of antennal bases (86): not polished → polished and shiny
	Type of bristles (87): long and bristle-like → short and spiniform
	Occipital bristles (96): 1, or 2-3 → absent
	Number of bristles in basicosta (158): 2 → 1
	Number of dorsal bristles in male procoxae (166): 3 → 2
Clade I	Male body length (1): 7.900 → 9.800-10.000
	Female body length (2): 8.000-8.200 → 8.500-9.500
	Head width/length (5/6): 0.697-0.771 → 0.667-0.691
	Head height/length (7/8): 0.591-0.634 → 0.556-0.569
	Male postcranium length/head length (16): 0.176-0.203 → 0.241-0.254
	Female postcranium length/head length (17): 0.170-0.175 → 0.229
	Female length of scape/head length (19): 0.119-0.130 → 0.133
	Male length of process of pedicel/length of pedicel (22): 0.400-0.434 → 0.250-0.263
	Female length of process of pedicel/length of pedicel (23): 0.417-0.469 → 0.286-0.321
	Thorax length/body length (28/29): 0.322-0.325 → 0.331-0.337
	Thorax width/length (31/32): 0.559-0.560 → 0.533-0.556
	Male third costal section/wing length (43): 0.135 → 0.113-0.127
	Female third costal section/wing length (44): 0.135-0.136 → 0.133

	Prebasal section of M_{1+2}/median section of M_{1+2} (48/49): 0.921-0.950 → 0.879
	Fore tibia length/fore femur length (58/59): 0.983 → 0.911-0.952
	Ovipositor width/length (72): 0.399-0.508 → 0.363-0.390
	Shape of inner process of pedicel (100): triangular and thin → triangular and broad
Clade J	Female body length (2): 8.500-9.500 → 11.500
	Head height/length (7/8): 0.556-0.569 → 0.533-0.546
	Female frons width/head width (10): 0.410-0.412 → 0.459
	Genae height/head height (14/15): 0.161-0.176 → 0.177-0.198
	Male postcranium length/head length (16): 0.241-0.254 → 0.266-0.288
	Female postcranium length/head length (17): 0.229 → 0.235-0.240
	Male length of pedicel/head length (20): 0.250-0.269 → 0.271-0.389
	Female length of process of pedicel/length of pedicel (23): 0.286-0.321 → 0.274
	Male width/length of postpedicel (26): 0.571-0.581 → 0.508
	Thorax width/length (31/32): 0.533-0.556 → 0.481-0.488
	Thorax height/length (33/34): 0.841-0.852 → 0.727-0.738
	Preabdomen length/body length (64/65): 0.375-0.383 → 0.357-0.371
	Female preabdomen width/length (67): 0.485-0.500 → 0.365
	Width/length of epandrium (70): 0.263-0.271 → 0.277-0.283
	Clear vitta in occiput (71): complete → incomplete
	Position of clear vitta (77): medial → inferior
	Shape of posterior margin of head (83): straight → curved
	vte bristle (91): present → absent
	Shape of first flagellomere (101): ovate → subrectangular
	Position of antennal arista (104): subapical → dorsoapical
	General pattern of coloration of antennal arista (108): brown → white
	Configuration of anterior region of thorax (116): scutum and antepronotal ridge ending at level of postpronotal carina → ending beyong level of postpronotal carina
	Shape of katepisternum (117): higher than wide → as wide as high
	Postpronotal bristles (119): absent → hair-like
	Relative lengths of *sa* and *pa* bristles (131): *pa* larger → approximately equal
	Shape of dorsal face of mid femur (164): straight/slightly convex → concave
Chaetonerius	Head width/length (5/6): 0.762-0.771 → 0.874-1.000
	Head height/length (7/8): 0.632-0.676 → 0.693-0.751
	Anterior/posterior region of frons (11): 0.974-1.000 → 0.933
	Female postcranium length/head length (16): 0.166-0.170 → 0.134-0.144
	Male length of process of pedicel/length of pedicel (22): 0.400-0.434 → 0.500-0.622
	Female length of process of pedicel/length of pedicel (23): 0.439-0.500 → 0.588
	Thorax length before suture/behind suture (30): 0.568-0.571 → 0.511-0.534
	Wing length/body length (35/36): 0.755-0.772 → 0.834-0.862
	Male third costal section/wing length (43): 0.135 → 0.144
	Preabdomen length/body length (64/65): 0.383 --> 0.400-0.409

	Antennal bases (85): present → absent
	Number of extra dorsocentral bristles posterior to transverse suture (126): 0 → 1
	Non-apical scutellar bristles (132): absent/vestigial → long
	Number of lateral bristles on metacoxae (169): 1 → 2
Glyphidops	Head length/body length (3/4): 0.208-0.215 → 0.195-0.199
	Head width/length (5/6): 0.697-0.771 → 0.827-0.843
	Male frons width/head width (9): 0.473-0.475 → 0.433-0.438
	Male length of scape/head length (18): 0.115-0.116 → 0.112
	Male length of process of pedicel/length of pedicel (22): 0.400-0.434 → 0.459
	Female width/length of postpedicel (27): 0.594-0.605 → 0.561
	Thorax length/body length (28/29): 0.322-0.325 → 0.316-0.321
	Female second costal section/wing length (42): 0.516-0.517 → 0.467-0.485
	Male third costal section/wing length (43): 0.135 → 0.142-0.149
	Prebasal section of M_{1+2}/median section of M_{1+2} (49/50): 0.921-0.950 → 1.014-1.064
	Preabdomen length/body length (64/65): 0.383 → 0.387-0.400
	Ovipositor length/body length (71): 0.217-0.218 → 0.196
	Ovipositor width/length (72): 0.399-0.508 → 0.548
	Upper margin of dark lateral vitta of occiput (78): running straight → fused with ocellar plate
	Proepisternal bristles (120): one short and inconspicuous pair → absent
	Number of bristles in basicosta (158): 1 → 0
Indonesicesa	Female body length (2): 10.300 → 10.600-10.700
	Head width/length (5/6): 0.762-0.771 → 0.857-0.907
	Head height/length (7/8): 0.591-0.640 → 0.683-0.716
	Anterior/posterior region of frons (11): 0.706-0.833 → 0.678
	Male eye length/height (12): 1.186-1.188 → 1.119
	Female eye length/height (13): 1.165 → 1.138-1.155
	Genae height/head height (14/15): 0.074-0.115 → 0.138-0.158
	Male postcranium length/head length (16): 0.167-0.169 → 0.157-0.166
	Female postcranium length/head length (17): 0.164 → 0.142-0.148
	Thorax length before suture/behind suture (30): 0.568-0.571 → 0.673
	Female second costal section/wing length (42): 0.605 → 0.547-0.564
	Male fourth costal section/wing width (45): 0.061-0.072 → 0.057-0.058
	Width/length of epandrium (70): 0.320-0.328 → 0.337-0.347
	Shape of posterior margin of head (83): curved → straight
	Relative lengths of *vti* and *vte* bristles (92): *vti* shorter → equal
	Occipital bristles (96): absent → 2-3
	Apex of first flagellomere (103): rounded → pointed
	Spinules on fore tibia (163): absent → present
	Club-shaped thickening in apex of male fore tibia (165): absent → present
	Color of meso and metacoxae (172): brown → yellow

	Apical dark stripe in tibiae (185): absent → present
Longina	Female body length (2): 11.600-13.100 → 13.400 Male length of scape/head length (18): 0.116-0.144 → 0.272 Female length of scape/head length (19): 0.133 → 0.236-0.284 Color of frontal vitta (74): apically yellow, or completely yellow → brown Clear vitta in occiput (76): incomplete → absent Vibrisae (94): absent → present Shape of scape (98): subglobose to obconical → semicylindrical Subcostal break (150): present → absent Bare areas of wing (157): absent, or small → large Ciliae on upper squamae (159): long → reduced Antero and posteroventral row of spines on fore femur (162): present → reduced and hair-like Number of dorsal bristles in female procoxae (167): 3, or 4 → 5 Apical dark stripe in tibiae (185): present → absent
Nerius	Genae height/head height (14/15): 0.161-0.176 → 0.160 Male width/length of postpedicel (26): 0.571-0.581 → 0.588-0.667 Wing width/wing length (37/38): 0.274-0.288 → 0.259-0.265 Female fourth costal section/wing width (46): 0.094-0.100 → 0.086-0.092 Ultimate section of M_{1+2}/third section of costal vein (47/48): 2.878-3.300 → 3.449-3.465 Length of syntergite 7+8/body length (68): 0.061-0.063 → 0.070-0.074 Length of epandrium/body length (69): 0.115 → 0.148-0.153 Width/length of epandrium (70): 0.263-0.271 → 0.179-0.194 Ovipositor length/body length (71): 0.217-0.224 → 0.243-0.256 Position of ocellar plate (80): on posterior eye margin → behind posterior eye margin Katepisternal bristle (128): absent → present Shape of crossvein *dm-cu* (153): straight to slightly convex → strongly convex Number of central rings in anterior femur (175): 1 → 0 Number of central rings in mid femur (179): 1 → 0 Apical dark stripe in tibiae (186): present → absent

Appendix S1: Character description

Continuous characters

The subsequent list shows all characters initially registered and included in the analysis of continuous characters. Characters marked with an asterisk are those computed using male and female data indistinctively. Contiguous male-female couples of characters that are shaded gray were collapsed into a single character given that they presented a lack of outliers for the criteria that led to the preferred phylogenetic hypothesis, *i.e.* a difference between at least one observed and expected value > 30% of the latter in the regression analysis. Non-shaded characters were retained as such, resulting in a quantitative dataset of 54 continuous characters that was used for the cladistic analysis of Neriidae. Some comments on the way in which certain measurements were taken can be found between brackets next to the respective characters.

