Zoologica Scripta

Chasing ghosts: the phylogeny of Amaurobioidinae ghost spiders (Araneae, Anyphaenidae)

Facundo M. Labarque, Eduardo M. Soto, Martín J. Ramírez & Miquel A. Arnedo

Submitted: 23 November 2014
Accepted: 19 April 2015
doi:10.1111/zsc.12119

Labarque, F. M., Soto, E. M., Ramírez, M. J., Arnedo, M. A. (2015). Chasing ghosts: the phylogeny of Amaurobioidinae ghost spiders (Araneae, Anyphaenidae). —*Zoologica Scripta*, *00*, 000–000.

The family Anyphaenidae, also known as ghost spiders, includes a diverse array of nocturnal cursorial spiders that actively hunt on vegetation. The family is mostly distributed in the Americas and has been traditionally divided into three subfamilies. The mostly tropical and North American Anyphaeninae and the Amaurobioidinae, primarily distributed in southern South America, hold the bulk of the diversity, while the Malenellininae includes a single Chilean species. Here, we use a combined morphological and molecular approach to infer the relationships of the subfamily Amaurobioidinae and examine the delimitation of contentious genera. The morphological characters include both genitalic and somatic morphology, whereas molecular data include four markers, two mitochondrial (COI, 16S) and two nuclear (28S, H3). All our analyses agree on the monophyly of Amaurobioidinae, Amaurobioidini, Gayennini, the genera *Negayan*, *Amaurobioides*, *Josa*, *Araiya*, *Arachosia* and *Monapia*, as well as the paraphyly of Anyphaeninae. The total evidence analysis supports the novel placement of *Josa* as the sister group of both tribes Amaurobioidini and Gayennini, most of the previously known intergeneric relationships within Gayennini, and a clade of Amaurobioidini with a projecting ocular area, including *Aysenoides*, *Axyracrus*, *Amaurobioides* and *Aysenia*. The sequence data solve the puzzling placement of *Philisca puconensis*, here transferred to *Tomopisthes*, and *Tasata chiloensis*, transferred to *Oxysoma*. The advantages of the total evidence phylogenetic approach and the evolution of the male copulatory organ are discussed.

Corresponding author: *Facundo M. Labarque, Entomology Department, California Academy of Sciences, 55 Music Concourse Drive, San Francisco, CA 94118, USA; Division of Arachnology, Museo Argentino de Ciencias Naturales "Bernardino Rivadavia", Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Av. Ángel Gallardo 470, C1405DJR, Ciudad Autónoma de Buenos Aires, Buenos Aires, Argentina. E-mail: facundo.labarque@gmail.com*

Facundo M. Labarque, and Eduardo M. Soto, Departamento de Ecología, Genética y Evolución, IEGEBA (CONICET-UBA), Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Autónoma de Buenos Aires, Buenos Aires, Argentina; and Division of Arachnology, Museo Argentino de Ciencias Naturales "Bernardino Rivadavia", Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Av. Ángel Gallardo 470, C1405DJR Ciudad Autónoma de Buenos Aires, Buenos Aires, Argentina. E-mail: edusoto@gmail.com, edusoto@gmail.com

Martín J. Ramírez, Division of Arachnology, Museo Argentino de Ciencias Naturales "Bernardino Rivadavia", Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Av. Ángel Gallardo 470, C1405DJR Ciudad Autónoma de Buenos Aires, Buenos Aires, Argentina. E-mail: ramirez@macn.gov.ar

Miquel A. Arnedo, Departament de Biologia Animal & Institut de Recerca de la Biodiversitat (IRBio), Universitat de Barcelona, Av. Diagonal 645, E-8028 Barcelona, Spain. E-mail: marnedo@ub.edu

Introduction

The spider family Anyphaenidae, also called ghost spiders because their pale legs blur at high speeds when they flee, is an abundant component of the spider fauna in forests and grasslands across the Americas. It comprises more than 500 species of medium to small wandering spiders with two tarsal claws and claw tufts. Of the 56 known genera, 22 belong to the subfamily Amaurobioidinae, mostly distrib-

uted in southern South America, 33 to Anyphaeninae, mostly tropical and North American, and the single Chilean species *Malenella nana* represents the third subfamily, Malenellinae (Ramírez 1995, 2003; Brescovit 1997; World Spider Catalog 2015). The bulk of anyphaenid diversity is American, with only a few species in the Palearctic, Asia and Polynesia, and one genus, *Amaurobioides*, inhabits the coasts and islands of southern continents. Neotropical biodiversity inventories have found that Anyphaenidae is one of the five most abundant families among cursorial hunting spiders and within savannah habitats (Silva 1996; Rubio *et al.* 2008). These nocturnal spiders actively hunt on vegetation but are easily collected using a beating tray during the daytime when they take refuge in the curled leaves of shrubs or suspended clusters of dry leaves (Silva & Coddington 1996). The anyphaenid fauna of the temperate areas in southern South America has similar natural histories to the tropical Anyphaeninae, with the addition of several soil and grassland specialists (Ramírez 2003) and the remarkable *Amaurobioides*, specific to the spray area of marine rocky shores (Lamoral 1968; Forster 1970). As expected, recent canopy fogging samples from Chile revealed a high abundance of anyphaenids (M. J. Ramírez, personal observation; Laborda *et al.* 2013).

Anyphaenidae are easily recognized by the combined presence of spatulate claw tuft setae and a well-developed tracheal system, externally evident by the wide tracheal spiracle well separated from the spinnerets (Platnick 1974; Ramírez 1995). The family was classically placed among a loosely defined group of two-clawed spiders, the Dionycha, for which the third claw of legs, useful for spiders that dwell on silk webs, was replaced with a pad of setae that is specialized for adhesion to the smooth surface of leaves (Wolff *et al.* 2013). Only recently have the relationships of dionychian spiders been tested in the context of a wide taxonomic scope (Ramírez 2014), and the classical subfamilies Amaurobioidinae and Anyphaeninae were recovered as monophyletic and sister to each other. Malenellinae remained as sister to all other anyphaenids, but other taxa with complex tracheal systems were also placed as putative close relatives of *Malenella*. However, the study did not find conclusive evidence for the interfamiliar relationships in the vicinity of anyphaenids.

The family was taxonomically revised at the generic level by Brescovit (1997; Anyphaeninae) and Ramírez (1995, 2003; Amaurobioidinae, Malenellinae). Starting from the phylogenetic analysis of amaurobioidines using morphological data by Ramírez (2003), a series of subsequent contributions produced species-level revisions of genera and studies of the male genitalia, all expanding on the same data set (Ramírez *et al.* 2004; Ramírez 2007; Izquierdo & Ramírez 2008; Werenkraut & Ramírez 2009; González Márquez & Ramírez 2012; Soto & Ramírez 2012; Laborda *et al.* 2013). All these analyses corroborated the monophyly of most of the genera and solved the large-scale relationships within the subfamily, yet consistently provided low support for many intergeneric relationships, and indicated that a few genera might benefit by being redelimited (*Sanogasta*, *Philisca*, *Oxysoma* and *Tasata*). Another result was that the subfamily Anyphaeninae is most likely paraphyletic, but this was only based on a few out-group representatives.

