

RESEARCH ARTICLE

Open Access



Origin and higher-level diversification of acariform mites – evidence from nuclear ribosomal genes, extensive taxon sampling, and secondary structure alignment

A R Pepato^{1*} and P B Klimov^{2,3*}

Abstract

Background: Acariformes is the most species-rich and morphologically diverse radiation of chelicerate arthropods, known from the oldest terrestrial ecosystems. It is also a key lineage in understanding the evolution of this group, with the most vexing question whether mites, or Acari (Parasitiformes and Acariformes) is monophyletic. Previous molecular studies recovered Acari either as monophyletic or non-monophyletic, albeit with a limited taxon sampling. Similarly, relationships between basal acariform groups (include little-known, deep-soil 'endeostigmatan' mites) and major lineages of Acariformes (Sarcoptiformes, Prostigmata) are virtually unknown. We infer phylogeny of chelicerate arthropods, using a large and representative dataset, comprising all main in- and outgroups (228 taxa). Basal diversity of Acariformes is particularly well sampled. With this dataset, we conduct a series of phylogenetically explicit tests of chelicerate and acariform relationships and present a phylogenetic framework for internal relationships of acariform mites.

Results: Our molecular data strongly support a diphyletic Acari, with Acariformes as the sister group to Solifugae ($PP = 1.0$; $BP = 100$), the so called Poecilophysidea. Among Acariformes, some representatives of the basal group Endeostigmata (mainly deep-soil mites) were recovered as sister-groups to the remaining Acariformes (i. e., Trombidiformes + and most of Sarcoptiformes). Desmonomatan oribatid mites (soil and litter mites) were recovered as the monophyletic sister group of Astigmata (e. g., stored product mites, house dust mites, mange mites, feather and fur mites). Trombidiformes (Sphaerolichida + Prostigmata) is strongly supported ($PP = 1.0$; $BP = 98-100$). Labidostomatina was inferred as the basal lineage of Prostigmata. Eleutherengona (e. g., spider mites) and Parasitengona (e. g., chiggers, fresh water mites) were recovered as monophyletic. By contrast, Eupodina (e. g., snout mites and relatives) was not. Marine mites (Halacaridae) were traditionally regarded as the sister-group to Bdelloidea (Eupodina), but our analyses show their close relationships to Parasitengona.

(Continued on next page)

* Correspondence: aepato@gmail.com; pklimov@umich.edu

¹Departamento de Zoologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Av. Antonio Carlos, 6627, 31270-901 Belo Horizonte, Brazil

²Department of Ecology and Evolutionary Biology, University of Michigan, 1109 Geddes Ave, Ann Arbor, MI 48109-1079, USA

Full list of author information is available at the end of the article



(Continued from previous page)

Conclusions: Non-trivial relationships recovered by our analyses with high support (*i.e.*, basal arrangement of endeostigmatid lineages, the position of marine mites, polyphyly of Eupodina) had been proposed by previous underappreciated morphological studies. Thus, we update currently the accepted taxonomic classification to reflect these results: the superfamily Halacaroidea Murray, 1877 is moved from the infraorder Eupodina Krantz, 1978 to Anystina van der Hammen, 1972; and the subfamily Erythracarinae Oudemans, 1936 (formerly in Anystidae Oudemans, 1902) is elevated to family rank, Erythracaridae *stat. ressur.*, leaving Anystidae only with the nominal subfamily. Our study also shows that a clade comprising early derivative Endeostigmata (Alycidae, Nanorchestidae, Nematalycidae, and maybe Alicorhagiidae) should be treated as a taxon with the same rank as Sarcoptiformes and Trombidiformes, and the scope of the superfamily Bdelloidea should be changed. Before turning those findings into nomenclatural changes, however, we consider that our study calls for (i) finding shared apomorphies of the early derivative Endeostigmata clade and the clade including the remaining Acariformes; (ii) a well-supported hypothesis for Alicorhagiidae placement; (iii) sampling the families Proterorhagiidae, Proteonematalycidae and Grandjeanicidae not yet included in molecular analyses; (iv) undertake a denser sampling of clades traditionally placed in Eupodina, Anystina (Trombidiformes) and Palaeosomata (Sarcoptiformes), since consensus networks and *Internode certainty* (IC) and *IC All* (ICA) indices indicate high levels of conflict in these tree regions. Our study shows that regions of ambiguous alignment may provide useful phylogenetic signal when secondary structure information is used to guide the alignment procedure and provides an R implementation to the Bayesian Relative Rates test.

Background

Mites, a ubiquitous and megadiverse group of chelicerates, are traditionally classified into two large assemblages: Acariformes (= Actinotrichida) and Parasitiformes (= Anactinotrichida), as reviewed by [1] and references therein. Recent analyses employing partial ribosomal and mitochondrial markers [2, 3], combined partial ribosomal genes and morphology [4], EST data sets [5] and genomic scale datasets [6, 7] suggest that Acari is a diphyletic group, something that contrasts with earlier studies employing partial ribosomal genes alone [8] or combined with morphology [9, 10]. Specifically, studies that recovered mites as a diphyletic group (except for [7]) proposed an unranked lineage, Acariformes + Solifugae, for which Pepato et al. [4] adopted the name Poecilophysidea.

Ingroup relationships of Acariformes were proposed based on morphology [11, 12] as follows: many basal Endeostigmata and all Oribatida (including Astigmata) form the Sarcoptiformes Reuter, 1909 (morphologically defined mainly by the presence of rutella); in contrast, another large radiation, the Trombidiformes Reuter, 1909, is represented by Sphaerolichida OConnor, 1984 and the megadiverse Prostigmata Kramer, 1877 (defined mainly by the lack of rutella and reduction of dorsal setation). Prostigmata were hypothesized to include three main groups: Labidostommatina Krantz, 1978 plus Eupodina Krantz, 1978, Anystina van der Hammen, 1972 and Eleutherengona Oudemans, 1909 [11, 13, 14], although Eupodina was placed either as sister group of Anystina [11] or as the basal lineage of the remaining Prostigmata [13]. Unfortunately, these hypotheses were not fully tested with molecular data.