1. Male body length.	(Excluding antennae and postabdomen)
2. Female body length.	
3. Male head length/body length.	
4. Female head length/body length.	
5. Male head width/head length.	
6. Female head width/head length.	
7. Male head height/head length.	
8. Female head height/head length.	
9. Male frons width/head width.	(Measured at vertex)
10. Female frons width/head width.	
11. Anterior/posterior region of frons. *	
12. Male eye length/eye height.	
13. Female eye length/eye height.	
14. Male genae height/head height.	(The height of the genae was measured below the eyes)
15. Female genae height/head height.	
16. Male postcranium length/head length.	
17. Female postcranium length/head length.	
18. Male length of scape/head length.	
19. Female length of scape/head length.	
20. Male length of pedicel/head length.	
21. Female length of pedicel/head length.	
22. Male length of process of pedicel/length of pedicel.	
23. Female length of process of pedicel/length of pedicel.	

24. Male length of first flagellomere/head length.	
25. Female length of first flagellomere/head length.	
26. Male width of first flagellomere/length of first flagellomere.	
27. Female width of first flagellomere/length of first flagellomere.	
28. Male thorax length/body length.	(Including scutellum)
29. Female thorax length/body length.	
30. Thorax length before transverse suture/thorax length behind suture. *	
31. Male thorax width/thorax length.	
32. Female thorax width/thorax length.	
33. Male thorax height/thorax length.	
34. Female thorax height/thorax length.	
35. Male wing length/body length.	
36. Female wing length/body length.	
37. Male wing width/wing length.	
38. Female wing width/wing length.	
39. Male first costal section/wing length.	
40. Female first costal section/wing length.	
41. Male second costal section/wing length.	
42. Female second costal section/wing length.	
43. Male third costal section/wing length.	
44. Female third costal section/wing length.	
45. Male fourth costal section/wing width.	
46. Female fourth costal section/wing width.	
47. Male ultimate section of M_{1+2}/third section of costal vein.	
48. Female ultimate section of M_{1+2}/third section of costal vein.	
49. Male prebasal section of M_{1+2}/median section of M_{1+2}.	
50. Female prebasal section of M_{1+2}/median section of M_{1+2}.	
51. Length of A_1+CuA_2/length of CuA_2. *	
52. Male fore femur length/thorax length.	
53. Female fore femur length/thorax length.	
54. Male mid femur length/thorax length.	
55. Female mid femur length/thorax length.	
56. Male hind femur length/thorax length.	
57. Female hind femur length/thorax length.	
58. Male fore tibia length/fore femur length.	
59. Female fore tibia length/fore femur length.	
60. Male mid tibia length/mid femur length.	
61. Female mid tibia length/mid femur length.	
62. Male hind tibia length/hind femur length.	
63. Female hind tibia length/hind femur length.	

64. Male preabdomen length/body length.	
65. Female preabdomen length/body length.	
66. Male preabdomen width/preabdomen length.	(The width of the preabdomen was measures at the posterior margin of the second tergite)
67. Female preabdomen width/preabdomen length.	
68. Length of syntergite 7+8/body length.	
69. Length of epandrium/body length.	
70. Width of epandrium/length of epandrium.	(The width of the epandrium was measured at the apex)
71. Ovipositor length/body length.	
72. Ovipositor width/ovipositor length.	(The width of the ovipositor was taken at the base of the structure)

Discrete characters

All multistate characters were treated as unordered unless otherwise stated. Autapomorphic characters were excluded.

73. Shape of frons between the eyes: convex (0), concave (1). The frons is more or less deeply impressed in all Neriidae and some Micropezinae (sometimes adopting a V-shape), while it is convex and outwardly rounded in other micropezids, as in the Taenipterinae and most Eurybatinae (Aczél, 1951).

74. Colour of frontal vitta: brown (0), apically yellow (1), completely yellow (2). The coloration markings on the frons vary considerably within and among genera. These markings sometimes include totally or partially the fronto-orbital plates, but this distinction was not taken into account. Some species present a characteristic yellowish wedge-shaped spot restricted to the anterior region of the frons (state 1), typical of the *Chaetonerius* but also present in some species of *Telostylinus* (Aczél 1954a, 1954b, 1955b); others, show a yellow marking that extends back at least until the ocellar tubercle (state 2). The character was treated as ordered, since the presence of a completely yellow frontal vitta includes having an apical spot.

75. Postcranium pigmentation pattern: homogenous (0), dark occiput and clear postgenae (1). All neriids present a head laterally divided into a darker upper third, encompassing the occiput behind the eyes, and a whitish postgenal area. Micropezids included in the analysis have laterally homogenous postcrania, while *Cypselosoma australis* has a clearer genal region.

76. Clear vitta in occiput: absent (0), incomplete (1), complete (2). Within the dark region of the occiput (char. 75), a clear yellowish vitta is generally present, more intensely marked

in Neotropical species. This vitta is nonetheless absent (state 0) in some species (eg.: some *Chaetonerius* and *Telostylinus*) most of them from the Australian-Oriental region, while it is incomplete and faint (state 1) in the genera *Odontoloxozus* and *Cerantichir*, not attaining the posterior margin of the head. The character was treated as ordered.

77. Position of clear vitta in occiput: medial (0), submedial (1), inferior (2). Most species with a clear vitta within the dark region of the occiput present it in a medial position (state 0). However, within some species of the genus *Nerius*, this vitta is placed submedially (state 1, Aczél, 1961), and in the genera *Cerantichir* and *Odontoloxozus* it is "(…) placed near lower margin of this blackish brown area" (Aczél, 1961: 298). The character was treated as ordered.

78. Upper margin of dark lateral vitta of occiput: moving downwards (0), running straight (1) fused with ocellar plate (2). Most species present a yellowish frontal vitta (char. 74), and even among those that do not, a clear stripe develops in the fronto-orbital plates. In both cases, the dark region of the occiput presents a well defined upper margin. This margin can move downwards as it progresses towards the back of the head (state 0), as in the genus *Rhoptrum* (Aczél, 1954a), run straight (state 1), or advance towards the dorsal region of the head, fusing with the ocellar plate, as in some *Glyphidops* (Aczél, 1961). In case this upper margin is absent, the state was codified as inapplicable.

79. Auxiliary ocellar lamina: absent (0), present (1). Among the Micropezinae and Eurybatinae, the ocellar lamina continues backwards up until the vertex, forming an elevation that is similarly strongly quitinized and pigmented as the ocellar tubercle. This structure is absent from the rest of the Nerioidea.

80. Position of ocellar plate: close to anterior eye margin (0), before posterior eye margin (1), on posterior eye margin (2), behind posterior eye margin (3). The placement of the ocellar lamina on the frons is clearly anterior among Taeniapterinae, while it adopts a more posterior position in Micropezinae, Eurybatinae and Neriidae (Aczél, 1951). However, even among these, the exact position with respect to the posterior eye margin can vary considerably depending on the elongation of the head and the postcranium. This character was treated as ordered.

81. Ocellar furrow: absent (0), present (1). Exclusively in the genus *Longina* (*L. abdominalis* and *L. semialba*), the ocellar tubercle is surrounded both laterally and posteriorly by a deep furrow (Aczél, 1961; Buck and Marshall, 2004).

82. Shape of posterior region of occiput: rounded (0), sharply truncated (1). Aczél (1951: 490) discussed the shape of the head among this group of flies, stating that only in the Neriidae was the posterior region of the occiput always sharply truncated, given that it does not participate in the elongation of the head. However, some genera (*Gymnonerius*,

Teloneria, *Paranerius* and *Cerantichir*) have species with an elongated occiput forming a rounded vertex.

83. Shape of posterior margin of head: curved (0), straight (1). The posterior margin of the head is straight in most Neotropical neriids, continuing the sharp angle of the occiput (char. 82). On the other hand, this margin is curved in most Australian-Oriental neriids and the *Cerantichir-Odontoloxozus-Longina* clade, with the maximum length of the head occurring at its center.

84. Superior region of mesofacial plate: short (0), prolonged (1). A character shared by all Neriidae and some Micropezinae is the prolonged mesofacial plate that, although not forming antennal bases in the latter nor in the Telostylinae (char. 85), gives nonetheless the head of both groups a conical and longitudinally elongated morphology (Aczél, 1951).

85. Antennal bases: absent (0), present (1). The anterior region of the mesofacial plate (upper face, or lunule) has become inflated, protruding and medially divided in a group of Neriidae, adopting the appearance of an extra segment of the antenna. This peculiar configuration of the head, usually referred to as antennal bases (Aczél 1951, 1961), divides the family into its two constituting subfamilies (Enderlein, 1922; Aczél, 1954a).

86. Texture of dorsal surface of antennal bases: not polished (0), polished and shiny (1). This character was proposed by Hennig (1937) to constitute the tribal limits of the family, but in Aczél's systematic revision of the family (1954) was downgraded, and subsequently used as the base for dividing the Neriinae into the *Eoneria* and the *Nerius* groups (1961).

87. Type of bristles: long and bristle-like (0), short and spiniform (1). Aczél (1951) considered the presence of short and strong bristles to be characteristic of the Neriidae, with only a few exceptions. However, after analyzing the African and Oceanic faunas, the presence of "long, bristle-like bristles" in many genera of the Old World was interpreted as plesiomorphic, and partly contributed to the conception of the Telostylinae and the *Eoneria*-group as basal clades (Aczél, 1954a; 1961).