With the aim to investigate the internal relationships of Amaurobioidinae using an additional data source, we generated a molecular data set for four markers: two nuclear genes, the ribosomal 28S rRNA and the protein-coding histone H3 and two mitochondrial genes, the ribosomal 16S rRNA and the protein-coding cytochrome *c* oxidase subunit I (COI). These sequences were analysed under Bayesian inference, maximum likelihood and parsimony, separately and with the accumulated morphological data matrix.

Materials and methods

Molecular data

The molecular data set included 67 species, 55 of which were sequenced in this study and 12 were provided by the ongoing Assembling the Tree of Life (AToL): Phylogeny of Spiders Project (W. Wheeler, personal communication). The Amaurobioidinae are represented by 60 species in 20 genera (Table S1). Out-group terminals included *M. nana* (Malenellinae) and six genera of Anyphaeninae. We succeeded in collecting both sequence and morphological data from 66 species. Unfortunately, we were unable to obtain suitable molecular data from the monotypic genera *Axyracrus* and *Selknamia* (Amaurobiodinae), and the species *Italaman santamaria* and *Wulfila albus* (Anyphaeninae), or suitable morphological data from *Sanogasta rufithorax* (Amaurobioidinae). For the total evidence analyses, we construed a chimera with the morphological information of *Amaurobioides maritima* and nucleotide sequences of a very similar, yet unidentified species of *Amaurobioides*.

Total DNA was extracted from one or two legs of freshly collected specimens that were preserved in 95–100% ethanol. Extraction, amplification and sequencing followed the protocols described by Arnedo *et al.* (2004). Partial fragments of the mitochondrial genes COI (~660 bp) and 16S rRNA (16S, ~450 bp) and the nuclear genes 28S rRNA (28S, ~800 bp) and Histone H3 (H3, ~330 pb) were amplified. Primers pairs and PCR annealing conditions for each locus are listed in Table S2.

Sequence errors and ambiguities were edited using the Staden Package ver. 1.4.0 (http://staden.sourceforge.net/). Each sequence was checked for contamination using an

NCBI BLAST search (http://ncbi.nlm.nih.gov/BLAST). Low-quality samples with ambiguous readings were removed from the matrix. Sequences were managed in BIOEDIT ver. 7.0.5.2 (Hall 1999). Stop codons were checked in DNASP ver. 5.10.1 (Librado & Rozas 2009). Taxonomic and sequence information of the study specimens are listed in Table S1.

Automatic multiple alignments were built using the FFT-NS-i slow search strategy in MAFFT ver. 6 (Katoh & Toh 2010). Several gap opening (GOP, 'gap opening penalty') and gap extension (GEP, 'offset value') costs were investigated to assess their influence on the results. The following parameter costs were employed (GOP/GEP): 1.00/0.5, 1.00/0.0, 1.53/0.5, 1.53/0.0 (default), 3.00/1.0, 3.00/0.5, 3.00/0.0. These values resulted in several alignments ranging from numerous short gaps ('gappy') to fewer and longer gaps in the sequences ('compacted'). The minimum value of the rescaled incongruence length difference (RILD) was used as a criterion to select the optimal alignment for each gene fragment (Wheeler & Hayashi 1998). The RILD measures the increase in homoplasy generated as a result of combining data sets and hence selecting the alignment combination that minimizes the RILD value and maximizes the congruence among partitions.

To avoid overweighting contiguous gap positions, gaps were recoded as presence/absence characters using Simmons and Ochoterena's simple code method (Simmons & Ochoterena 2000) as implemented in the program GAPCODER (Young & Healy 2002).

Morphological data

We fused the data sets of González Márquez & Ramírez (2012) and Soto & Ramírez (2012), which already contained the homology statements proposed in all previous studies. The morphological matrix includes 260 characters, 11 of which are multistate additive. Twenty-five per cent of the characters came from the male copulatory bulb, and the rest came from female sexual structures, spine patterns, coloration, general morphology, spinnerets, tracheal system and sexual behaviour. WINCLADA ver. 1.00.09 (Nixon 2002) was used to build the morphological data set and produce character mappings. We also added scorings for the additional molecular terminals for which we had sufficient specimens for dissections (all except *S. rufithorax*). The final morphological character matrix includes 118 taxa, 107 Amaurobioidinae species covering all known genera and 11 additional representatives of the other subfamilies (*M. nana* and ten species of Anyphaeninae) (Table S3).

Phylogenetic analyses

Bayesian analyses were performed using MRBAYES 3.2.1 (Ronquist & Huelsenbeck 2003), run remotely at the California Academy of Sciences Center of Comparative Genomics (Phylocluster, http://phylocluster.calacademy.org/) and CIPRES Science Gateway (Miller *et al.* 2010; https://www.phylo.org/). The program JMODELTEST ver. 2.1.1 (Guindon & Gascuel 2003; Darriba *et al.* 2012) was used to select the best-fitting model of evolution for each partition using the Akaike's information criterion (Akaike 1973; Buckley *et al.* 2002). Ribosomal genes (16S and 28S) were each treated as a single partition. Three different partition schemes were implemented for the protein-coding genes (COI and H3), namely *gene* (each gene a single partition), *3rd codon* (two partitions: 1st + 2nd codons and 3rd codon) and *codon specific* (one partition to each codon position), and evaluated using Bayes factors (BF) (Brown & Lemmon 2007). Acceptance or rejection of each strategy was based on the following cut-off: $BF \geq 10$ (strong evidence against the competing hypothesis); $10 < BF \geq -10$ (ambiguous, select least complex strategy); and $BF \leq -10$ (strong evidence for the competing hypothesis). The partitions corresponding to the gaps scored as absence/presence characters and to the morphological characters were treated with a standard discrete model. The substitution estimates were allowed to vary independently between each partition. For phylogenetic analysis, two independent runs with four simultaneous Markov chain Monte Carlo (MCMC) chains (one cold and three heated), each with random starting trees, were conducted simultaneously, sampling every 1000 generations until the standard deviation of the split frequencies of these two runs dropped below 0.01 (10^7 generations). The program TRACER ver. 1.5 (Rambaut & Drummund 2009) was used to ensure that the Markov chains had reached stationarity by examining the effective sample size values (above 200) and to determine the correct number of generations to discard as a burn-in (first 10% of generations).

Maximum-likelihood analysis was performed using the program RAXML ver. 7.3.0 (Stamatakis 2006; Stamatakis *et al.* 2008). We used the RAXMLGUI (Silvestro & Michalak 2011) graphical interface to define partitions and conduct the analyses. Partition schemes were implemented as above, recoded gaps were treated as binary characters, and morphological data were treated as multistate characters. Each gene partition was assigned an unlinked general time-reversible model with gamma distributed among site rate variation (GTR + Γ). The best likelihood tree was obtained out of 100 random iterations, and support was assessed by conducting 1000 nonparametric bootstrap replicates for each analysis. Both independent gene trees and concatenated analyses were run under the different partition schemes.