Here, we used nearly complete sequences of the small and large subunits of nuclear rDNA (SSU and LSU) of all major lineages of acariform mites and numerous chelicerate and non-chelicerate outgroups to infer chelicerate relationships, test the hypotheses on the placement of Acariformes, and their internal relationships. Recently published studies on the Acariformes internal phylogeny put relatively little emphasis on outgroup sampling, taking a monophyletic Acari as granted and including just a few Parasitiformes terminals [15, 16], whereas those presenting broader taxon sampling [2, 4] missed crucial acariform families such as Lordalychidae and deep soil Endeostigmata (never sequenced previously, except for *Alicorhagia*) and were based on SSU and small stretches of the LSU gene (~300 bp); the only published study on Chelicerata using large portions of the LSU gene [3] has a biased in-group Acariformes sampling and did not attempt to employ the secondary structure in order to align sequences and improve phylogenetic analyses [17, 18]. Our study accounts for these shortcomings, offering an opportunity to investigate questions related to the origin of Acariformes with a greater accuracy.

Results

Data set characteristics

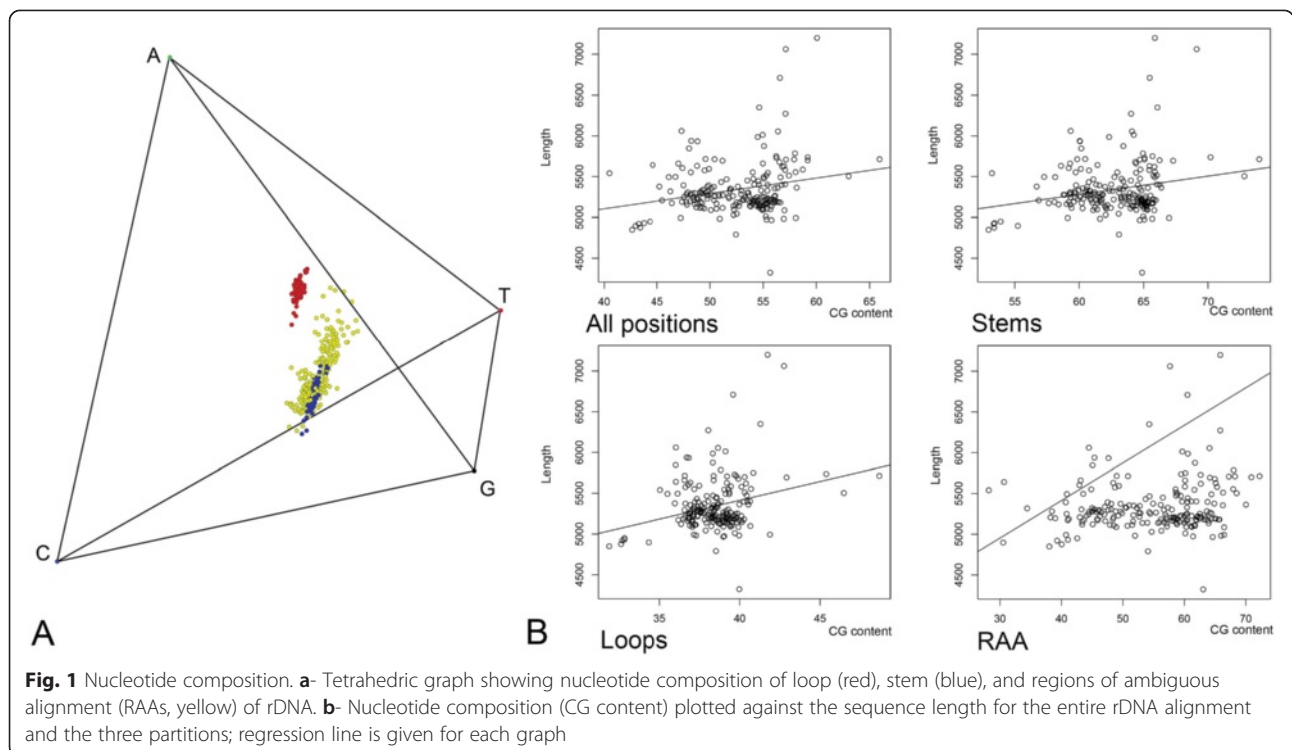
The length of SSU rDNA sequences (helices H47 to H1399) of the 228 species used in this study ranged from 1691 bp in *Macrocheles* (Parasitiformes, Mesostigmata) to 2429 bp in *Peripatus* (Onychophora). Among the Chelicerata, the longest sequence (2135 bp) was observed in *Haplochthonius* (Acariformes, Oribatida). For LSU rDNA (helices H224 to H2675), sequences ranged from 3204 bp in *Amblyseius* (Parasitiformes, Mesostigmata) to

5367 bp in *Penaeus* (Decapoda). Among the Chelicerata, the longest sequence was found in *Brevipalpus* (Acariformes, Trombidiformes) with at least 3989 bp (a short region at the 3' end was not sequenced for this species). The two genes were aligned using a secondary structure consensus approach (see below). The final aligned matrix (after exclusion of regions of ambiguous alignment, RAAs) contained 4503 nt positions; of them, 1240 were constant, and 2421 were parsimony informative. There were 2251 paired (the number is uneven because a small stretch of pairing region was not sequenced) and 2252 non-paired nucleotide positions in the alignment.

We detected a strong bias toward adenines in the loop regions (colored by red on Fig. 1a) relative to stems (blue); values for regions of ambiguous alignment (yellow) overlapped with those for stems but were more spread. The sequence length and GC content were positively correlated ($\rho = 0.23, P < 0.01$, Fig. 1b), similarly to that reported previously [17]. A weak positive correlation ($\rho = 0.15, P = 0.0337$) was detected for RAA; these regions account for most of the observed length variation. The observed differences in nucleotide frequency among secondary structure defined partitions suggest separate models of molecular evolution; irrespective of modeling covarying sites in stem regions (see the Methods and Discussion sections below). From the 228 taxa, 116 deviated from the maximum likelihood stationary composition. Despite this, arguably spurious groupings due to

base composition were not detected in the phylogenetic analyses (see the following section).