88. Number of superior anterior orbital bristles: range 0 - 4. The fronto-orbital bristles in all Nerioidea are located in a simple reclinate series (D. K. McAlpine, 2000). The anterior pairs (if present) always stand on a barely quitinized region of the fronto-orbital plate (Aczél, 1951), while a single pair is generally present on the posterior, strongly quitinized region. Although many authors no longer discriminate between the anterior pairs (superior anterior orbital bristles or *orsa* according to Hennig, 1937), and the posterior one (superior superior orbital bristle or *orss*), this distinction was maintained here since it allowed to separate the fronto-orbital bristles into positionally homologous elements. The character was treated as ordered.

89. Superior superior orbital bristles: absent (0), present (1). See char. 88. This pair of bristles is only absent in the Micropezinae.

90. Vertical interior bristles (vti): absent (0), present (1).

91. Vertical external bristle (vte): absent (0), present (1).

92. Relative length of vti and vte bristles: vti larger (0), equal (1), vti shorter (2). The relative lengths of both cephalic and thoracic bristles were often stated by Aczél in his taxonomic descriptions. In particular, *vti* bristles are more commonly elongated (state 0) in micropezids, while among Neriidae they are generally equal (state 1) to *vte* in the Neotropical genera and considerably reduced (and therefore shorter than *vte*, state 2) in most Old World genera. This character was treated as continuous.

93. Postocellar bristles: divergent (0), convergent (1). Postocellar bristles (referred to as postvertical bristles or *pvt* by Aczél) are invariantly present in all included taxa. Nonetheless, in Neriidae these are generally strongly convergent (state 1) and even crossed if long enough, considered by many as an apomorphic state (Aczél, 1951; Hennig, 1958, McAlpine, 1974). Among Micropezidae and Cypselosomatidae they are always divergent (state 0).

94. Vibrissae: absent (0), present (1).

95. Genal bristles: absent (0), present (1).

96. Occipital bristles: absent (0), 1 (1), 2-3 (2), many (3). This character was divided into four states, representing the absence of occipital bristles (state 0), the presence of a single bristle, of only a few or of many (states 1, 2 and 3 respectively). This grouping corresponds to the variability observed in revised specimens, with many of them having 2 occipital bristles on one side of the head and three on the other, whereas the distinction with respect to the presence of many occipital bristles is very clear. This character was treated as ordered.

97. Antennae: pending and relatively short (0), porrect and short (1), porrect and elongated (2). The elongated and porrect antenna (state 2) of the Neriidae is one of the most well established synapomorphies of the family (Cresson, 1938; Aczél, 1951, 1961; Steyskal, 1968; Barraclough, 1993a; Buck, 2010). In contraposition, the antennae of the other families within Nerioidea is relatively short (Aczél, 1951), with pending pedicel and first flagellomere (state 0), except in the Cypselosomatidae (J. F. McAlpine, 1981b), which have porrect antennae yet retaining its short length (state 1). The character was treated as ordered.

98. Shape of scape: linear (0), subglobose to obconical (1), semicylindrical (2). This character reflects the relationship between the length and the height of the scape. A given

scape was taken to be linear (state 0) if it was significantly higher than long (length = 1/3 – 2/3 of height), subglobose to obconical (state 1) if both dimensions were approximately equivalent (length = 0.8 - 1.5 the height), and semicylindrical (state 2) when the length was at least twice the height (proposed as a synapomorphy of the *Longina* by Buck and Marshall, 2004).

99. Inner process of pedicel: absent (0), present (1). The inner face of the pedicel of the neriid antennae presents a projection that attaches to the median region of the first flagellomere (Steyskal, 1987a). This effectively separates the pedicel into a membered shaped at the base (referred to as body) and an apical prolongation (referred to as process; see chars. 22, 23) (Aczél, 1951), a synapomorphic antennal configuration that is not present in any other dipteran family (Aczél, 1954a).

100. Shape of inner process of pedicel: triangular and broad (0), triangular and thin (1), finger-like (2). A lot has been written on the shape, length, width and apex of the inner process of the pedicel of the different genera of neriid flies. However, many authors have used the same terminology to describe slightly different morphologies, as is the case of the term "finger-like" in the work of Enderlein (1922), Aczél (1951, 1954a) and Sepúlveda *et al.* (2013a). Some authors have even changed the way in which they described this structure throughout their work, leading to inconsistent depictions; e.g.: the use of "finger-like" between Aczél 1951 and 1954, the criteria relating to the broadness of the process between Aczél 1954 and 1961, among others. Here, processes were divided into two different general shapes: a triangular morphology, with both sides more or less straightly tapering into an apex, and a finger-like morphology, in which both sides run parallel through some portion of the structure. By this definition, all Neotropical Neriidae present triangular processes, in contrast with what Sepúlveda *et al.* (2013a) described. On the other side, this definition of "finger-like" is more congruent with what Aczél (1954a, 1955) described as processes developing a "longitudinal keel". Triangular processes were further subdivided into those which are broad (state 0) and those which are thin (state 1), considering broad all processes whose base occupy the entire inner side of the pedicel. Since all finger-like processes are thin, the character was treated as ordered, being therefore equivalent to a couple of characters, one relating to its shape and the other to its broadness.

101. Shape of first flagellomere: ovate (0), obovate (1), subrectangular (2). First flagellomeres can be broadly divided into those presenting an ovoid shape, and those whose dorsal and ventral margins run parallel (Sepúlveda *et al.*, 2013a), creating a subrectangular shape. Ovoid postpedicels were further classified by Aczél (1961) as ovate or obovate depending on whether the maximum width is attained near the base or the apex of the segment, respectively, and we maintained this classification.

102. Base of first flagellomere: rounded (0), truncate (1). The base of the first flagellomere is in most species rounded, given that the apex of the pedicel is clearly concave when seen

in a lateral view (state 0). However, in a few species of Neotropical neriids, the apex of the pedicel ends in a sharp and transverse apex (Aczél, 1961), therefore the first flagellomere emerges from a truncate base (state 1).

103. Apex of first flagellomere: pointed (0), rounded (1), truncate (2). This important feature has been widely used in taxonomic descriptions and genera definitions (Aczél, 1961). Although related to both position of antennal arista (char. 104) and shape of the first flagellomere (char. 101), the shape of the apex of the postpedicel is not unequivocably determined by these other attributes. The apex of this segment was considered pointed when it presents a width equivalent to the arista, and truncate when it ends in a surface more or less perpendicular to the longitudinal axis. On the other hand, the majority of first flagellomeres were considered to present a rounded apex, without discriminating those that were described as "widely rounded" (Aczél, 1961; Sepúlveda *et al.*, 2013a), given that this attribute depends only on the width of the postpedicel near the apex (a variable already included in chars. 26 and 27), and does not represent a difference in morphology.

104. Position of antennal arista: apical (0), subapical (1), dorsoapical (2), dorsal (3). One of the most distinguishing (and profoundly plesiomorphic) feature of the Neriidae is the apical/subapical placement of the antennal arista (Aczél, 1951; Steyskal, 1968; McAlpine, 1981). In his monography on American Neriidae, Aczél (1961) further subdivided this trait, discriminating those species for which the antenna is placed at the dorsoapical angle of the first flagellomere (state 2), and others have followed this distinction (Buck and Marshall, 2004; Sepúlveda *et al.,* 2013a; 2013b). This character was treated as ordered, given that the states represent discrete points of a positional cline.

105. Basal flagellomeres of arista: slender (0), thickened (1). Another of the synapomorphic characters of the neriid antenna is the enlarged configuration of the three (sometimes only two) most basal flagellomeres of the arista (Aczél, 1961).

106. Pubescence of antennal arista: absent (0), microscopically short (1), long (2). The degree of development of the pubescence of the antennal arista has been proposed as a key feature in the delineation of groups within the family (Aczél, 1961). Three major types of pilosity occur within the neriids: a completely naked arista, typical of the genera *Nerius* and *Glyphidops* (with some exceptions); an arista covered by microscopically short yet conspicuous pubescence, only visible at 30-40x magnification (e.g.: American branch of the *Eoneria*-group); or covered with long and fine hairs, that although varying in length, are always visible at naked eye or using lower magnifications. Among the analyzed Micropezidae the arista is pilose exclusively in the Eurybatinae (Aczél, 1959), while the extremely short and sparce pubescence of the included Cypselosomatidae (only visible at 200x magnification; McAlpine, 1966) was also codified as absent. This character was treated as ordered.

107. Pattern of pubescence of antennal arista: basally dense and apically sparse/naked (0), homogenously dense (1). Among species with a hairy arista, excluding the basal enlarged aristomeres which are always naked, the pubescence can either homogenously cover the structure, or vary in density along it. In this last case, the density of hairs is invariably reduced towards the apex of the arista.

108. General pattern of coloration of antennal arista: brown (0), white (1), basally white and apically brown (2). Aczél (1961) originally described two general patterns of coloration of the antennal arista among the Neotropical Neriidae. According to him, a group of genera present a dark brown arista, bare or covered with brown pubescence, while the other present a white arista, always covered in hair. However, he noted that in case the arista gets sparcely covered towards the apex (char. 107), this region can develop a brown coloration. Due to this last pattern being also common among Old World neriids (Aczél, 1954a; 1955a), it was taken as a third state. Although a few taxonomic descriptions acknowledge different colors for the arista and for the pubescence it carries, the modularity of these two characters is doubtful, and in most species both change in a congruent fashion. This character therefore simply describes the general coloration of the antennal arista as a whole.