Parsimony analyses were performed using TNT version 1.1 (Goloboff *et al.* 2008a). The data were analysed under both equal and implied weights (Goloboff 1993) with a

concavity constant of the weighting function value $k = 6$, reproducing previous analysis conditions. Ramírez (2003), Lopardo (2005) and Goloboff *et al.* (2008b) have shown that mild concavity values (e.g. $k = 6$) yielded higher topological congruence indices for most morphological data sets. DNA sequences are usually more homoplasious than morphological data; hence, we further explored k values of 6, 20, 50 and 100 for the molecular and combined data sets, as suggested by Goloboff *et al.* (2008b). All tree searches were driven to hit independently 15 times the optimal scoring, using the default values of the 'New Technologies' search in TNT with sectorial searches, tree fusing and ratchet, followed by TBR branch swapping (string of commands *hold 20000; xmult = hits 15 ratchet 10; bb = fill only;*). Support values were estimated by jackknifing frequencies; each of the 1000 pseudoreplicates used three random addition sequences plus TBR, followed by TBR collapsing to calculate the consensus.

For the total evidence data set, morphological and molecular characters were combined in a single data matrix; taxa with no available molecular or morphological data were scored as missing entries. The combined morphological and molecular (plus gaps as absent/present state) data matrix was analysed with the three criteria as described above. Morphological characters were treated with a standard discrete model in Bayesian analyses (Lewis 2001), or as multistate characters (MULTI) in likelihood inference. For the comparison of tree resolution and support levels, we also analysed a reduced data set including only the 66 complete species with both molecular and morphological data, 'overlap' data set. This data set was analysed with the three criteria as described above.

Trees were edited with the program FIGTREE ver. 1.4 (Rambaut 2006–2014) and WINCLADA. All analyses were rooted with the branch separating *M. nana* (Malenellinae) from the remaining taxa. The data sets can be obtained in TREEBASE with the accession number 17447 (available at http://purl.org/phylo/treebase/phylows/study/TB2:S17447).

Homology of morphological structures

Analysis using dynamic homologies was performed to test the homology correspondences of male copulatory structures among Anyphaenidae: the primary and secondary conductors of the male palp (C1 and C2, respectively) and the prolateral projection of the paramedian apophysis (pPMA). Briefly stated, this works by trying alternative homology schemas, analysing a total evidence data set, and retaining the homology correspondences and trees that produce the most parsimonious solutions. The two alternative homology correspondences proposed in consecutive studies by Ramírez (2003, 2007) are coded in the morphology data set as characters 181–252 and 260–331, respectively. Each homology schema was tested activating the corresponding block of characters and inactivating the other.

Results

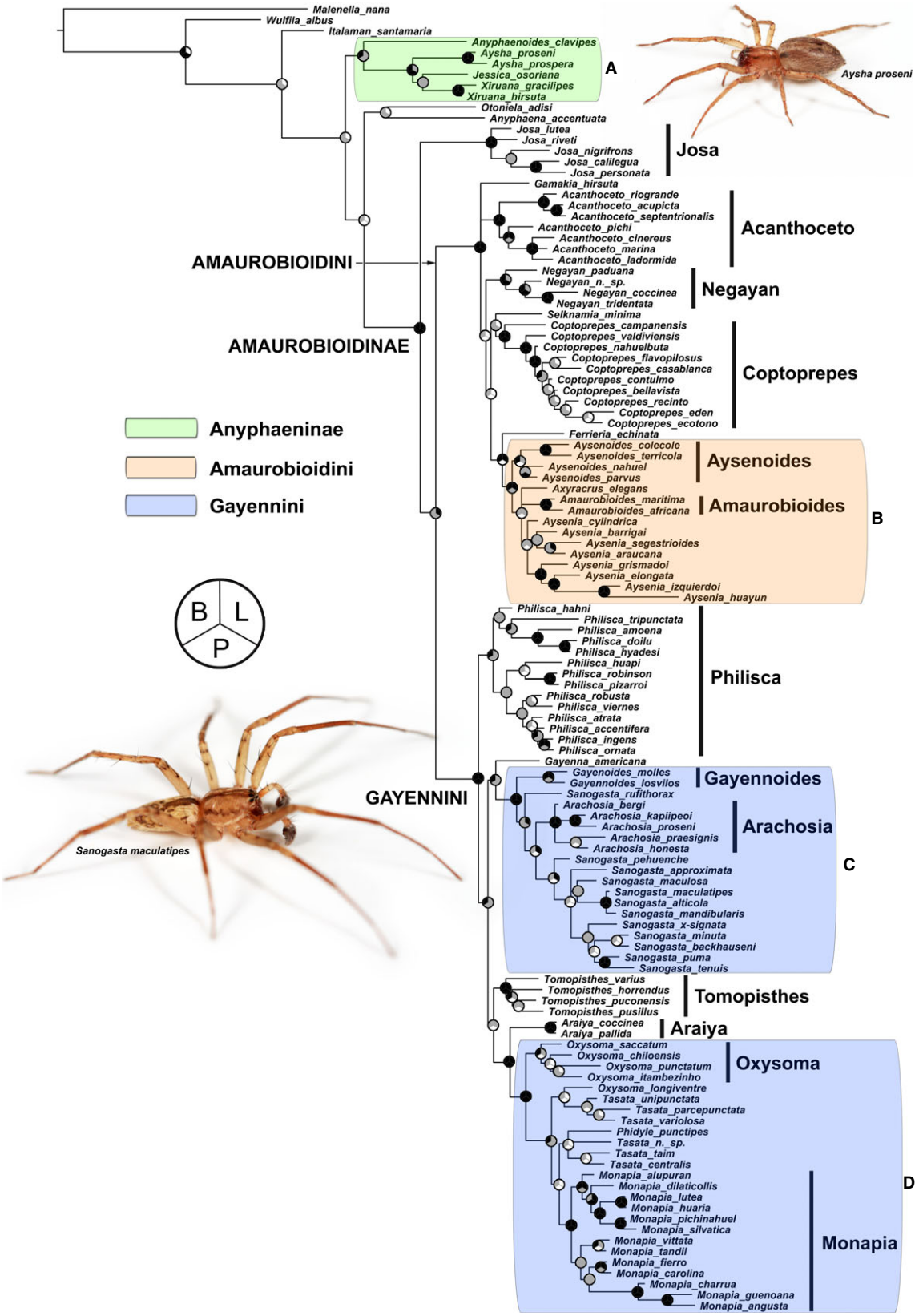
A summary of results, statistics and conditions for each phylogenetic analysis is presented in Table S4. We analysed the different sources of data (gene fragments and morphology) separately and in combination. We focus our discussion on the groups with higher support, hereby defined as those with a Bayesian posterior probability (PP) > 0.95, bootstrap proportions (BS) > 0.75 and jackknifing proportions (JS) > 0.75 (denoted with black in circles on Fig. 1, Figs S1–S3 and S6), and also on clades that were not well supported but were consistently recovered in Bayesian, likelihood and parsimony analyses (grey circles on Fig. 1, Figs S1 and S6 and black squares on Fig. 2).