To evaluate the saturation level in our dataset, we calculated the index of substitution saturation (Iss), a critical value of the parameter at which a phylogenetic analysis will begin to fail to recover the true tree for either symmetrical (Iss.c sym) or asymmetrical (Iss.c asym) topology. If Iss of a particular dataset is substantially smaller than both Iss.c sym and Iss.c asym, no significant saturation is presumed [19, 20]. Our analysis suggests that none of the secondary structure defined partitions are saturated – Iss, Iss.C Sym, and Iss C Asym were as follows (ambiguous positions excluded): Loops 0.159, 0.700, and 0.395; Stems 0.218, 0.686, and 0.366; and RAA after filtered using *Aliscore* 0.640, 11.936, 19.025. Likelihood mapping (LM) and consensus networks (CN) analyses conducted on bootstrap trees displayed contrasting results with regard to the phylogenetic signal in individual partitions and combined alignment. The LM analysis (Fig. 2b) indicated that the stem and loop partitions are similarly informative, leaving small fractions of quartets in the star-like tree region (1.5 and 1.4 %, respectively), and RAA regions with a larger proportion of unsolved quartets irrespective of being analyzed before or after *Aliscore* filtering (11.9 and 13.1 %, respectively). The CN analysis (Fig. 3) showed fewer conflicting bipartitions and a resolution near to that observed for the RAA concatenated tree. It is interesting that a major acariform lineage, Astigmata, was inferred basal



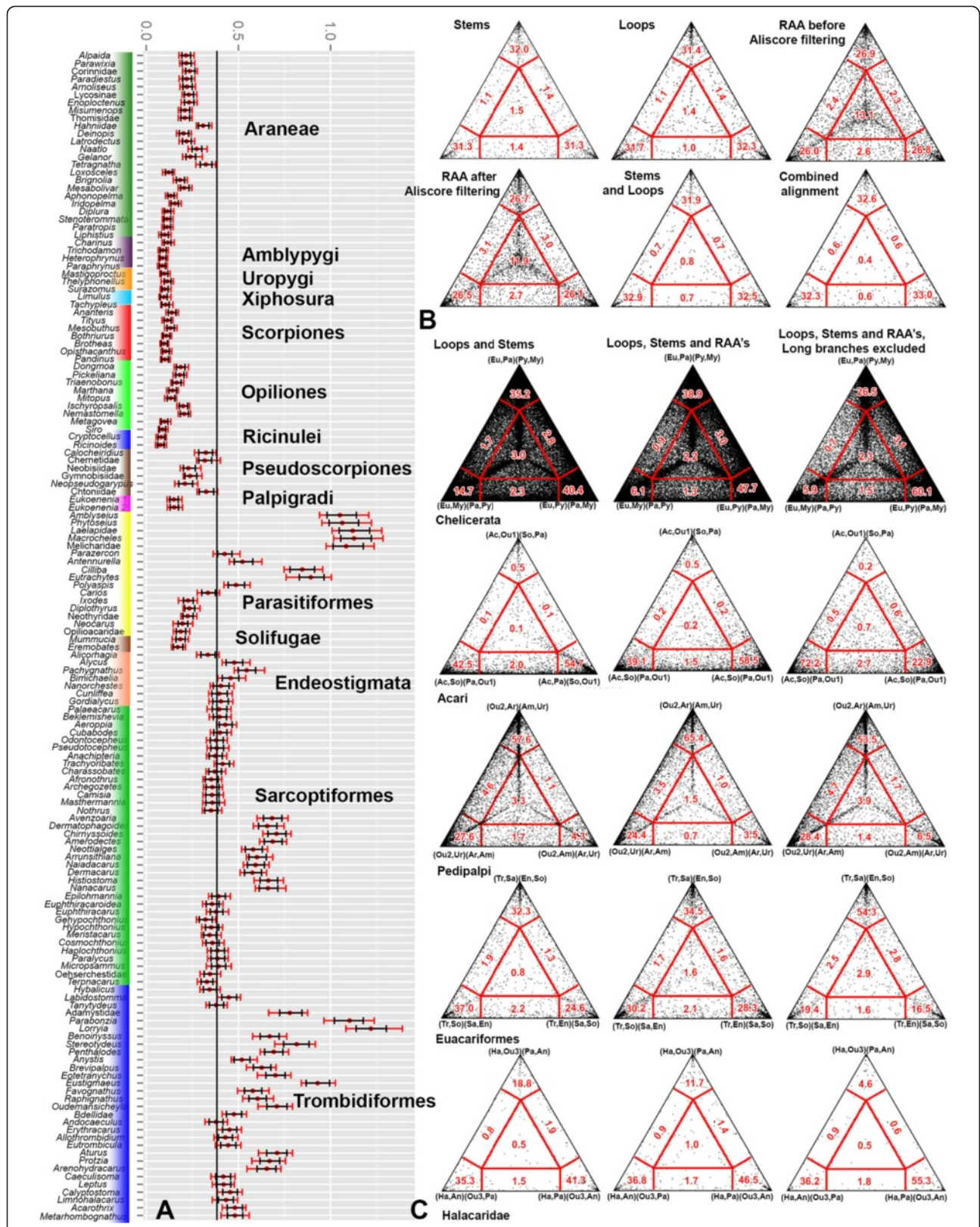


Fig. 2 (See legend on next page.)

(See figure on previous page.)

Fig. 2 Bayesian Relative Rates test and Likelihood Mapping. **a-** Bayesian Relative Rates of substitutions. For each horizontal bar, the mean rate (red dot), the credible interval (black) and the range (red) are given. **b-** Likelihood mapping for stems, loops, RAA partitions before and after *Aliscore* filtering, combined stems and loops, and combined stems, loops and filtered RAAs. **c-** Likelihood mapping for selected bipartitions. The mapping was performed for concatenated stems and loops, all partitions, and all partitions with rapidly evolving lineages (long branches) excluded. Abbreviations. Chelicerata: Eu, Euchelicerata; My, Myriapoda; Py, Pycnogonida; Pa, Pancrustacea. Acari: Ac, Acariformes: So, Solifugae; Pa, Parasitiformes; Ou1, outgroup Euchelicerata. Pedipalpi: Ou2, Outgroup Euchelicerata; Ur, Uropygi; Ar, Araneae; Am, Amblypygi. Euacariformes: Tr, Trombidiformes; So, Solifugae; Sa, Sarcoptiformes; En, Endeostigmata. Ha, Halacaridae; An, Caeculidae + Erythracarinae; Pa, Parasitengona; Ou3, all other Acariformes

at the Acariformes tree when the loop partition alone was analyzed. This position was recovered previously using unpartitioned analysis of the 18S gene [21], but is considered spurious based on morphology (*i.e.*, the presence of the opisthonotal glands suggests its origin within derived Oribatida). Similar warnings concerning the loop regions were issued previously, although they were linked to high levels of homoplasy detected by saturation tests [22], something that was not observed here (see above). In fact, tree certainty for the loop partition (Tree certainty including all conflicting bipartitions (TCA) for this tree: 90.4, Relative TCA for the best loop ML tree: 0.40) is considerably lower than that for stem partition (respectively 107.0 and 0.48), hence our analysis highlights potential non-phylogenetic signal in the loop regions, which may solve quartets in LM analyses but may introduce conflicts in the topology. These results indicate the need for careful evaluation of trees based on the total evidence approach because they can be influenced by a single rDNA partition.