109. Differentially pigmented region of antennal arista: absent (0), encompassing basally enlarged region (1), distad to enlarged region (2). Apart from the general coloration described previously (char. 108), most species of Neriidae present a small region of differentially pigmented region at the base of the antennal arista. This region attains a color that varies between testaceous yellow and dark brown, and stands out as a darker region in whitish aristae or as a bright region in brown ones. In most cases, the differentially pigmented zone is found at the base of the arista, where the enlarged aristomeres are (state 1), although in some members of the genus *Telostylinus* it is placed right after this enlarged region (state 2).

110. Size of buccal cavity: small (0), large (1). All groups included in the analysis have oval-shaped buccal cavities, of which that of Neriidae and Taeniapterinae is much bigger, with a total surface at least 3 times (and up to 10 times) larger than that of Micropezinae and Eurybatinae (Aczél, 1951).

111. Shape of prementum: as long as wide (0), slightly longer (1), considerably longer (2). The prementum of Micropezinae is approximately as long as wide, 1.25 – 1.5 times longer than wide in the Taeniapterinae and Eurybatinae, and becoming noticeably thin and elongated in the Neriidae, 2 – 3 longer than wide according to Aczél (1951: 495), although it can even be 4 times longer than wide in some *Chaetonerius* and *Telostylinus*. This character was treated as ordered.

112. Shape of clypeus: U-shaped (0), shield-shaped (1). Only in the Micropezinae one finds the small and U-shaped clypeus that is typical of the Muscomorpha (McAlpine, 1981), whereas in Neriidae, Taeniapterinae and Eurybatinae the clypeus is relatively large and shield-shaped (Aczél, 1951).

113. Lateral borders of the clypeus: thickening backwards (0), uniform (1), thickening forwards (2). Despite both Neriidae and Taeniapterinae having similarly shaped clypeus (char. 101), they differ in the structure of the lateral borders of the sclerite, which become strongly thicker posteriorly in Neriidae and anteriorly in Taeniapterniae (Aczél, 1951). The Micropezinae and Eurybatinae present an intermediate state, with the lateral borders of the clypeus being uniform. The character was treated as ordered.

114. Length of maxillary palpi: reaching clypeus' anterior margin (0), not reaching (1). Another feature shared between Neriidae and Taeniapterinae is the presence of elongated maxillary palpi (state 0), which almost invariantly reach the anterior margin of the clypeus. Micropezinae and Eurybatinae on the other hand, always present short maxillary palpi (state 0).

115. Shape of maxillary palpi: compressed (0), subcylindrical (1). Contrary to what happens with the length of the maxillary palpi (char. 103), the shape of this structure allies the Neriidae and Micropezinae, both of which have subcylindrical maxillary palpi, while it is always compressed in the Taeniapterinae (Aczél, 1951).

116. Configuration of anterior region of thorax: scutum and antepronotal ridge ending at level of postpronotal carina (0), scutum and antepronotal ridge ending beyond level of postpronotal carina (1). Among some species with an elongated thorax, the presutural scutellum and the antepronotal ridge have an anterior ending beyond the level of the postpronotal (humeral) carina. This was taken to be an apomorphic configuration of the thorax by Buck (2010), and the basis for his transfer of *Odontoloxozus peruanus* to the genus *Cerantichir*. However, as Sepúlveda *et al.* (2013a) noticed, this character is also shared by the genus *Longina*. Among Old World neriids it is also present in the species *Gymnonerius fuscus* (Aczél, 1955a: 35, Fig. 4).

117. Shape of katepisternum: as wide as high (0), higher than wide (1). The katepisternum of all Neotropical taxa with elongated thorax (except the genus *Nerius*) is about as wide as it is high, being higher than wide in the rest (Aczél, 1961), as well as in the entire Old World fauna.

118. Copulatory processes constituted by fifth sternite: absent (0), present (1). In general, most male micropezids present forceps-like processes of variable sizes and shapes, which arise from the sternite 5 (Steyskal, 1987b). These processes, which are involved in the copula, are absent from Neriidae and Cypselosomatidae, and from some genera of

Micropezinae and Taeniapterinae (Aczél, 1951), although present in those included in the analysis.

119. Postpronotal bristles: absent (0), hair-like (1), strong (2). This pair of bristles, also referred to as humeral bristles, were cited as absent from the Neriidae and the Mycropezidae by Aczél (1951: 486). However, he afterwards described the presence of this bristles as a synapomorphic condition of the genus *Odontoloxozus* (Aczél, 1961). Since then, other authors (Buck and Marshall, 2004; Sepúlveda *et al.*, 2013a) described the presence of small bristles on the postpronotal lobe for species of the genera *Cerantichir* and *Longina*. These were codified as reduced (state 1), with only *Od. longicornis* and *C. peruana*, former *Od. peruanus*, having strong (state 2) postpronotal bristles. The character was treated as ordered.

120. Proepisternal bristles: absent (0), one short and inconspicuous pair (1), one strong and spine-like pair (2), many (3). The presence of proepisternal bristles (referred to as propleural by Aczél) is an important feature in the taxonomy of the higher Diptera (McAlpine, 1981). Most species of the Australian-Oriental and Afrotropic regions present strong proepisternal bristles, contrasting with the inconspicuous pair present in Neotropical Neriidae, which is furthermore absent in many species. A patch of bristles in this region is common among the Taeniapterinae.

121. Anterior pair of notopleural bristles: absent (0), hair-like (1), equal/subequal to posterior pair (2). The posterior pair of notopleural bristles is invariantly present and strong in all species included in the analysis (Aczél, 1951). However, the anterior pair has a tendency towards reduction (Aczél, 1954a, 1961), being totally absent (state 0) in a few genera of the Australian-Oriental region (e.g.: *Telostylinus*, *Rhoptrum*, etc.); hair-like (state 1) in many genera of the Neotropics (e.g.: *Glyphidops*, *Nerius*, etc.); and as strong as the posterior pair, although generally shorter in length, in genera from both the Neotropical and Afrotropical regions (e.g.: *Longina*, *Eoneria*, *Chaetonerius*, etc.). The character was treated as ordered.

122. Suprahumeral protuberances: absent (0), present (1). Aczél (1961) acknowledged a close relationship between *Cerantichir peruana* and *Longina abdominalis* given the shared presence of a suprahumeral protuberance on which a strong spine sits. Later, Buck and Marshall (2004) described the presence of this structure in *L. anguliceps* as well. They furthermore discussed the presence of an enlarged first dorsocentral (lacking a protuberance) as a possible synapomorphy of a *Longina* - *Cerantichir* clade. However, suprahumeral bristles are common in many genera, and their strength is very polymorphic within species, so only the presence of the protuberance was codified.

123. Presutural dorsocentral bristle: absent (0), present (1). The Neriidae present a wide variety of dorsocentral bristle patterns, ranging from only one to as much as nine (six if one

leaves out the small bristles before the presutural pair in *Eoneria* and *Eoloxozus*, char. 127). The presence of such "complete series of dorsocentrals" (D. K. McAlpine, 1974: 232) common to some Neriidae, Cypselosomatidae and Pseudopomyzidae, have been regarded as plesiomorphic with respect to the reduced pattern present in most micropezids. Dorsocentral bristles were codified separately based on their position (chars. 123 to 127), following the same principle adopted for the fronto-orbital bristles (chars. 88, 89). This character refers to the bristle present before the transverse suture.

124. Postsutural dorsocentral bristle: absent (0), present (1). See char. 117.

125. Prescutellar dorsocentral bristle: absent (0), present (1). This bristle pair, situated just before the scutellum, is only absent in the genera *Micropeza* (Aczél, 1951).

126. Number of extra dorsocentral bristles posterior to transverse suture: range 0 - 2. Between the postsutural and prescutellar bristle pairs, one or two more pairs of bristles may be present, as large as the former. In case one bristle pair is present, this one is situated in the middle of the postsutural area of the scutellum (characteristic of *Chaetonerius*); if two are present, these divide this area into three more or less equally spaced regions (characteristic of *Eoneria* and *Eoloxozus*). The character was treated as ordered.

127. Number of extra dorsocentral bristles anterior to transverse suture: range 0 - 3. A patch of small dorsocentral bristles are present in the genera *Eoneria* and *Eoloxozus* in front of the prescutellar pair. This bristles are short yet strong and bristle-like, unlike the inconspicuous hairs that are present in many other species of neriids, as for example *Longina anguliceps* (Buck and Marshall, 2004) and some *Telostylinus* species (Aczél, 1959). The character was treated as ordered.

128. Katepisternal bristle: absent (0), present (1).

129. Horizontal line of bristles in posterior margin of katepisternum: absent (0), present (1). The presence of this line of small bristles (referred to as hypopleural by Hennig, 1934; and sternopleural by Aczél, 1951) is restricted to the Taeniapterinae and was taken to be a symplesiomorphic character by Aczél (1951), who did not considered them to be homologous to the single katepisternal bristle (char. 128) of the other Nerioidea, given that they stand at different positions. In the Eurybatinae *Crosa yapensis*, he found what he considered to be an intermediate form, having well differentiated katepisternal bristles (char. 128) while retaining a reduced version of the taeniopterine "fan of erect bristle-like hairs" (Aczél, 1959: 89; D. K. McAlpine, 1974).

130. Supra-alar bristle: absent (0), present (1). Unlike the post-alar (*pa*) bristle, which is always present, the supra-alar (*sa*) bristle is absent in *Glyphidops filosus* and *G. ochreus* (Aczél, 1961).

131. Relative length of sa and pa bristles: approximately equal (0), pa larger (1). Although the *pa* bristles are generally a little bit longer than the *sa* bristles, only in certain genera of Neotropical Neriidae do the *pa* bristles attain lengths considerably longer (30 - 50%) than that of the *sa* bristles.