Analysis of molecular data

The alignment of the protein-coding genes COI and H3 was trivial because they show no evidence of indel mutations. For the 16S ribosomal gene, the alignment with the parameter values GOP = 1.53 and GEP = 0, minimized RILD when combined with the unambiguously aligned protein-coding genes, whereas for the 28S ribosomal gene, the preferred parameter values were GOP = 3 and GEP = 0 (Table S5). The total length of the concatenated aligned data matrix was 2373 characters, of which 2269 corresponded to nucleotides (657 bp of COI, 327 bp of H3, 474 bp of 16S and 811 bp of 28S) and 104 to gaps coded as absence/presence (42 from 16S and 62 from 28S).

Summary results of the evolutionary models selected by the Akaike's information criterion in JMODELTEST are presented in Table S6, and for Bayes factor, partitioning strategies are provided in Table S7. Comparison of the

Fig. 1 Majority rule consensus of the trees sampled in the Bayesian analysis of the combined molecular alignments and morphology, with the preferred model for each gene fragment and gaps scored as absence/presence. The first 10% of generations were removed as burn-in. The information of the results of the analyses under maximum likelihood and parsimony ($k = 100$) is summarized in the node support. Circles at nodes indicate support levels subdivided by analysis (B, Bayesian; L, maximum likelihood; P, parsimony; see inset). Black indicates posterior probabilities >0.95, bootstrap proportions >0.75 and jackknifing proportions >0.75; grey indicates that the clade was recovered but with lower support than the previous values; white indicates that the clade was not recovered. Supported clades and genera discussed in the text are highlighted with shaded boxes and lateral bars, respectively.



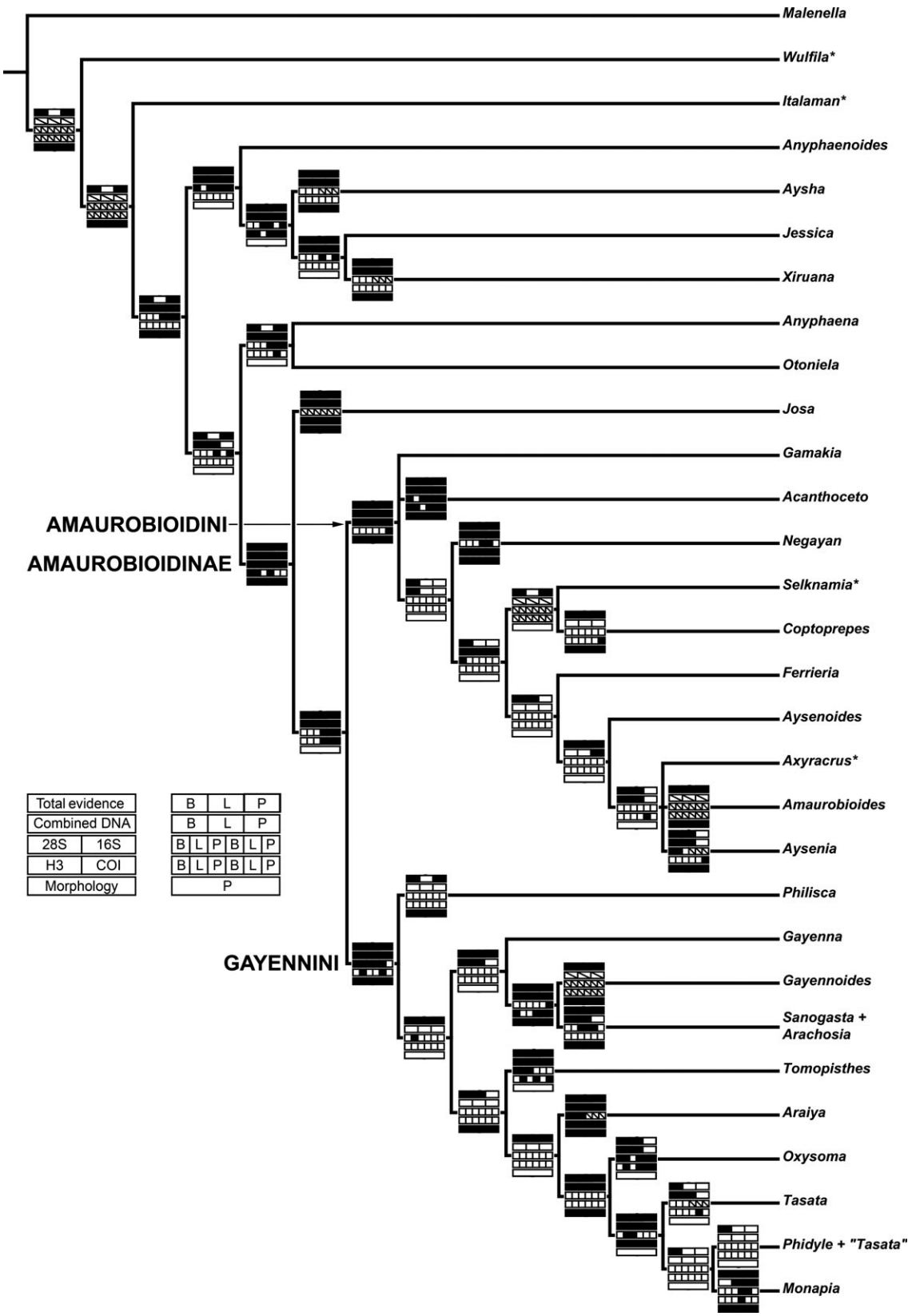


Fig. 2 Summary phylogeny of Anyphaenidae with congruence among data partitions and optimality criteria. Topology and clade composition is based on Bayesian analysis of total evidence (see Fig. 1). Boxes at nodes indicate each partition (morphology, combined DNA, individual genes or total evidence) subdivided by analysis method (B, Bayesian; L, maximum likelihood; P, parsimony; see inset). Black squares indicate that the clade was recovered (regardless of support) in the given analysis, white squares indicate that the clade was not recovered, and a diagonal line indicates that the clade was not tested (missing sequence data). Terminals with '*'are missing sequence data.

alternative partitioning strategies revealed an ambiguous preference for the codon-specific partition instead of the 3rd codon partition in COI, and a strong preference for the 3rd codon partition in H3 (Table S7). An ambiguous preference indicates that the least complex strategy should be selected, in this case the 3rd codon partition for COI (Table S4). Bayesian analysis of independent and concatenated genes achieved convergence within 5×10^7 generations (standard deviation of split frequencies <0.01). The maximum-likelihood analysis of the alternative partition schemes revealed better likelihoods for codon-specific partitions for both the COI and H3 genes (Table S4). Analyses of the concatenated data matrix with a/p gaps yielded one single tree with a score of –lnL = −25 433.291 (Fig. S1, Table S4). The maximum-likelihood results of the concatenated molecular matrix closely mirrored the results from the Bayesian analysis (Fig. S1). In the parsimony analysis, as expected, milder down-weighting against homoplasy (higher k values) resulted in topologies more similar to equally weighted trees. The k = 50 and 100 trees were identical to each other, and to one of the 97 equal-weighted trees, and the k = 20 tree only differed in the resolution of a few, weakly supported nodes. Generally, the parsimony results of the concatenated molecular matrix mostly differ from the Bayesian analysis in nodes with low support (Fig. S1).