Molecular phylogenetic analyses

Bayesian and Maximum Likelihood trees inferred from the two-partition (stems, loops) and three-partition (stems, loops, RAAs) datasets were largely congruent (Figs. 4 and 5, Additional file 4). In order to test for potential Long Branch Attraction Artifact (LBA), terminals for which the Bayesian Rates test showed no overlap of their credible intervals to those of the remaining Acariformes (Fig. 2a) were removed. Then this dataset was partitioned (stems, loops, RAA) and re-analyzed. This reduced dataset was also subjected to a likelihood analysis mapping for some alternative resolutions of relevant bipartitions (Fig. 2c). No grouping due to putative LBA could be detected in this way.

Arthropod relationships

Most of the well-supported outgroup relationships were similar to previous molecular studies, indicating the presence of useful, phylogenetically informative signal in our target genes. Pancrustacea was recovered with a very strong support (100 % of BP and PP in all analyses), whereas Myriochelata (Myriapoda plus Chelicerata) was

recovered with a low support in both analyses (two partitions, Bayesian and ML: $PP = 78$, $BP = 40$; three partitions, Bayesian and ML: $PP = 82$, $BP = 46$). Chelicerata in its traditional sense (Pycnogonida + Euchelicerata) was recovered only after the inclusion of the RAA data ($PP = 100$, $BP = 69$). We also observed an increasing proportion of quartets ((Pancrustacea, Myriapoda)(-Chelicerata, Pycnogonida)) from the pre-aligned, three-partition, and long branch pruned dataset (Fig. 2c).

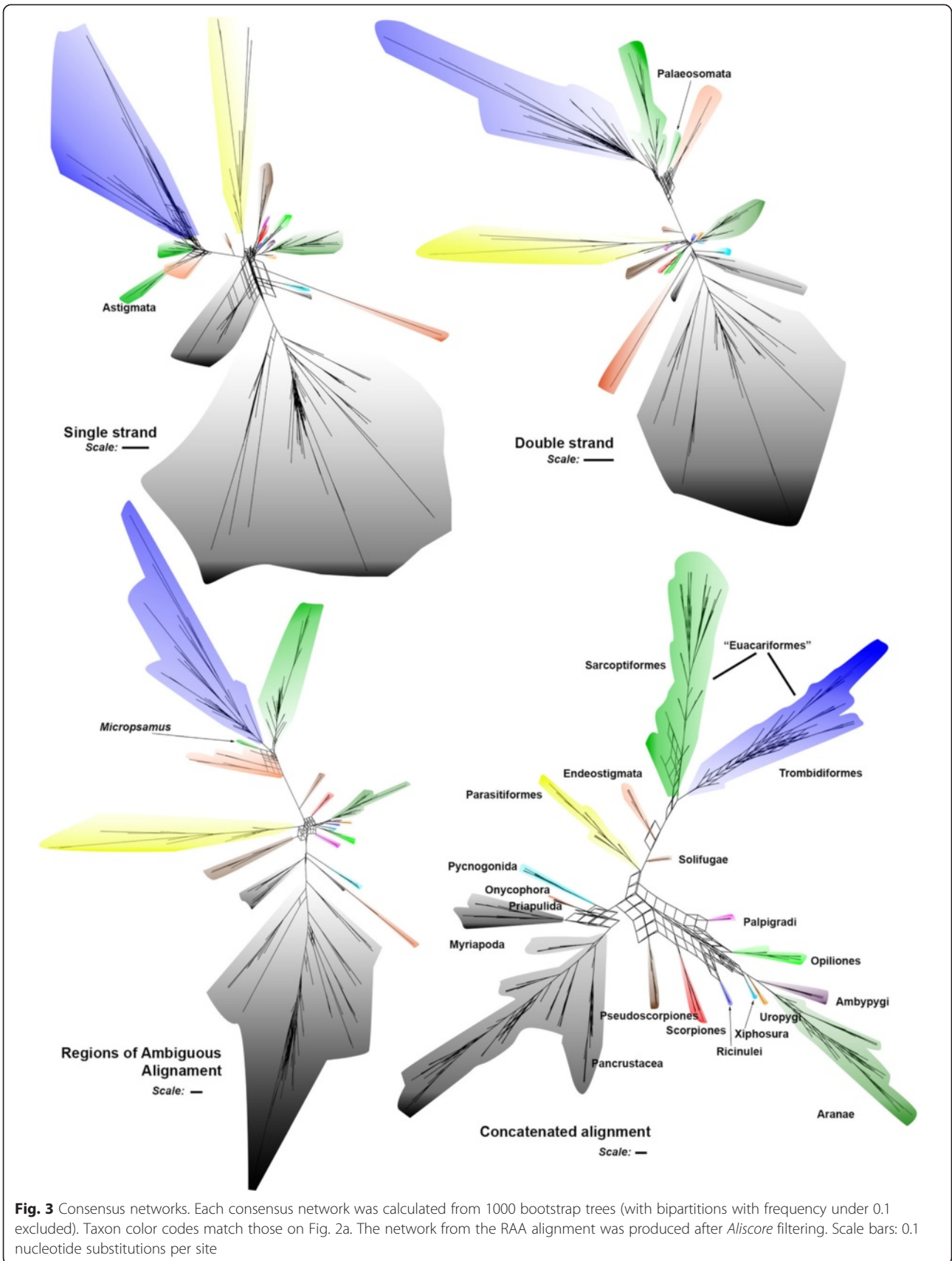
Two alternative hypotheses to Myriochelata, Cormogonida (= Euchelicerata (Myriapoda("Crustacea", Hexapoda))) [23, 24] and Mandibulata (= Myriapoda ("Crustacea", Hexapoda)) [25–27], could not be rejected statistically by the AU test given our data (Table 1). Euchelicerata was recovered under all analytical approaches and datasets, with a high support.