132. Non-apical scutellar bristles: absent/vestigial (0), long (1). In the scutellum there is always an apical bristle pair, which is also generally the longest bristle pair on the thorax. However, in many species an extra bristle may be present. This can adopt the shape of a vestigial bristle, very small and hair-like (whose presence is generally polymorphic within species, and was therefore not codified as a separate state), or be strong and well-developed, although always shorter that the apical pair. This long non-apical scutellar pair is exclusively found in the Telostylinae and in *Gymnonerius fuscus*, all of which are described as having 2 pairs of scutellar bristles (Aczél, 1954a; 1955a), as well as in some Cypselosomatidae (D. K. McAlpine, 1966).

133. Protuberances on apical scutellar bristles: absent (0), present (1). These protuberances represent projections of the surface of the scutellum, in the shape of a cone or semisphere, in which the apical bristles stand, sometimes even longer than the bristle itself (Aczél, 1961).

134. Protuberances on hind notopleural bristles: absent (0), present (1).

135. Coloration of katatergite: brown (0), yellow (1). Aczél (1961) discussed the presence of an intensely yellowish coloration on the katatergite (or inferior pleurotergite) of some species of the genus *Glyphidops*, proposing that this character may be of taxonomic importance (Aczél, 1961). Among Old World neriids, the horizontal yellow vitta of the thorax is broader than in Neotropical species, sometimes including the katatergite. However, among many species in which this horizontal vitta encompasses the entire postnotum, the katatergite is nonetheless dark brown in colour, showing that the coloration of this region of the thorax is independent of the rest. Consequently, the presence of a yellow katatergite was also acknowledged for some species of the genus *Telostylinus* and *Chaetonerius* (Aczél 1954a, 1959).

136. Central stripe of pruinosity on mesonotum: absent (0), present (1). A very complex and important feature of the neriid thorax is the pattern of pigmentation on the mesonotum (Aczél, 1961). The vast majority of species (except the genera *Gymnonerius*, *Teloneria* and *Telostylus*) have a central stripe covered by intense clear dusting, covering approximately the central third of the mesonotum.

137. Color of pruinosity of mesonotum: yellow (0), gray (1). Most Neotropical neriids have the central region of the mesonotum (char. 136) covered with whitish to grayish dusting, except in the genera *Cerantichir* and in *Glyphidops limbata*. In these and all species of the Old World, the pruinosity of the mesonotum is distinctly yellow.

138. Longitudinal dustless medial line in mesonotum: absent (0), present (1). The central stripe of pruinosity of the mesonotum (char. 136) may be interrupted by a medial line which lacks the clear dusting and therefore presents the same coloration as the lateral regions of the mesonotum.

139. Position of dustless medial line: anterior to transverse suture (0), posterior to transverse suture (1), complete (2).

140. Margins of dustless medial line: converging backwards (0), straight (1), diverging backwards (2). Among some species of the genus *Glyphidops*, the dustless medial line of the mesonotum varies in shape along the longitudinal axis, sometimes becoming narrower posteriorly (state 0), sometimes becoming wider (state 2). Among all other species that posses this medial line, the borders of the line are straight, therefore not changing shape along the thorax.

141. Longitudinal dustless lateral lines in mesonotum: absent (0), present (1). A different pattern of pigmentation in the mesonotum results from the presence of two dustless lateral lines in the pruinose central stripe. This leads to a three-stripe pattern if the medial line is absent (char. 138), as in the *Telostylinus* or some *Nerius*, or a four-stripe pattern if the medial line is present, as in *Antillonerius*.

142. Position of dustless lateral lines: anterior to transverse suture (0), posterior to transverse suture (1), complete (2).

143. Central yellowish vitta in scutellum: absent (0), present (1). Among most species, the central region of the scutellum has a yellowish vitta, which presents diverse degrees of wideness. Among species with yellowish mesonotal pruinosity (char. 137) this vitta develops as a prolongation of the posterior region of the central stripe, but it is also present in many species with grayish pruinosity.

144. Lateral pruinosity in scutellum: absent (0), present (1). As has been described before, some species with grayish mesonotal pruinosity nonetheless develop a yellow vitta in the center of the scutellum. However, the two lateral stripes of grayish pruinosity among these species also continue into the lateral sides of the scutellum (state 1), which are therefore very different in color from the lateral dark brown regions of the mesonotum.

145. Dustless regions at the base of bristles and hairs: absent (0), present (1). Aczél (1961) drew attention to the similarities between *Eoloxozus sabroskyi* and *Odontoloxozus longicornis*, a resemblance based on many distinguishing characters, one of which is the insertion of the bristles and hairs of the pleurae, mesonotum and preabdominal tergites on dustless dots. This character is not present in any other species.

146. Anal lobe and alula: reduced (0), developed (1). The proximal region of the wing, including the anal lobe and the alula, is considerably reduced in the species of

Micropezinae, resulting in a quite different wing morphology that significantly narrows towards the wing stalk (see Aczél, 1951: lamina III). The wings of Cypselosomatidae and Neriidae retain the plesiomorphic well-developed alula and anal lobe, while the same wing configuration is interpreted as a consequence of a secondarily broadened wing-base in some Taeniapterinae and Eurybatinae (D. K. McAlpine, 1974).

147. Crossvein bm-cu: absent (0), vestigial/weak (1), conspicuous (2). The crossvein bm-cu is according to Aczél (1951) the only wing vein whose presence is variable within the Nerioidea. Among the groups included in the analysis, this vein is completely absent in the Micropezinae and in *Cypselosoma astralis*, reduced or vestigial in many Taeniapterinae and all Neriidae, and conspicuous only in some genera of Taeniapterinae (e.g.: *Taeniaptera*) and Eurybatinae (Aczél, 1959). The character was treated as ordered.

148. First costal section: absent (0), present (1). The apical parts of the Sc and R_1 veins are divergent in the Taeniapterinae and in all Neriidae, determining the first costal section (char. 39, 40) or pterostigma according to Aczél (1951, 1954b). On the other hand, the apical parts of these veins run parallel, never diverging, in the Micropezinae, the eurybatine fly *Crosa yapensis* and the Cypselosomatidae (state 0).

149. Longitudinal veins R_{4+5} and M_{1+2}: fusing at or near wing apex (0), not fusing (1). Among all families of Nerioidea included in the analysis, the veins R_{4+5} and M_{1+2} are convergent, resulting in an r_{4+5} cell that narrows towards wing apex. In some genera of Micropezidae (e.g.: *Taeniaptera*, *Micropeza*) these veins fuse with each other, sometimes at the level of costal vein, sometimes before it, resulting in an apical crossvein (Aczél, 1951). These species therefore lack a fourth costal section (chars. 45 and 46).

150. Subcostal break: absent (0), present (1). The costal vein is more or less strongly interrupted just proximal to where the subcostal vein joins it. This character was first reported by Aczél (1951) for all revised species of Neriidae except for the genus *Longina*. The character was further used as evidence to elevate Neriidae to the level of family (McAlpine, 1974), and latter considered as a groundplan plesiomorphy of the Nerioidea, given its shared presence in Pseudopomyzidae and Cypselosomatidae (McAlpine, 2000). Aczél (1961: 270) stated that the subcostal break was present in most genera of neriids, and some trace of it could be found in all studied specimens except for the genera *Longina*, for which it was taken to be absent.

151. Direction of vein CuA_2: distal (0), proximal (1). Among the Taeniapterinae and Eurybatinae, the vein CuA_2 meets the A_1 vein at a position that is distal with respect to the point where it originates from CuA_1, therefore having a distal direction of development (state 0). This direction is inverted in the rest of the included taxa (state 1), with the CuA_2 vein developing towards the wing base (Aczél, 1951; D. K. McAlpine, 1974).

152. Length of A_1+CuA_2 vein: reaching posterior wing margin (0), short and not reaching posterior margin (1), vestigial (2). The vein A_1+CuA_2 (pedicel of the anal cell in Aczél, 1951) is long among the Taeniapterinae and the Micropezinae, attaining the posterior margin of the wing (state 0). On the other hand, among the Neriidae, Eurybatinae and Cypselosomatidae it is always short (Aczél, 1951; D. K. McAlpine, 1966), generally having about half the length it would take to attain hind wing margin (state 1), except in the genera *Teloneria*, in which it is "vanishingly small" (Aczél, 1955a: 32 and plate 2), almost absent (state 2). The character was treated as ordered.

153. Shape of crossvein dm-cu: straight to slightly convex (0), strongly convex (1), undulated (2). The shape of the crossvein dm-cu is typically straight (state 0), sometimes a little bit outwardly convex, in most species studied. Only does this vein attain a relatively high degree of curvature in the genus *Nerius* (Aczél, 1961) and in *Gymnonerius fuscus* (Aczél, 1955a). Both in *Eoneria* and in *Eoloxozus*, and especially in the later, the vein is undulated (state 2), adopting an S-shape (Aczél, 1961; Sepúlveda *et al.*, 2013b). This feature, although never specified as such, is also shared by *Odontoloxozus longicornis* (see Buck, 2010).

154. Angle of crossvein dm-cu: transverse (0), slightly oblique (1), strongly oblique (2). The Neotropical genera *Loxozus*, *Odontoloxozus* and *Eoloxozus* have a dm-cu crossvein that is placed strongly oblique (state 2) with respect to the longitudinal veins, almost moving parallel to posterior wing margin (Aczél, 1961). The rest of the species, have a vein that is either normally placed (state 0), or only slightly oblique (state 1). This character was treated as ordered.