Analysis of morphological data

The heuristic search strategies under equal weights produced 13 056 trees of length 1151. Heuristic searches in TNT using implied weights (k = 6) yielded one optimal tree of adjusted homoplasy 70.76107 (1183 steps long). Analyses under equal and implied weights (Figs S2 and S3, respectively) found the same supported groups, but differed in the less supported groups.

Analysis of the total evidence data

Results from Bayesian analysis using the preferred partition scheme is shown in Fig. 1 and Table S4. Maximum-likelihood analysis closely resembled the Bayesian results, but with lower node supports (Fig. 1). The parsimony analysis under equal weights resulted in 1944 trees of length 6629 (Table S4). The implied weights analysis with k = 100 obtained a single tree, also optimal under equal weights, but the topology of the k = 20 and 50 was slightly different in the resolution of a few groups with low bootstrap values. A summary of the synapomorphies of the main groups discussed below are presented in Table S8 and Fig. S4, as mapped on the total evidence Bayesian tree. Bayesian, maximum-likelihood and parsimony analyses of the 66 species with morphological and molecular data produced trees with higher average support values (i.e. more well-supported groups respective to the total number of groups), but otherwise, there were few differences with the total evidence tree (Fig. S6).

Robust clades and support

Several clades were consistently recovered in most analyses (Figs 1 and 2). The clades robust to the alternative inference methods were the subfamily Amaurobioidinae, the tribes Amaurobiodini and Gayennini, clade C (*Gayennoides*, *Sanogasta* and *Arachosia*), clade D comprising genera with densely spinated forelegs (*Oxysoma*, *Tasata*, *Phidyle* and *Monapia*) and the genera *Negayan*, *Acanthoceto* and *Araiya*. The genus *Josa* was also consistently recovered as the sister group of both tribes, although only protein-coding genes were sequenced. The monotypic genera *Axyracrus* and *Selknamia*, and the species *I. santamaria* and *W. albus* lacked molecular information and hence were not considered in the sensitivity assessment (Fig. 2; taxa without sequences are indicated with asterisks).

We compared the clade support for different data partitions by examining the average of posterior probabilities (PP), bootstrap (BS) and jackknifing (JS) values under Bayesian inference, maximum likelihood and parsimony with equal weights, respectively (Fig. S5A). All partitions yielded similar results, independently of the phylogenetic analyses used. The morphological data partition yielded lower average support (70%, 55% and 32%, respectively) than the molecular data partition (85%, 68% and 60%, respectively), whereas the combined data set yielded higher posterior probability values (90%) but intermediate bootstrap and jackknifing values (66% and 41%, respectively). The analysis of the data matrix including only species for which we had both morphological and molecular data ('overlap' data set, 66 terminals) showed the largest average values (93%, 79% and 68%, respectively). Because the average support value is sensitive to the total numbers of

clades, supported or not, we also counted the number of clades with highly supported groups (PP > 0.95, BS > 0.75, and JS > 0.75) as a measure of total support (Fig. S5B). The morphological, molecular and 'overlap' data sets all yielded a similar number of highly supported clades (29–40, 30–42 and 36–51 groups, respectively). By contrast, the total evidence analysis yielded more well-supported clades (79, 56 and 42 clades, respectively).

Homology of male copulatory structures

We re-analysed our total evidence data under alternative homology schemas for a variety of partitions and analytical conditions (Table S9), and in all cases, the preference for the new homology correspondences is maintained. In the preferred total evidence Bayesian tree, the three sclerites with contentious homology optimize as follows: The C1 is present in all Amaurobioidinae, with ambiguous optimization in the out-groups (Fig. S7A); the C2 is absent in the backbone phylogeny, with several convergent acquisitions (Fig. S7B); and the pPMA is present for all Amaurobioidinae, with several independent losses (Fig. S7C).

Discussion

This study increases the number of spider families that have been examined in a total evidence framework (Bruvo-Madaric *et al.* 2005; Bond & Hedin 2006; Agnarsson *et al.* 2007; Álvarez-Padilla *et al.* 2009; Arnedo *et al.* 2009; Lopardo *et al.* 2011; Bond *et al.* 2012; Wood *et al.* 2012). We have analysed our data using a variety of analytical approaches and data partitions to better understand phylogenetic patterns that are robust (e.g. less sensitive) to different analytical approaches and to identify the main sources of information for such clades.

Total evidence: resolution, support and missing data

Analytical comparison across several studies (Wortley & Scotland 2006) has revealed that the inclusion of morphological data has a significant positive effect upon resolution in combined analyses and our analysis is consistent with this. Here, we showed that the molecular data set (67 taxa) yielded lower support and resolution than the 'overlap' data set (66 taxa; i.e. only species scored for both molecules and morphology) across alternative phylogenetic analyses: fewer nodes were resolved and fewer clades found to be highly supported (Figs S1, S5B and S6). Despite the fact that our results mirror the increase in the number of supported clades found in the majority of studies compared by Wortley & Scotland (2006), they also shown that the inclusion of morphological data lacks a significant effect upon nodal support in combined analyses. Simulations (Wiens 2003, 2006) have shown that when the overall number of characters in a data set is high (i.e. 2000), the entire tree can be reconstructed correctly even when half of the taxa have 90% of their data cells lacking data. Adding new terminals (i.e. 52 taxa lacking 2373 DNA characters and one taxon lacking 260 morphological characters) to the 'overlap' data set (i.e. 66 taxa with 2633 combined characters) behaved in the same way; both matrices produced similar general topologies despite the methodology (i.e. Bayesian, parsimony) used in their reconstruction (Fig. 1 and Fig. S6). Also, the total evidence analysis (119 taxa) increased the resolution (more nodes) and the nodal support of the tree (Fig. 1 and Fig. S5B). These results differ from Wiens' simulations (2003, 2006), who observed that the inclusion of highly incomplete taxa seemed to have little effect on either the placement of taxa or on levels of support for this placement. However, the increase in resolution and nodal support on our total evidence results may be more related with taxon sampling than with missing data. While Wiens' simulations (2003, 2006) compared different set of character matrices with the same number of terminals, in our study, new taxa were added to a previous complete matrix. Several studies have indicated that introducing additional taxa into a phylogenetic analysis results (on average) in more accurate estimates (Hedtket *et al.* 2006). Thus, even at the expense of adding substantial missing data (i.e. lacking DNA sequences), our results suggest to include as many terminals as possible scored morphologically on a combined data set to increase the resolution of the tree without compromising the nodal support.

Total evidence: phylogenetic methods and congruence

Meta-analysis of 500 empirical DNA data studies (Rindal & Brower 2011) has revealed that, despite the phylogenetic algorithms (e.g. Bayesian, maximum likelihood, parsimony) used to infer relationships among taxa, the vast majority of studies produced topologies that exhibited minor differences and, when present, the incongruent nodes showed low support. Furthermore, they observed that when one analytical method yielded inconsistent results, the other method also behaved inconsistently. Our analysis again is congruent with these observations. Here, we show (Fig. 2) that inconsistent topologies are more common between data sets (namely genes, morphology, 'overlap' and total evidence) rather than between alternative phylogenetic inference methods applied to the same data set (e.g. Bayesian, parsimony). The incongruent nodes present on the preferred topologies of each method within a particular data set (Fig. 1 and Fig. S1) were consistent across data sets both in number (7–8 nodes) and position in the tree (i.e. *Aysenoides*, *Philisca* and *Oxysoma*).