Euchelicerata relationships and monophyly of Acari

All chelicerate orders were well supported. In contrast, the class Arachnida was not recovered, although the AU test could not rule out this grouping. This situation is due to the unstable position of Xiphosura and most other Euchelicerata. Only two supra-ordinal groupings, Tetrapulmonata and Poecilophysidea (*i.e.* Acariformes + Solifugae) are stable and received substantial support.

Tetrapulmonata is a well-defined and uncontroversial clade, whereby the topology recovered here was (Uropygi (Amblypygi + Araneae)), the so-called Labellata [28] rather than the Pedipalpi clade (Amblypygi + Uropygi) typically found in morphological and combined morphological + DNA analyses [4, 9, 10, 29]. Pedipalpi was rejected by the data after including the RAA partition (Table 1).

The best supported topology did not recover mites as a monophylum. Parasitiformes is unstable, since it was found associated with Palpigradi in the ML analyses including only stems and loops ($BP = 32$) or Poecilophysidea ($PP = 1.0$, $BP = 71$) after including RAAs. Regardless, the AU test cannot reject most of alternative placements, such as a sister-group relationship with Pseudoscorpiones [2], Opiliones [6], Ricinulei (the so called Cryptognomae [30, 31]) (Table 1). Our data, however, reject Acari [32] and Acaromorpha (Acari + Ricinulei) [9, 10, 28, 29, 32], as well as alternative placements of Solifugae, such as the Haplocnemata (= Solifugae + Pseudoscorpiones [9, 10, 28, 29])



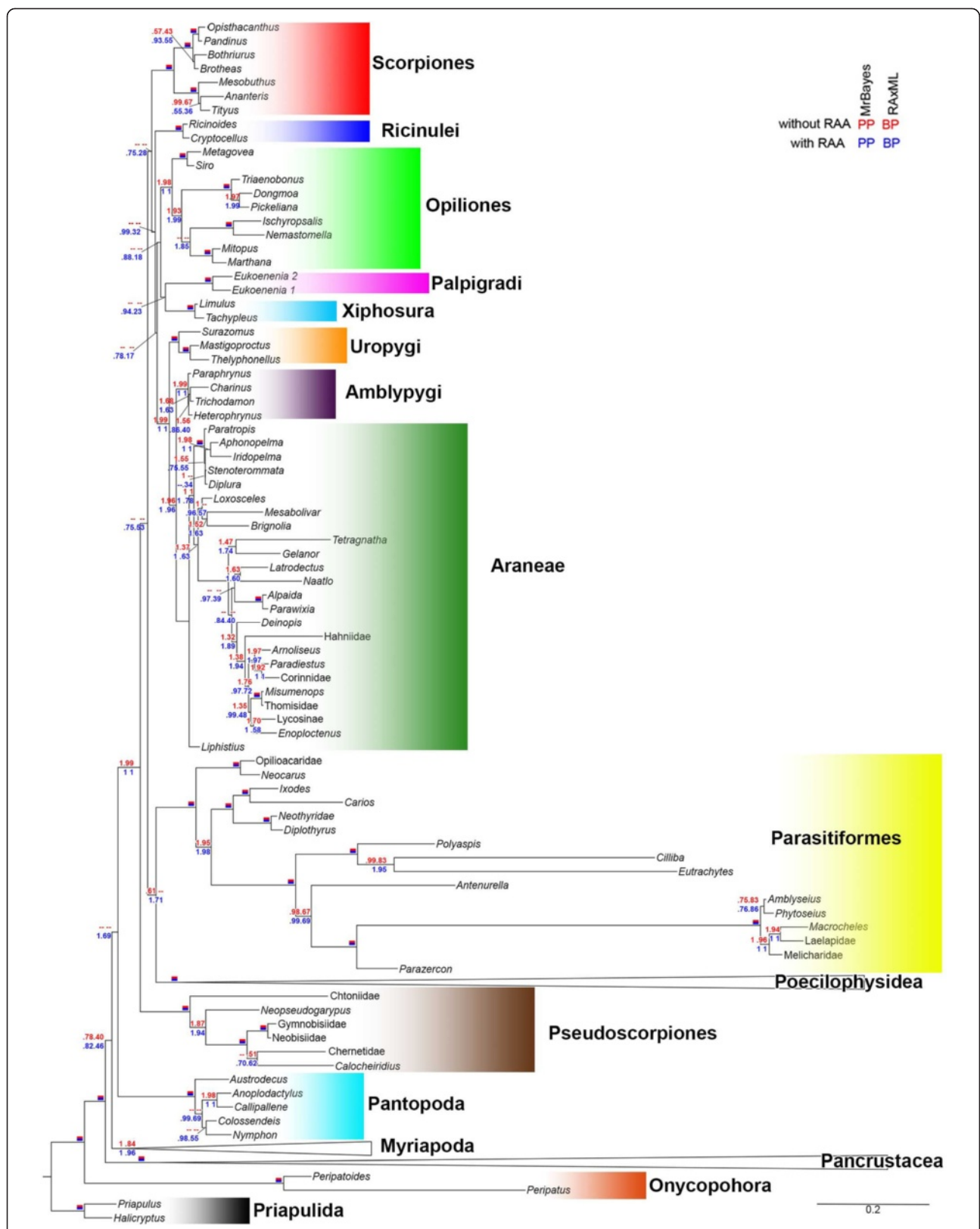


Fig. 4 Relationships between chelicerate orders. Topology inferred by the partitioned (stems, loops and RAAs) ML analysis. For each node, posterior probabilities and bootstrap support values are given for three (blue) and two (red) partition analyses. If all values are 100 %, the node is marked with a solid red and blue rectangle. Dashes indicate internodes not recovered in all analyses