155. Infuscation of crossvein dm-cu: absent (0), present (1). Only one included species, *Telostylinus zonalis*, presents an isolated brownish infuscation in the dm-cu cross vein (Aczél, 1954a). However, both species that present supernumerary crossveins (char. 156) also present a brown lined dm-cu vein, as part of a broader pattern of infuscation (Aczél, 1961; Sepúlveda *et al.*, 2013b).

156. Supernumerary crossveins: absent (0), present (1). Only three species within the Nerioidea have been described with supernumerary crossveins arising from the R_{2+3} and M_{1+2} veins, all of which are neriids (Aczél, 1961). Two of these species, *Eoneria maldonadoi* and *Glyphidops pluricellata*, have been included in the analysis.

157. Bare areas of wing: absent (0), small (1), large (2). Buck and Marshall (2004) described the absence of microtrichia in most of the basal region of the wings of the genus *Longina* (state 2), extending beyond level of r-m cross vein in the r_1, r_{2+3} and r_{4+5} cells, and proposed it as a putative synapomorphy of the genus. However, they also discussed the presence of less pronounced bare areas in *Cerantichir enderleini* (state 1). The rest of

Neotropical Neriidae, as well as all Old World species revised, have completely microtrichose wings (state 0). The character was treated as ordered.

158. Number of bristles in basicosta: range 0 - 3. The number of bristles in the wide basicosta of the Nerioidea varies considerably. It is always 3 in the Taeniapterinae, 0 in the Micropezinae and Eurybatinae, and a number between 0 and 2 in the Neriidae. Among the last, the tendency towards reduction is stronger among Neotropical neriids, and generally circumscribed to the *Nerius*-group, while the *Eoneria*-group and the Telostylinae retain the presence of 2 bristles. The character was treated as ordered.

159. Ciliae on upper calypter: absent (0), short (1), long (2). The general shape of the calypteres does not vary within the Neriidae, being the upper calypter rounded or oval and the inferior one reduced to a linear strip (or frenulum squamulare). However, some degree of variation in the length of the marginal hairs present in the calypteres can be found. In general, the upper calypter has long hairs (state 2), while the lower one has short ones (state 1), or they are completely absent (state 0). However, in the species *Longina anguliceps* this tendency is reversed (Buck and Marshall, 2004), while in *Glyphidops ochreus* both calypteres have equally long ciliae (Aczél, 1961). This character was treated as ordered.

160. Ciliae on lower calypter: absent (0), short (1), long (2). See char. 159. This character was treated as ordered.

161. Shape of halteres: lobulose capitulum and short pedicel (0), compact capitulum and long pedicel (1). The morphology of the halters is modified only in the Taeniapterinae (state 1), which posses a small and compact capitulum (or knob) that is supported by a relatively longer pedicel (or stem).

162. Antero and posteroventral row of spines on fore femur: absent (0), reduced and hair-like (1), present (2). Another of the distinguishing features of the Neriidae is the presence of two rows of conspicuous spines on the ventral region of the femora (Aczél, 1951; McAlpine, 1981; Buck, 2010). These bristles are nonetheless not present in all species, having been lost (state 0) in some species of *Telostylinus*, *Teloneria* and *Antillonerius*. In some other species, these rows only bear weak and hair-like bristles (state 1, present also in the Cypselosomatidae and Eurybatinae), a pattern very common for Old World species, in contrast with the stout spines of many Neotropical genera, which often sit on cylindrico-conical protuberances. This character is profoundly sexually dimorphic, being involved in the male-male aggressive displays (Eberhard, 1998), and was therefore only codified for male specimens of each species. The presence of spines on mid and hind femora covaries with the strength of their development on the fore femur, and where therefore also excluded from the data. The character was treated as continuous.

163. Spinules on fore tibia: absent (0), present (1). Among some few species of neriids the antero and posteroventral rows of spines of the fore femur (char. 162) continue into the tibia, in the shape of smaller spinules.

164. Shape of dorsal face of mid femur: concave (0), straight to slightly convex (1). The dorsal face of the mid femur is only concave in the *Logina* - *Odontoloxozus* - *Cerantichir* clade, giving this segment an arched shape (Aczél, 1961). In the rest of the included species, the mid femur presents a straight to convex shape, being more strongly curved on species with short femora. This character also determines the place at which the femur attains its largest width, a character often described by Aczél (1955a, 1961), being always apical in species with concave mid femur and medial to distomedial in the rest.

165. Club-shaped thickening in apex of male fore tibia: absent (0), present (1). The presence of a thickened apex of the fore tibiae was first described by Hennig (1937) as a characteristic of the genus *Rhoptrum*. Aczél (1954a) later concluded that this character is sexually dimorphic, being only present in the male sex.

166. Number of dorsal bristles in male procoxae: range 0 - 8. All neriids have 1-2 apical bristles on the procoxae. However, the number of dorsal bristles present is extremely variable, and according to Aczél (1961) may be of taxonomic value. This character presents an enormous degree of intraspecific variation, with larger specimens always having more dorsal procoxal bristles than smaller ones. However, if this variability is included in the analysis, phylogenetic signal may be retrieved. The character also presents a high degree of sexual dimorphism, and was therefore codified separately for males and females (char. 167). Characters 166 and 167 were treated as ordered.

167. Number of dorsal bristles in female procoxae: range 0 - 6. See char. 166.

168. Number of lateral bristles in mesocoxae: range 0 - 3. The number of lateral bristles (that is, excluding the apical ones) on the meso and metacoxae is very constant within species and has a very interesting pattern of variation among taxa (Aczél, 1961). Characters 167 and 168 were treated as ordered.

169. Number of lateral bristles in metacoxae: range 1 - 3. See char. 168.

170. Small dorsal bristles in mid and hind tibiae: absent (0), present (1). All along the length of the mid and hind tibiae of the studied Micropezidae and Taenipaterinae there is a row of tiny black bristles, that is always absent in Neriidae (Aczél, 1951).

171. Color of procoxa: yellow (0), brown (1). Among all species studied, the coxae can present the same light to dark brown coloration of the pleurae (state 0), or they can be yellowish, clearly different from the surrounding coloration (state 1). However, the color of the meso and metacoxae is always the same (char. 172), whereas the color of the procoxa

varies independently. Therefore, the color of the procoxae was divided into two separate characters.

172. Color of meso and metacoxae: yellow (0), brown (1). See char. 171.

173. Color of fore femur: yellow (0), brown (1). The coloration of the legs is an important taxonomic character, being commonly used in several generic keys (Aczél, 1954b, 1961). Two attributes of the femora coloration were used, the base color and the presence of bands or rings of differential coloration. Since the positional homology of these rings was often difficult to establish, they were divided into three categories: basal (char. 174) and apical (char. 176) stripes, usually broad and always attaining the base or apex of the corresponding femur; and the number of central rings (char. 175), thinner, usually lacking a conspicuous margin, and restricted to the central region of the femur. Furthermore, since differences in patterning of the three femora are not uncommon, these four character were codified separately for each femur (chars. 173-184).

174. Basal stripe in fore femur: absent (0), present (1). See char. 173.

175. Number of central rings in fore femur: range 0 - 3. See char. 173. The character was treated as ordered.

176. Apical stripe in fore femur: absent (0), present (1). See char. 173.

177. Color of mid femur: yellow (0), brown (1). See char. 173.

178. Basal stripe in mid femur: absent (0), present (1). See char. 173.

179. Number of central rings in mid femur: range 0 - 3. See char. 173. The character was treated as ordered.

180. Apical stripe in mid femur: absent (0), present (1). See char. 173.

181. Color of hind femur: yellow (0), brown (1). See char. 173.

182. Basal stripe in hind femur: absent (0), present (1). See char. 173.

183. Number of central rings in hind femur: range 0 - 3. See char. 173. The character was treated as ordered.

184. Apical stripe in hind femur: absent (0), present (1). See char. 173.

185. Color of tibiae: yellow (0), brown (1). Unlike what happens with the femora, the color of the tibiae is very homogenous, and was therefore merged into a single character. This character nonetheless has shown to vary independently with respect to the coloration of the femora.

186. Apical dark stripe in tibiae: absent (0), present (1). Most species have a pattern of tibial coloration in which it becomes darker towards the tip (state 1). This happens both on yellowish or brown tibiae (char. 185). Among other species, the tibiae are homogenously dark brown (state 0).

187. Preabdominal median longitudinal vita: absence (0), presence (1). Some species of the genera *Telostylinus* and *Chaetonerius* have been inconsistently described by Aczél as having brown or yellow background coloration in the preabdominal tergites (see for example Aczél, 1954b and 1954c). Here, we adopt the hypothesis that the background coloration of all neriid species (as well as the included Micropezidae and Cypselosomatidae) is brown, since even among genera described as having yellow background coloration there are species with entirely brown preabdominal tergites (e.g.: *Ch. niger* and *T. papuanus*). Therefore, the pigmentation pattern of all neriid species can be described by the presence/absence of 3 types of longitudinal vittae: a median thin vitta, two wide lateral vittae (char. 188) and two marginal vittae, which are very wide and usually divide the preabdomen into thirds (char. 189); all of these markings are whitish to yellowish in color. When defined in this way, the presence of the median longitudinal vitta is exclusive of the *Eoneria – Eoloxozus* clade (although it is also present as polymorphism in some *Glyphidops*), while the lateral vittae are only found among those species of *Chaetonerius* and *Telostylinus* that were described as having yellow tergites.