Systematics

Our total evidence analyses reproduced previous results (Ramírez 2003, 2007) suggesting that anyphaenines are not

a monophyletic group. Although our sampling of anyphaenines is sparse, a clade of four genera (clade A) was consistently recovered in all analyses containing sequence data, by itself or in combination with morphology. A close relationship of *Aysha*, *Xiruana* and *Anyphaenoides* appears reasonable, as all have similar male copulatory organs, with an apical embolar division making a transversal loop, but the association with *Jessica* is unexpected and will require further scrutiny (see Brescovit 1997).

We recovered the monophyly of the subfamily Amaurobioidinae in the total evidence and the partial analyses of the molecular and the morphological partitions, under all analytical conditions. The results are identical for the tribes Gayennini and Amaurobioidini; Gayennini was even recovered with the individual genes, at least under some of the analytical conditions (Fig. 2). A novel placement was recovered for *Josa*, as sister to the remaining Amaurobioidinae in the total evidence and partial analyses of molecular partitions, differing from the morphological partition. It should be noted that only the protein-coding genes were scored for this genus.

The internal relationships of Amaurobioidini were poorly supported in general. The total evidence analyses supported the inclusion of *Amaurobioides*, *Aysenia*, *Aysenoides* and *Axyracrus*, all with a characteristically projecting ocular area, in clade B.

Within Gayennini, the genus *Philisca* was recovered as sister to the remaining members of the tribe, both in the total evidence and the molecular analyses. The monophyly of the genus, however, was challenged by the exclusion of a species tentatively placed in the genus (*Philisca puconensis*). Its new placement in *Tomopishtes* is consistently recovered and supported in the molecular and total evidence analyses (Fig. 1 and Fig. S1). This result was already recovered as a slightly suboptimal solution in the analysis of Ramírez (2003: 177). Because both genera are fully revised and the results are consistent across the analytical space, here, we propose the formal taxonomic transfer. The monophyly of the remaining *Philisca* was recovered only by the total evidence data set. The molecular analyses, however, supported the placement of *Philisca tripunctata*, atypical among its congeners by having a more elongated body and different genitalia to other Gayennini genera, as sister to clade D containing the spinose genera. Curiously, *P. tripunctata* is the least spiny species (legs I and II) in the entire subfamily, and it is similar in this regard to the other species of *Philisca* that were formerly placed in *Liparotoma* (see Soto & Ramírez 2012). Alternatively, the total evidence analyses favour the placement of *P. tripunctata* within *Philisca*, albeit with low support and a long branch (Fig. 1). A denser sampling of sequences of other *Philisca* species will hopefully contribute to clarify this issue.

A clade including members of the genera *Oxysoma*, *Tasata* and *Phidyle* (clade D) was recovered in all data sets and supported in the total evidence and molecular analyses. The molecular data shed new light on the limits of *Oxysoma* and *Tasata*, as the latter was very weakly supported in previous morphological analyses. The molecular data, by itself or in combination with morphology, suggest that *Tasata chiloensis* is, in fact, a member of *Oxysoma*, and *Oxysoma longiventre* may instead be a member of *Tasata*. The monophyly of *Tasata*, however, is compromised in the total evidence analyses, in which some *Tasata* species are placed closer to *Phidyle punctipes*, albeit with low support. We propose here to transfer *T. chiloensis* to *Oxysoma*, but additional taxonomic amendments in *Tasata* will have to wait until a denser sampling for this clade is achieved.

The total evidence analyses confirmed the close relationship of the genera *Gayennoides*, *Sanogasta* and *Arachosia* (clade C), and the monophyly of *Arachosia*. *Sanogasta*, however, was recovered as paraphyletic, with *Arachosia* nesting within it in most analyses. The actual taxonomic status of the species *S. rufithorax* must be re-examined in the context of denser sampling. Our analyses further support the close relationship of the Chilean genus *Gayenna* with these three genera, a relationship that was challenged in the recent analysis of morphological data of Soto & Ramírez (2012).

Homology of male copulatory structures

The male copulatory organ of anyphaenids is a complex structure providing as much as 25% of the morphological characters in the data set. Although highly informative, the complexity and variability of those structures pose a special challenge in establishing homologies. While the general morphology is conserved within each tribe, the correspondences between Gayennini, Amaurobioidini, *Josa* and the anyphaenines are more speculative. Ramírez (2007) tested those correspondences in an analysis using dynamic homologies. That analysis led to the conclusion that the homology of the primary and secondary conductors of the male palp (C1 and C2, respectively), and a prolateral projection of the paramedian apophysis (pPMA) should be corrected from the previous interpretation of Ramírez (2003). The results obtained here, after the addition of sequence data, have the potential to modify those conclusions, especially after the updated relationships of *Josa* and the anyphaenine out-groups. Here, we show that under alternative homology schemas for a variety of partitions and analytical conditions (Table S9), the preference for the new homology correspondences is maintained. In the preferred total evidence Bayesian tree, the three sclerites with contentious homology optimize as follows: The C1 is present in all Amaurobioidinae, with ambiguous optimization in the out-

groups (Fig. S7A); the C2 is absent in the backbone phylogeny, with several convergent acquisitions (Fig. S7B); the pPMA is present for all Amaurobioidinae, with several independent losses (Fig. S7C). As it was obtained in Ramírez (2007), the alternative homologies of the male genitalia, although affecting many characters at the same time, only produced slight topological changes in the resulting phylogenies and in groups of low support.

Taxonomy changes

Here, we propose two formal taxonomic transfers of species with ambiguous morphology, whose relationships were clarified by the sequence data.

Tomopisthes puconensis (Ramírez) new combination. Originally described as *P. puconensis* Ramírez (2003: 191). This species is very similar in habitus to *Tomopisthes pusillus*. Both have contrasting coloration, making them remarkably cryptic on epiphytic lichens, where they are frequently collected (Ramírez 2003).

Oxysoma chiloensis (Ramírez) new combination. Originally described as *T. chiloensis* Ramírez (2003: 237). This species is similar in habitus to *Oxysoma punctatum*, by its elongate pale body covered with small dark dots. They are very cryptic on the dry leaves of *Chusquea* spp. bamboos, where they are usually found (Ramírez 2003).