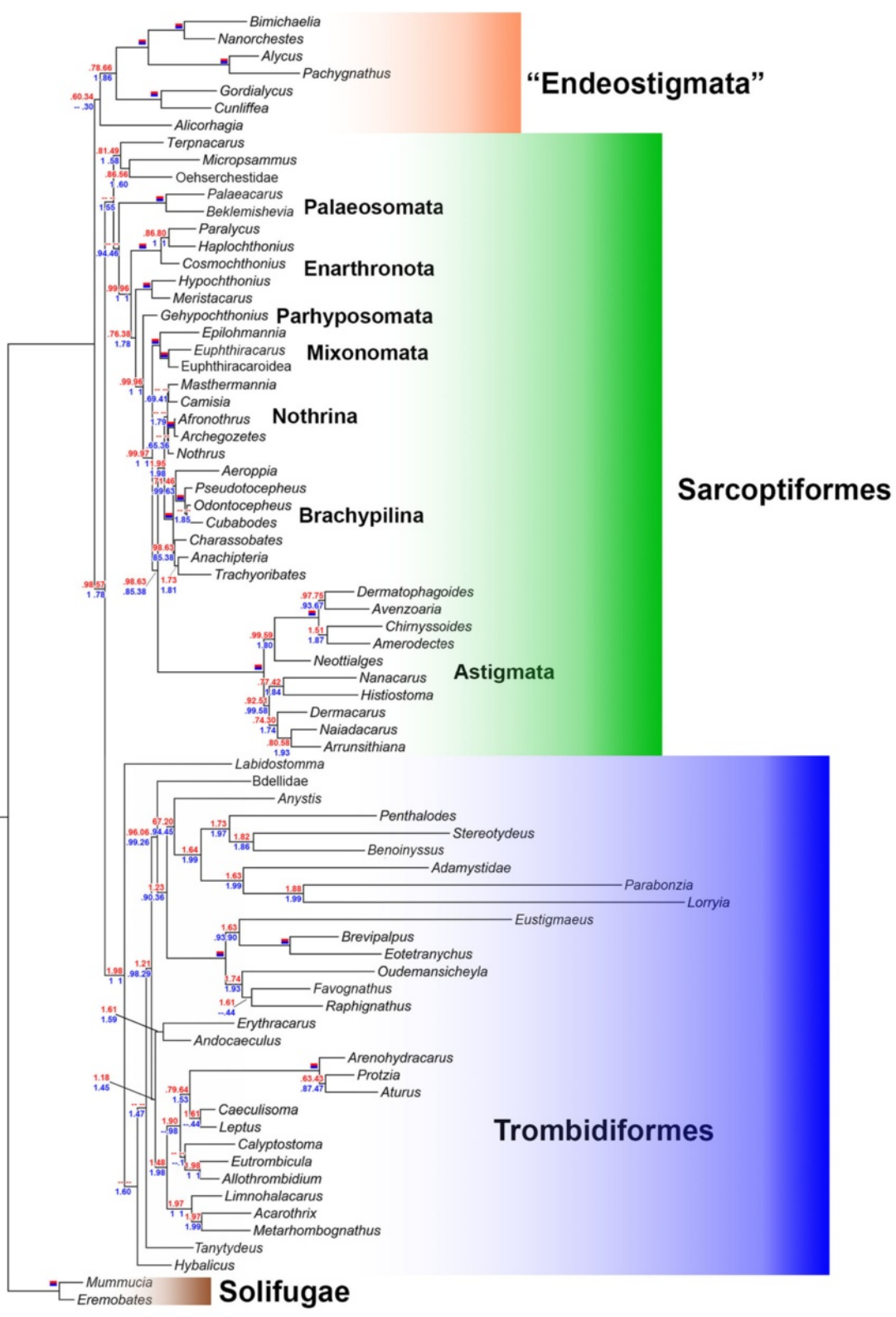


Fig. 5 Relationships of Poecilophysidea (Acariformes and Solifugae). Maximum likelihood tree, partitioned analysis (stems, loops and RAAs). See the Fig. 4 caption for conventions in showing node support values

Table 1 Hypothesis testing using AU statistics

Hypothesis	Without RAAs		With RAAs	
	δLnl	AU	δLnl	AU
Best tree without RAA	–	0.768	82.6	0.027
Best tree with RAA	43.3	0.135	–	0.819
Cormogonida	11.3	0.518	12.3	0.358
Mandibulata	18.7	0.353	5.5	0.724
Arachnida	22.2	0.263	27.5	0.219
Pedipalpi	39.8	0.104	145.9	0.002
Haplocnemata	66.3	5e-005	72.8	0.004
Dromopoda	68.9	0.014	86.8	3e-004
Dromopoda (Solifugae excluded)	9.3	0.622	17.1	0.403
Monophyletic Acari	34.5	0.074	43.6	0.038
Acaromorpha	56.4	0.003	79.8	0.001
Parasitiformes + Opiliones	14.7	0.485	29.7	0.127
Parasitiformes + Pseudoscorpiones	2.0	0.729	13.9	0.392
Cryptognomae	14.6	0.430	33.6	0.173
Non-Euacariformes	60.9	0.124	11.6	0.379
Sarcoptiformes	14.7	0.434	13.0	0.471
Eupodina	74.6	0.007	130.2	3e-004
Bdelloidea	37.5	0.208	46.3	0.041
Anystidae	39.1	0.217	35.7	0.208
Anystina	37.2	0.062	51.5	0.009

P-values below $P < 0.05$ (bold) indicate that a particular hypothesis can be statistically rejected.

and Dromopoda (Haplocnemata + Scorpiones + Opiliones [9, 10, 29]).

Higher-level relationships of Acariformes

Our molecular tree of the Acariformes disagrees with influential, morphology-based ideas on their systematics [14]. In our topology, the earliest basal divergence of Acariformes was “Endeostigmata” (Fig. 5), which has been considered as basal Sarcoptiformes based on morphological evidence [12]. We recovered “Endeostigmata” as a monophyletic sister group to the remaining Acariformes (a grouping referred hereafter to as “Euacariformes” for sake of simplicity) in all analyses, except for the stem and loop Bayesian analysis, which could not resolve the position of *Alicorhagia*. The non-endeostigmatan acariform clade received moderate (two partitions: $PP = 98$, $BP = 57$) to high support (three partitions: $PP = 100$, $BP = 78$). When applying a negative constraint against “Euacariformes”, neither the most probable tree sampled by Markov chain, nor the majority rule consensus tree recovered Sarcoptiformes (most of Endeostigmata + Oribatida + Astigmata), but a topology with most of Endeostigmata grouped with Trombidiformes. Sarcoptiformes sensu [12] could not be rejected in the AU test, however.