188. Preabdominal lateral longitudinal vittae: absent (0), present (1). See char. 187.

189. Preabdominal marginal longitudinal vittae: absent (0), present (1). See char. 187.

190. Shape of epandrium: flattened to hemispheric (0), semicylindrical (1). The epandrium of the Micropezidae is generally shorter than that of the Neriidae (char. 69), yet its shape is also quite different, with the epandrium of the Micropezidae lacking the clearly semicylindrical shape common to all Neriidae (state 1). The shape of the epandrium of Cypselosomatidae is much closer to that of the Neriidae than to the rest of included taxa (McAlpine, 1974), and was therefore also codified as semicylindrical.

191. Lateral border of semicylindrical epandria: narrowing at the center (0), nearly straight (1). The only significant difference in structure in the genitalia of all known neriid species is present in the epandrium of three *Chaetonerius* species: *Ch. apicalis*, *Ch. collarti* and *Ch. ghesquièrei* (known as the *apicalis*-group, Aczél, 1954b). This are have a modified, bulky epandrium (state 1), whose lateral margins run more or less straightly from base to apex, and with enlarged and folded lateroapical regions. The rest of the family has a more slender epandrium, clearly narrowing at the center, with a distinctive hour-glass shape.

192. Width of ventral furrow of epandrium: thin (0), intermediate (1), broad (2). The entire ventral surface of the epandrium presents a longitudinal cleft in which the folded aedeagus reposes. This “genital pouch” (Aczél, 1961: 272) is thin and linear in the Taeniapterinae,

intermediately wide in Tylinae and Eurybatinae, and attaining its maximum broadness among Neriidae (Aczél, 1951). The character was treated as ordered.

193. Surstyli: absent (0), vestigial/reduced (1), big (2). The presence and development of the surstyli are features that show significant differences among groups of Nerioidea (Aczél, 1951). This lobes are completely absent in the Taenipaterinae, while present in other groups included in the analysis. However, the surstyli of Neriidae are much reduced, almost vestigial, while this structure is well-developed in the Tylinae, Eurybatinae and Cypselosomatidae. The character was treated as ordered.

194. Shape of ventral face of ovipositor: concave (0), convex (1). The ventral surface of the ovipositor is concave and sunken in most species studied. However, Aczél (1951) noted that the ovipositor of the genera *Eoneria* and *Longina* was convex. Although this character was not mentioned in subsequent descriptions, the ventral surface of the ovipositor proved to be concave in all other revised species. The state of *L. anguliceps* and *L. semialba*, not known by Aczél, were codified as missing data, since the shape of the female terminalia was not specified in their taxonomic descriptions.

Appendix S2: Material examined

All revised material is deposited in the Colección Entomológica, Instituto y Fundación Miguel Lillo, Tucumán, Argentina (curator Emilia Pérez). Acronyms are: BPBM, Bernice P. Bishop Museum, Honolulu; IRSN, Institut Royal des Sciences Naturelles de Belgique, Brussels; KUEC, Kyushu University, Fukuoka; MRAC, Musée royal de l'Afrique Centrale, Tervuren.

Cerantichir peruana – **Peru**: Valle Chanchamayo, 25.II.1929, L. Weyrauch, 1 male.

Chaetonerius apicalis – **Democratic Republic of the Congo**: Eala, I.1935 (ex IRSN, I. G. 10.482), J. Ghesquière, 1 male. Eala, VII.1935 (ex IRSN, I. G. 10.482), J. Ghesquière, 2 male and 1 female.

Chaetonerius brachialis – **Democratic Republic of the Congo**: Bambesa, 10.XII.1938 (ex IRSN, I. G. 12.234), J. Vrydagh, 1 male.

Chaetonerius collarti – **Democratic Republic of the Congo**: Eala, III.1936 (ex IRSN, I. G. 10.482), J. Ghesquière, 2 males (paratypes).

Chaetonerius inermis – **Indonesia**: West Java, Idjen plateau, Blawan, 950 m, VI.1924, K. W. Dammerman, 1 male and 1 female.

Chaetonerius latifemur – **Democratic Republic of the Congo**: Eala, 12.II.1935 (Ex IRSN, I .G. 10.482), J. Ghesquière, 1 female. Eala, III.1935 (ex IRSN, I .G. 10.482), J. Ghesquière, 1 male and 1 female. Eala, 7.IV.1935 (ex IRSN, I .G. 10.482), J. Ghesquière, 2 females. Eala, VII.1935 (ex IRSN, I .G. 10.482), J. Ghesquière, 1 male. Eala, 6.V.1936 (ex IRSN, I .G. 10.482), J. Ghesquière, 1 male. Rutshuru, XI.1937 (ex IRSN, I .G. 10.482), J. Ghesquière, 2 males. Rutshuru, 4.XII.1937 (ex IRSN, I .G. 10.482), J. Ghesquière, 1 female.

Cheatonerius niger – **Democratic Republic of the Congo**: Rutshuru, Kilinga, VI.1936 (ex MRAC), L. Lippens, 1 female. Kivu, Nzombe Amont, 200 m near Mwana, 1952 (ex MRAC), A. Froidebise, 1 male.

Chaetonerius perstriatus – **Democratic Republic of the Congo**: Upemba National Park, Kalule-Nord, near Klamalwa, 3-4.III.1949 (Mis. G. F. de Witte 2401a), R. Bowa, 1 male and 1 sex unknown (postabdomen missing; paratypes, labeled *Chaetonerius wittei*).

Eoloxozus sabroskyi – **Peru**: Quebrada Verde, Lurín, 10.X.1950, L. Weyrauch, 1 female.

Eoneria blanchardi – **Argentina**: Chacho, II.1974, 1 male. Corrientes, I.1950, D'Angelo, 1 female (paratype).

Eoneria maldonadoi – **Argentina**: La Rioja, XI.1952, 1 female (holotype). Catamarca, Andalhuallas, 2000 m, 19.I.1968, R. Goldbach , A. L. Terán and H. Willink, 1 female. "Vinagre D, La S., 9.IX.1965" (?), 1 female.

Glyphidops neuter – **Argentina**: Dpto. San Pedro, Carumbé, 28.I-10.III.1963, R. Golbach, 1 female (labeled *Oncopsia neutra*).

Gymnonerius fuscus – **Indonesia**: Mentawai Is., Sipora, 14.X.1924, H. M. Karny, 1 female. North Sumatra, Medan, Saengei Krio, IV.1928, J.C. v.d. Meer Mohr, 1 female. Western Java, Djampang-Tengah, Goenoeng Tjisoeroe, 600-800 m, III.1933, M. E. Walsh, 2 males and 1 female. Western Java, Djampang-Tengah, Goenoeng Tjisoeroe, 600-800 m, IX.1933, M. E. Walsh, 1 male.

Longina abdominalis – **Argentina**: Misiones, Puerto Bemberg, 12-29.I.1945, H. Willink and R. Golbach, 1 male (labeled *Longina peletieri*). **Paraguay**: Caaguazú, Paso Yobai, 280 m, VI.1951, J. Foerster, 3 males (2 labeled *Longina vittata*, 1 labeled *Longina peletieri*). Dpto. San Pedro, Carumbé, 28.I-10.III.1963, R. Golbach, 8 males and 3 females.

Nerius czernyi – **Argentina**: Misiones, Puerto Aguirre, 19.I.1931, K. Hayward, 1 female (holotype).

Nerius pilifer – **Argentina**: Misiones, Puerto Bemberg, 12-29.I.1945, K. Hayward, H. Willink and R. Golbach, 3 males and 3 female. Tucumán, Burruyacú, Villa Padre Monti, 17.I-7.II.1948, R. Golbach, 1 female. Jujuy, Caimancito, 26.V.1949, N. Kusnecow. Jujuy, "Monrés, La-Minb" (?), 11.II.1951, 3 males. Salta, Orán, Abra Grande, 10.I-28.II-1967, R. Golbach, 1 female. **Bolivia**: Santa Cruz, El Cidral, 1-28.I.1962, R. Golbach, 1 female. **Brasil**: São Paulo, Paraná River, Porto Cabral, 1-25.IV.1944, M. Carrera and E. Dente, 2 males. São Paulo, Paraná River, Porto Cabral, 1-25.IV.1944, M. Carrera, 1 male and 3 females. São Paulo, Paraná River, Porto Cabral, 20-31.III.1944, Trav. Fo., M. Carrera and E. Dente, 1 male and 4 females. **Paraguay**: Dpto. San Pedro, Carumbé, 28.I-10.III.1963, R. Golbach, 1 male and 1 female.

Paranerius fibulatus – **Indonesia**: New Guinea, Papua, Araucaria Camp, 800 m, 25.III.1939 (Archbold's Netherland Indian – American New Guinea Expedition), L. J. Toxopeus, 1 male.

Rhoptrum annulipes – **Indonesia**: New Guinea, Papua, Bernhard Camp, 50 m, 12.IX.1938 (Archbold's Netherland Indian – American New Guinea Expedition), J. Olthof, 3 males and 2 females. New Guinea, Papua, Araucaria Camp, 800 m, 7.III.1939 (Archbold's Netherland Indian – American New Guinea Expedition), L. J. Toxopeus, 1 female.

Rhoptrum lieftincki – **Indonesia**: New Guinea, Papua, Rattan Camp, 1500 m, 12.III.1939 (Archbold's Netherland Indian – American New Guinea Expedition), L. J. Toxopeus, 1 male (paratype). New Guinea, Papua, Araucaria Camp, 800 m, 20.II.1939 (Archbold's Netherland Indian – American New Guinea Expedition), L. J. Toxopeus, 1 female (paratype).