Acknowledgements

Thanks for hospitality to Nuria Macías-Hernández, Leticia Bidegaray-Batista, Carles Ribera and Gema Blasco-Melia (arachnological laboratory, UB) during FML's visit to Barcelona, and especially to Nuria Macías-Hernández and Leticia Bidegaray-Batista, who guided and trained FML through the molecular laboratory techniques. The AToL Spiders team (see below) kindly provided unpublished sequences used in this work; we especially appreciate the help of Ward Wheeler, Lorenzo Prendini, Torsten Dikow, Dana Price, and Ellen Trimarco (American Museum of Natural History). We would like to thanks Sarah Crews and Sara Ceccarelli for an early friendly revision of the manuscript and comments regarding English. We also like to thanks Lutz Bachmann and two anonymous reviewers for the comments and suggestions that improve the quality of this manuscript. Several friends and colleagues joined us in field trips: we are grateful to Cristian Grismado, Matías Izquierdo, Andrés Ojanguren (MACN) and Rosa María Fernández García (Harvard University), Iván Leiva, Silvia Moreno, Paola González de Rodt, Mascimiliano Recabarre Green, Guillermo Araya Arredondo, Mauricio Gana (CONAF, Juan Fernández Archipelago National Park) and Jaime Pizarro Araya (Universidad de La Serena). We further thank the authorities at the Administración de Parques Nacionales (APN, Argentina) and the Corporación Nacional Forestal of Chile (CONAF, Chile) for collecting permits and assistance in the field, and Mauricio Gana (Oficina Juan Fernández y Gestión, Ministerio del Interior y Seguridad Pública, Gobierno de Chile) for help in logistics in Juan Fernández. Museum specimens for the molecular work were kindly provided by the California Academy of Sciences (Charles Griswold), American Museum of Natural History (Norman Platnick) and Museo Argentino de Ciencias Naturales 'Bernardino Rivadavia' (Cristina Scioscia). Brian Simison helped with the access to the California Academy of Sciences Center of Comparative Genomics Phylocluster (Brian Simison). This work was supported by grants FONCyT PICT-2007 01393, 2011 1007, CONICET PIP 112–200801–03209 and UBA UBACyT 01/1240 to MJR, NSF EAR-0228699 to WW, and Fundación BBVA (3ª Convocatoria de Ayudas a la Investigación en Biología de la Conservación) to Joan Pons and MA, and by a doctoral grant from CONICET to FML.

References

Agnarsson, I., Maddison, W. P. & Avilés, L. (2007). The phylogeny of the social *Anelosimus* spiders (Araneae: Theridiidae) inferred from six molecular loci and morphology. *Molecular Phylogenetics and Evolution*, *43*, 833–851.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & F. Csáki (Eds) *2nd International Symposium on Information Theory*, *Tsahkadsor*, *Armenia*, *USSR* (pp. 267–281), September 2–8, 1971, Budapest: Akadémiai Kiadó.

Álvarez-Padilla, F., Dimitrov, D., Giribet, G. & Hormiga, G. (2009). Phylogenetic relationships of the spider family Tetragnathidae (Araneae, Araneoidea) based on morphological and DNA sequence data. *Cladistics*, *25*, 109–146.

Arnedo, M. A., Coddington, J., Agnarsson, I. & Gillespie, R. G. (2004). From a comb to a tree: phylogenetic relationships of the comb-footed spiders (Araneae, Theridiidae) inferred from nuclear and mitochondrial genes. *Molecular Phylogenetics and Evolution*, *31*, 225–245.

Arnedo, M. A., Hormiga, G. & Scharff, N. (2009). Higher-level phylogenetics of linyphiid spiders (Araneae, Linyphiidae) based on morphological and molecular evidence. *Cladistics*, *25*, 231–262.

Bond, J. E. & Hedin, M. (2006). A total evidence assessment of the phylogeny of North American euctenizine trapdoor spiders (Araneae, Mygalomorphae, Cyrtaucheniidae) using Bayesian inference. *Molecular Phylogenetics and Evolution*, *41*, 70–85.

Bond, J. E., Hendrixson, B. E., Hamilton, C. A. & Hedin, M. (2012). A reconsideration of the classification of the spider infraorder Mygalomorphae (Arachnida: Araneae) based on three nuclear genes and morphology. *PLoS One*, 7, e38753.

Brescovit, A. D. (1997). Revisão de Anyphaeninae Bertkau a nível de gêneros na região Neotropical (Araneae, Anyphaenidae). *Revista Brasileira de Zoologia*, *13*, 1–187.

Brown, J. M. & Lemmon, A. R. (2007). The importance of data partitioning and the utility of Bayes factors in Bayesian phylogenetics. *Systematic Biology*, *56*, 643–655.

Bruvo-Madaric, B., Huber, B. A., Steinacher, A. & Pass, G. (2005). Phylogeny of pholcid spiders (Araneae: Pholcidae): combined analysis using morphology and molecules. *Molecular Phylogenetics and Evolution*, *37*, 661–673.

Buckley, T. R., Arensburger, P., Simon, C. & Chambers, G. K. (2002). Combined data, Bayesian phylogenetics, and the origin of the New Zealand cicada genera. *Systematic Biology*, *51*, 4–18.

Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, *9*, 772.

Forster, R. R. (1970). The spiders of New Zealand. Part III. *Otago Museum Bulletin*, *3*, 1–184.

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

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

Goloboff, P. A., Carpenter, J. M., Arias, J. S. & Miranda Esquivel, D. R. (2008b). Weighting against homoplasy improves phylogenetic analysis of morphological data sets. *Cladistics*, *24*, 1–16.

González Márquez, M. E. & Ramírez, M. J. (2012). A revision and phylogenetic analysis of the spider genus *Aysenia* Tullgren (Araneae: Anyphaenidae, Amaurobioidinae). *Zootaxa*, *3201*, 1–26.

Guindon, S. & Gascuel, O. (2003). A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. *Systematic Biology*, *52*, 696–704.

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium*, *41*, 95–98.

Hedtket, S. M., Townsend, T. M. & Hills, D. M. (2006). Resolution of phylogenetic conflict in large data sets by increased taxon sampling. *Systematic Biology*, *55*, 522–529.

Izquierdo, M. A. & Ramírez, M. J. (2008). Two new spider species of the genera *Aysenia* and *Aysenoides* from southern Chile and Argentina: description and phylogenetic relationships (Araneae: Anyphaenidae, Amaurobioidinae). *Zootaxa*, *1861*, 29–43.

Katoh, K. & Toh, H. (2010). Parallelization of the MAFFT multiple sequence alignment program. *Bioinformatics*, *26*, 1899–1900.

Laborda, A., Ramírez, M. J. & Pizarro-Araya, J. (2013). New species of the spider genera *Aysenia* and *Aysenoides* from Chile and Argentina: description and phylogenetic relationships (Araneae: Anyphaenidae, Amaurobioidinae). *Zootaxa*, *3731*, 133–152.

Lamoral, B. H. (1968). On the ecology and habitat adaptations of two intertidal spiders, *Desis formidabilis* (O.P. Cambridge) and Amaurobioides africanus Hewitt, at "The Island" (Kommetjie, Cape Peninsula), with notes on the occurrence of two other spiders. *Annals of the Natal Museum*, *20*, 151–193.

Librado, P. & Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, *25*, 1451–1452.

Lewis, P. (2001). A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology*, *50*, 913–925.

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

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

Miller, M. A., Pfeiffer, W. & Schwartz, T. (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: *Proceedings of the Gateway Computing Environments Workshop (GCE)* (pp. 1–8), 14 November 2010, New Orleans, LA.

Nixon, K. C. (2002). *WinClada*. Ithaca, NY: Author. Available via http://www.cladistics.com

Platnick, N. I. (1974). The spider family Anyphaenidae in America, north of Mexico. *Bulletin of the Museum of Comparative Zoology*, *146*, 205–266.