Except for Paleosomata, the remaining sarcoptiform backbone tree is well supported, especially after including RAAs, and roughly congruent with the relationships inferred from morphology: (Palaeosomata, (Enarthronota, (Parhypsomata, (Mixonomata, (Astigmata, (Brachypilina, Nothrina)))))). Here we note a major improvement in our tree over previously published DNA-based phylogenies with respect to the position of Parhypsomata; the latter is now consistent with morphological evidence (i. e., the presence of the opisthotal glands). The hypothesis of a paraphyletic Desmonomata with respect to Astigmata remains to be tested since a key taxon, Malaconothroidea, was not included.

Trombidiformes (Sphaerolichida plus Prostigmata) was recovered with high support ($PP = 100$, $BP = 98-100$). The trombidiform tree is unstable in the basal portion, depending on whether or not RAAs were included. Regardless of different analyses, Labidostommatina comes close to the trombidiform root. Among the major divisions of Prostigmata, Eleutherengona was recovered with a high support ($PP = 100$, $BP = 100$); Eupodina was not recovered, Halacaridae (marine mites) was placed as the sister group of Parasitengona, receiving 100 % of PP in all Bayesian analyses and moderate to high support in ML analyses ($BP = 48-89$); Adamystidae and Anystinae, usually classified in Anystina, were found nested within several Eupodina; and Bdellidae was recovered as basal Prostigmata. Neither the family Anystidae nor the superfamily Bdelloidea was recovered as a monophyletic group. The best tree recovered after constraining monophyletic Eupodina was rejected by the AU test. Hypotheses obtained after constraining the clade Caeculidae+ Erythracaridae+ Parasitengona (i.e., excluding the sister group relationship between Halacaridae and Parasitengona), and Bdelloidea (Cunaxidae + Bdellidae) were rejected by the analyses including RAAs. The tree obtained after constraining monophyletic Anystidae could not be rejected.

Assessing the phylogenetic signal: likelihood mapping, consensus networks, and internode certainty

To facilitate visual comparison among alternative topologies after including RAAs and excluding the fast evolving lineages detected through the Bayesian Relative Rates test (Fig. 2a), we employed Likelihood Mapping (Fig. 2c). In three cases the proportion of quartets resolved as the preferred hypothesis (i.e. the three-partition analysis) increases after including RAA and excluding rapidly evolving lineages: Chelicerata (=Pycnogonida + Euchelicerata) – 40.4, 47.7, and 60.1 %, respectively); “Euacariformes” (32.3 %, 34.5 %, 54.3 %); and the placement of Halacaridae as sister group of Parasitengona (41.3 %, 46.5 %, 55.3 %). The other two hypotheses Labellata (=Amblypigi + Araneae) and Poecilophysidea (=Acariformes + Solifugae) have a slight decrease in support when

RAAs are included and present a slight (Labellata, 27.6 %, 24.4 %, 28.4 %) or great (Poecilophysidea, 42.5 %, 39.1 %, 72.2 %) improvement in support after exclusion of long branches.

Consensus networks based on bootstrap trees detected similar regions of conflict for stems and RAAs (Fig. 3), i.e. the basal portion of the tree; the relationships among Eichelicerate orders (except Tetrapulmonata and Poecilophysidea); and Endeostigmata relative to basal “Euacariformes”, especially Sarcoptiformes. The consensus networks from the analyses of the loop partition show the poorest resolution, especially for the basal splits (relative to Pycnogonida, Myriapoda, Pancrustacea), Eichelicerata, and for Trombidiformes relative to Sarcoptiformes. Each partition had a representative of taxa otherwise recovered as “Sarcoptiformes” misplaced closer to “Endeostigmata”. For loops it was Astigmata, for stems it was Palaeosomata, and for RAAs it was the genus *Micropsammus*.

The consensus network based on the combined dataset shows no conflicting bipartitions between “Endeostigmata” and “Euacariformes”. The “Sarcoptiformes” network contrasts to the reticulate structure of the Trombidiformes consensus network by being much more tree-like.

Numerical results for the Internode Certainty (IC) and Internode Certainty All (ICA) match perfectly the consensus networks in detecting the most conflicting regions of the three and show a gain in resolution after including RAAs (Additional file 4).

Discussion

Strengths and limitations

The inclusion of a large outgroup sample allowed us to compare the performance of our analyses to those of published studies utilizing single-gene, multi-gene or genomic data. The topology presented here for major arthropod lineages —(Pancrustacea, (Myriapoda, Pycnogonida, Eichelicerata)), the so-called Myriochelata or Paradoxosopoda —is commonly recovered by ribosomal gene analyses [33], although support for the alternative hypothesis, Mandibulata, was obtained from combined analyses of morphology and microRNAs [27 and references therein].

The Pancrustacea topology (Additional file 4) is largely congruent to that presented previously [34], sharing its limits and strengths. Copepods are reported as sister-group of Hexapoda with a high support ($PP = 100$, $BP = 86–98$), a controversial result also found when non-stationary models were applied to ribosomal sequences [35]. No putative morphological apomorphy can be found supporting this clade, nor was it recovered in multi-locus analyses [36, 37]. On the other hand, similarly to genomic studies, our results agree in recovering the endognath hexapods Nonoculata (Protura + Diplura),

Oligostraca (Branchiura + Ostracoda + Mystacocarida), and the close relationship between barnacles and Malacostraca [37].

For relationships among the orders of Eichelicerata, we found no resolution, except for Tetrapulmonata and Poecilophysidea [2–6]. Recent genomic-scale studies [5, 7] also failed in reaching a well-supported hypothesis on ordinal relationships. They do support, however a distinct solution for in-group Tetrapulmonata, with Pedipalpi instead of Labellata, and only one study [5] recovered Poecilophysidea with moderate support. The most surprising result for intraordinal relationships was the recovery of Palpatores (Eupnoi + Dyspnoi) only after inclusion of the RAA dataset, a clade recovered in combined analyses including fossils and genomic scale datasets [38].