Telostylinus gressitti – **Federal States of Micronesia**: Faraulep atoll, Faraulep Is., 21.-.1952, N. Krauss, 1 male (paratype). Ifaluk atoll, Ifaluk Is., 4.XI.1953, M. Bates, 1 male (paratype). **Palau**: Babelthuap Is., Ulimang, 10.XII.1947, H. S. Dybas, 1 female (paratype). Peleliu Is., Mt. Amiangel, 22.XII.1952, J. L. Gressitt, 1 male (paratype). Koror, 18-20.IV.1955, J. W. Beardeley, 1 female (paratype).

Telosylinus lineolatus – **Federal States of Micronesia**: Kusaie, Mutunlik, 22 m, 15.II.1953, J. F. G. Clarke, 1 female. **Indonesia**: Western Java, Bandoeng, 700 m, 17.III.1940, J. Olthof, 2 males.

Telostylinus longicoxa – **Federal States of Micronesia**: Kusaie, Lelu Is., Mt. Fenkol, 30.I.1936, Z. Ono, 1 male (ex BPBM).Truk (Chuuk) State, Weno Is., S. Valley, Mt. Tonaachau, 4.IV.1949, R. W. L. Potts, 1 male. Pohnpei (Ponape) Is., Kolonia, VI-IX.1950, P. A. Adams, 1 male. Kusaie, Mutunlik, 22 m, 15.II.1953, J. F. G. Clarke, 1 female. Truk (Chuuk) State, Weno Is, Mt. Teroken, Nantaruil, J. L. Gressitt, 1 female. **Marshall Islands:** Arno Atoll, Ine Is., 28.VII.1950, I. La Rivers, 1 male. Namu atoll, Namu Is., 24.X.1953, J. W. Beardaley, 1 female. Lae atoll, Lae Is., 24.X.1953, J. W. Beardaley, 1 female. **Northern Mariana Islands**: Saipan Is., Matansha-Calabera, 3.V.1940, Yasumatsu and Yoshimura, 1 male (ex KUEC).

Telostylinus papuanus – **Indonesia**: New Guinea, Papua, Bernhard Camp, 50 m, 7.XI.1938 (Archbold's Netherland Indian – American New Guinea Expedition), J. Olthof , 1 female. New Guinea, Papua, Sigi Camp, 1500 m, 2.II.1939 (Archbold's Netherland Indian – American New Guinea Expedition), L. J. Toxopeus, 1 female. New Guinea, Papua, Rattan Camp, 1200 m, 4.II.1939 (Archbold's Netherland Indian – American New Guinea Expedition), L. J. Toxopeus, 2 males.

Telostylinus ponapensis – **Federal States of Micronesia**: Pohnpei (Ponape) Is., southeast of Nanpohnmal, 70 m, 11.I.1953, 1 female (paratype). Pohnpei (Ponape) Is., Mt. Tamatamansakir, 180 m, 17.I.1953, J. L. Gressitt, 1 male (paratype).

Telostylinus spinicoxa – **Indonesia**: New Guinea, Papua, Jayapura (ex Hollandia), VII.1938 (Archbold's Netherland Indian – American New Guinea Expedition), L. J. Toxopeus, 1 male (paratype).

Telostylinus yapensis – **Federal States of Micronesia**: Yap, Yap Is., Mt. Matade, 95 m, 1.XII.1952, J. L. Gressitt, 1 female (paratype). Yap, Yap. Is., 1952, N. L. H. Krauss, 1 female.

Appendix S3: Scripts

SPR distance tuning (SPRsensitivity.run)

The following script generates 100 Wagner trees, groups them in pairs, and calculates the same 50 weighted SPR distances under a variety of combinations of the two parameters that determine its efficiency in finding the shortest distance, *i.e.* number of replicates and levels of stratification. Results are saved to a comma-delimited CSV file which can be opened directly with MS-Excel. The weight assigned to moves can be easily modified (or eliminated for unweighted SPR distances), as well as the values of the parameters explored.

```
macro = ;

mult 100 = wagner keepall ;

var =
     + numrepl
     + numstrat
     + sprmoves
     + tree1
     + tree2 ;

sprdiff [3 ;

report - ;
silent = file ;
log SPRdistances.csv ;
silent - file ; quote Tree1, Tree2, Stratifications, Replicates,
Movements ; silent = file ;

loop 0 30
     set numstrat #1 ;

     loop 1 20
          set numrepl #2 * 1000 ;

          loop 0 ntrees
               if ( #3 < 50 )
                    set tree1 #3 * 2 ;
                    set tree2 #3 * 2 + 1 ;
                    if ( #1 == 0 )
                         set sprmoves sprdiff [ 'tree1' 'tree2'
'numrepl' ] ;
                    else
```

```
                              set sprmoves sprdiff [ 'tree1' 'tree2'
'numrepl'x'numstrat' ] ;
                         end
                    silent - file ; quote 'tree1', 'tree2',
'numstrat', 'numrepl', 'sprmoves' ; silent = file ;
                    else
                         continue ;
                    end
               stop
          stop
stop

proc/;
```

Collapsing of continuous character (collapsing.run)

This script allows collapsing couples of continuous characters into a single one without modifying the original matrix. To do so, the smallest and largest values of both are stored as a range on one of the characters, with the other becoming inactivated. In its present form, the script does so only for contiguous couples of characters, but can be modified for different matrix configurations. The number of characters being collapsed can be modified easily by stating a list of exceptions (`collexceptions.txt`) that is read by the script, and the number of exceptions has to be stated as argument.

```
macro = ;

var =
     + exceptions [%1]
     + numExceptions
     + numcont
     + min
     + max
     + exception
     + continueLoop
     + nextLoop
     + secondchar ;

proc collexceptions2.txt ;

set numExceptions %1 ;
set numExceptions -- ;

loop 0 nchar
if ( !iscont[#1] ) set numcont #1 - 1 ;
endloop end
stop
```

```
loop 0 ('numcont' - 1)

      quote #1;

      set continueLoop 0 ;

      loop 0 'numExceptions'
            set exception 'exceptions[#2]' ;
            if(#1 ==  'exception')
                  set continueLoop 1 ;
                  continue;
            end
      stop

      if ('continueLoop' == 1) continue; end

      loop 0 ntax

            if ( contmaxs[#1 #2] == 65.535 )
                  if ( contmaxs[(#1+1) #2] == 65.535 )
                        continue ;
                  else
                        set min contmins[(#1+1) #2] ;
                        set max contmaxs[(#1+1) #2] ;
                  end
            else
                        set min contmins[#1 #2] ;
                        set max contmaxs[#1 #2] ;
                  if ( !(contmaxs[(#1+1) #2] == 65.535) )
                        if(contmins[(#1+1) #2] < 'min') set min
contmins[(#1+1) #2] ; end
                        if(contmaxs[(#1+1) #2] > 'max') set max
contmaxs[(#1+1) #2] ; end
                  end
            end

            quote MIN-MAX ;
            quote 'min'-'max' ;

            xread = #1 #2 'min'-'max' ;

      stop

      set secondchar #1 + 1 ;

      ccode ] 'secondchar'  ;

      if (#1 < ('numcont' - 1))
            set nextLoop #1 + 2 ;
            setloop 'nextLoop' ;
      end
```

```
stop

loop 0 'numExceptions'
     set exception 'exceptions[#1]' ;
     ccode [ 'exception' ;
stop

proc/;
```

collexceptions.run:

```
macro = ;

set exceptions[0] 10;
set exceptions[1] 29;
set exceptions[2] 50;
set exceptions[3] 67;
…
proc/;
```

Balance of characters (subsampleSPR.run)

This script searches for the optimal tree of the complete continuous partition and saves it to a TRE file. It then generates *n* pseudoreplicates of *m* continuous characters by activating them randomly (*n* and *m* have to be stated as agruments), searches for up to 10 optimal trees supported by each pseudoreplicate, calculates the weighted SPR distance of all of the trees found to the saved one, and retains the shortest value. This is done in order to avoid the use of consenses, which may result in the subestimation of SPR distances. Values are subsequently saved to a comma-delimited CSV file which can be opened directly with MS-Excel. The script can be easily modified to be used for entire matrices or different kinds of partitions. Tree search and SPR distance parameters can be also changed with ease.

```
macro = ;


var =
     + numrepl
     + numcharac
     + randomCharac
     + sprmove
     + sprparcial
```

```
     + sprsimilarity ;

set numrepl %1 ;
set numcharac %2 ;

macfloat 8 ;
taxname - ;
col 0 ;
rseed 0 ;

sprdiff [3 ;

loop 0 nchar
     if ( !iscont [#1] )
          ccode ] #1 ;
     else
          continue ;
     end
stop

mult = tbr ratchet replic 10 hold 10 ;
tsave *trees.tre ;
save ;
tsave/ ;

ccode ] . ;

report - ; silent = file ; log SPRresults.csv ; silent - file ;
quote Moves ; silent = file ;

loop 1 'numrepl'

     keep 0 ;

     loop 1 'numcharac'
          set randomCharac getrandom [ 0 nchar ] ;
          if ( (iscont['randomCharac']) &&
(!isact['randomCharac']) )
               ccode [ 'randomCharac' ;
          else
               setloop #2 - 1 ;
          end
     stop

     mult = tbr noratchet replic 20 hold 10 ;
     reroot 0 ;

     proc trees.tre ;
     set sprmove 10000 ;

     loop 0 (ntrees - 1)
          set sprparcial sprdiff [ #2 ntrees 13000x18 ] ;
```

```
                if ( 'sprparcial' < 'sprmove' )
                     set sprmove 'sprparcial' ;
                end
           stop

           silent - file ; quote 'sprmove' ; silent = file ;

           ccode ] . ;

      stop

      proc/;
```