Rambaut, A. (2006–2014). Tree Figure Drawing Tool Version 1.4. Institute of Evolutionary Biology, University of Edinburgh. Available via http://tree.bio.ed.ac.uk/

Rambaut, A. & Drummund, A. J. (2009). Tracer Version 1.5: MCMC trace analysis tool. Institute of Evolutionary Biology, University of Edinburgh. Available via http://tree.bio.ed.ac.uk/software/tracer

Ramírez, M. J. (1995). A phylogenetic analysis of the subfamilies of Anyphaenidae (Arachnida, Araneae). *Entomologica Scandinavica*, *26*, 361–384.

Ramírez, M. J. (2003). The spider subfamily Amaurobioidinae (Araneae, Anyphaenidae): a phylogenetic revision at the generic level. *Bulletin of the American Museum of Natural History*, *277*, 1–262.

Ramírez, M. J. (2007). Homology as a parsimony problem: a dynamic homology approach for morphological data. *Cladistics*, *23*, 588–612.

Ramírez, M. J. (2014). The morphology and phylogeny of dionychan spiders (Araneae, Araneomorphae). *Bulletin of the American Museum of Natural History*, *390*, 1–374.

Ramírez, M. J., Ansaldi, M. J. & Puglisi, A. F. (2004). Description of the females of *Oxysoma itambezinho* Ramírez and *Monapia tandil* Ramírez, and their effects on the generic relationships of Gayennini (Araneae, Anyphaenidae, Amaurobioidinae). *Zootaxa*, *668*, 1–8.

Rindal, E. & Brower, A. V. Z. (2011). Do model-based phylogenetic analyses perform better than parsimony? A test with empirical data. *Cladistics*, *27*, 331–334.

Ronquist, F. & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, *19*, 1572–1574.

Rubio, G. D., Corronca, J. A. & Damborsky, M. P. (2008). Do spider diversity and assemblages change in different contiguous habitats? A case study in the protected habitats of the humid Chaco Ecoregion, Northeast Argentina. *Environmental Entomology*, *37*, 419–430.

Silva, D. (1996). Species composition and community structure of peruvian rainforest spiders: a case study from a seasonally inundated forest along the Samiria river. *Revue Suisse de Zoologie*, *Hors série*, 597–610.

Silva, D. & Coddington, J. A. (1996). Spiders of Pakitza (Madre de Dios, Peru): species richness and notes on community structure. In D. E. Wilson & A. Sandoval (Eds) *Manu. The Biodiversity of Southeastern Peru* (pp. 253–311). Washington: Smithsonian Institution.

Silvestro, D. & Michalak, I. (2011). raxmlGUI: a graphical frontend for RAxML. *Organisms Diversity & Evolution*, *12*, 335–337.

Simmons, M. P. & Ochoterena, H. (2000). Gaps as characters in sequencebased phylogenetic analyses. *Systematic Biology*, *49*, 369–381.

Soto, E. M. & Ramírez, M. J. (2012). Revision and phylogenetic analysis of the spider genus *Philisca* Simon (Araneae: Anyphaenidae, Amaurobioidinae). *Zootaxa*, *3443*, 1–65.

Stamatakis, A. (2006). RAxML-VI-HPC: maximum Likelihood-based Phylogenetic Analyses with Thousands of Taxa and Mixed Models. *Bioinformatics*, *22*, 2688–2690.

Stamatakis, A., Hoover, P. & Rougemont, J. (2008). A fast bootstrapping algorithm for the RAxML web-servers. *Systematic Biology*, *57*, 758–771.

Werenkraut, V. & Ramírez, M. J. (2009). A revision and phylogenetic analysis of the spider genus *Coptoprepes* Simon (Araneae: Anyphaenidae, Amaurobioidinae). *Zootaxa*, *2212*, 1–40.

Wheeler, W. C. & Hayashi, C. Y. (1998). The phylogeny of the extant chelicerate orders. *Cladistics*, *14*, 173–192.

Wiens, J. (2003). Missing data, incomplete taxa, and phylogenetic accuracy. *Systematic Biology*, *52*, 528.

Wiens, J. (2006). Missing data and the design of phylogenetic analyses. *Journal of Biomedical Informatics*, *39*, 34–42.

Wolff, J. O., Nentwig, W. & Gorb, S. N. (2013). The great silk alternative: multiple co-evolution of web loss and sticky hairs in spiders. *PLoS One*, *8*, e62682, 1–13.

Wood, H. M., Griswold, C. E. & Gillespie, R. G. (2012). Phylogenetic placement of pelican spiders (Archaeidae, Araneae), with insight into evolution of the "neck" and predatory behaviours of the superfamily Palpimanoidea. *Cladistics*, *28*, 598–626.

World Spider Catalog (2015). World Spider Catalog. Natural History Museum Bern, Available via http://wsc.nmbe.ch, version 16 (accessed on 22.1.2015)

Wortley, A. H. & Scotland, R. W. (2006). The effect of combining molecular and morphological data in published phylogenetic analyses. *Systematic Biology*, *55*, 677–685.

Young, N. D. & Healy, J. (2002). GapCoder. Available via http://www.trinity.edu/nyoung/GapCoder/Download.html

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Fig. S1. Majority rule consensus of the trees sampled in the Bayesian analysis of the combined molecular alignments and morphology, with the preferred model for each gene fragment and gaps scored as absence/presence.

Fig. S2. Strict consensus of the 13056 most parsimonious trees (MPTs) obtained from the morphological parsimony analyses under equal weights.

Fig. S3. Single tree obtained from the parsimony analyses under implied weights ($k = 6$) of the morphology.

Fig. S4. Majority rule consensus of the trees sampled by the Bayesian analysis of the combined molecular alignments and morphology, with the preferred model for each gene fragment and gaps scored as absence/presence (Fig. 1), with morphological character changes mapped using Unambiguous Changes (length 1218 steps, CI = 0.20 and RI = 0.67).

Fig. S5. Variation of support levels of the different data partitions, for Bayesian, maximum likelihood and parsimony analyses under equal weights, respectively.

Fig. S6. Majority rule consensus of the trees sampled in the Bayesian analysis of 66 terminals ('overlap') of the combined molecular alignments and morphology, with the preferred model for each gene fragment and gaps scored as absence/presence.

Fig. S7. Mapping of male palp organ characters on the preferred total evidence Bayesian tree (Fig. 1).

Table S1. List of the specimens used for the molecular phylogenetic analysis in the study, localities and vouchers codes.

Table S2. Primer sequences, source and annealing temperatures.

Table S3. List of the specimens used for the morphological phylogenetic analysis in the study.

Table S4. Summary tree statistics and conditions for each analysis.

Table S5. Summary of the results of the parsimony analysis of the alignments obtained in with MAFFT under different parameter combination and similarity values with the results of the combined data matrix (ribosomal gene + protein coding genes).

Table S6. Evolutionary models used for Bayesian analysis as selected by Akaike information criterion (AIC) in JMODELTETS ver. 2.1.1.

Table S7. Results of Bayes factor hypothesis testing for partitioning strategies.

Table S8. Unambiguos synapomorphies of the genera and internal nodes referred in Fig. S4.

Table S9. Alternative homology schemas for the male palpal organ.