Phylogenetic placement of parasitiform mites remains extremely unstable. Parasitiformes was recovered as a basal lineage among Eichelicerata [3, 4, 7], as the sister group of Pseudoscorpiones [2], or in a group including Opiliones and Tetrapulmonata [6]. All of these phylogenies suggest a large gap in the fossil record for Parasitiformes, since the oldest parasitiform mites, ticks and Opilioacarida, were discovered in Burmese amber from the Cretaceous (ca. 100 Ma) [39–42]. By contrast, acariform mites are known from the early Devonian (ca. 410 Ma) Rhynie Chert of Scotland [43, 44], and Paleozoic fossils have been discovered for most of the arachnid orders [45, 46]. The traditional morphology-based groupings, Acari and Acaromorpha (Acari + Ricinulei), were statistically rejected by the AU test employing the three-partition dataset (Table 1). Furthermore, alternative placements of Solifugae also could be rejected.

Multiple alignments of ribosomal gene sequences

In contrast to introns or coding regions, alignment of ribosomal DNA, to be biologically sensible, should account for secondary structure information (*i.e.*, mature ribosomal RNA folded to form stems and loops maintained both by paired RNA regions and ribosomal proteins). Non-conserved RAA regions of rDNA are especially challenging for alignment because both secondary rDNA structures and sequences evolve in these regions. To address this, two alternative analytical approaches, consensus and individual secondary structure based alignments were developed. As part of initial exploratory analyses we tried both methods, but only the former approach is further reported here. In this method, the rationale for homology assessment is accounting for the fact that secondary structure is functionally constrained in ribosomal RNA. This method, however, usually leaves large stretches where the secondary structure is not conserved (Regions of Ambiguous Alignment, RAAs) out.

Congruence of the secondary structure consensus analyses with previous studies, the better resolution of phylogenetic trees, and bootstrap consensus networks with filtered RAAs all indicate that it is possible to take advantage of RAA regions if two conditions are met. First, based on existing eukaryotic secondary structure rDNA models, conserved regions and RAAs should be identified (bracketed), and then each RAA should be independently aligned using an automatic algorithm. It was already noted [4] that if two variable in length regions are separated by very short stretches of conserved secondary structure of sequence, most algorithms fail to recover those conserved regions. Second, accurate guide trees and realistic cost regimes should be employed. All automatic alignment methods rely on a guide tree to order pairwise alignments and a cost regime to produce scores for comparing alternative alignments. In this study, we employed trees from the 95 % credible interval of trees sampled from the posterior for each RAA to avoid biasing our output alignments to a single topology (see the Methods section below). For gap opening and extension costs, any choice is arbitrary since it is not data driven [47]. In a study on direct optimization accuracy, Liu et al. [48] observed that it performed best under a cost regime penalizing four times the gap opening relative to gap extension and substitution. In any case, the use of a single cost regime for all RAAs chosen without regard for data is far from ideal; hence the alignment was masked using a program that identifies random similarity within multiple sequence alignments based on Monte Carlo resampling within a sliding window [49].

Another method of multiple rDNA alignment is the individual secondary structure approach. In current literature this framework has been tailored to the internal transcribed spacer (ITS2), a region that is spliced out from mature rRNA. This method typically includes a pipeline starting with the *ITS2 database* (source of sequence–structure data), *foursale* or *4SALE* (alignment) and *ProfDistS3* (distance method for phylogenetic reconstruction) or a custom substitution matrix to be implemented in a maximum likelihood environment [50, 51]. This pipeline was extensively applied to ITS2 [52–54] or rarely to other markers, like mitochondrial 16S [55]. The main issue precluding a large scale application of this approach for structural RNA is that automatic free-energy folding algorithms (UNAFold, mfold, RNAstructure, ViennaRNA) may ignore secondary structure constraints if there is too much conflict with free-energy structures. As a result inferring individual secondary structures for smaller regions followed by a manual quality check should be done (a time consuming task). Secondary structure models are available for a large set of almost complete SSU and fragments of LSU in the SILVA database [56] but most variable regions of LSU

remain to be inferred individually. We hope that the consensus secondary structure alignments presented herewith stimulate future studies in the area.

Concerning the procedure employed here for incorporating regions of ambiguous alignment (RAAs), a better solution may be achieved by jointly inferring trees and alignments in a Bayesian framework. In this case, the guide tree and cost regime problem is solved through the joint estimation of phylogenetic tree and sequence alignment, something that is implemented in *Bali-Phy* [57]. Despite problems concerning how to model indels, analyses of large data sets are computationally unfeasible for this program (the largest data set analyzed to date contained 117 terminals, aligned sequences with ~200 bp) [58]. Bracketing the RAAs and aligning them independently would save computational time (since most of non-RAA regions are pre-aligned), would allow parallel calculations for each RAA (91 in our dataset), leading to a biologically more realistic, independent estimate for the parameters of the indel model for each RAA and turning Bayesian alignment into a useful and computationally feasible tool for rDNA genes.

Evaluating previous hypotheses on acariform relationships

The results presented here are in agreement with several insightful previous works but disagree, in several aspects, with current classifications and published molecular phylogenies. Based on morphological characters and *a priori* transformation series, OConnor [12], presented a higher-level phylogeny of Acariformes showing a basal split among Trombidiformes and Sarcoptiformes (including most Endeostigmata). This dichotomic view of Acariformes phylogeny, however, was first proposed by Enzo Reuter in 1909 [59] and is followed in current, broadly accepted classifications [14]. In OConnor's scheme [12], Sarcoptiformes was supported by the toothed rutellum, differentiated prodorsal region, and the loss of solenidia from tarsi IV. The implied character polarity, however, is not congruent with the relationships inferred in this study. The “Endeostigmata”—Alycidae, Nanorchestidae, Nematalycidae, Alicorhagiidae — share the above mentioned characters (the rutellum was thought to be absent in Nematalycidae, but this was recently dismissed [60]) and was recovered as sister group of remaining Acariformes — the Euacariformes. Moreover, the differentiated prodorsal region is present in the Acariformes' sister group, Solifugae, hence turning this character state into a plesiomorphy.

The problem here is the absence of unequivocal morphological apomorphies supporting either the clade including most Endeostigmata or Euacariformes. We found a single character in support to the Endeostigmata clade: Nematalycidae presents cuticle projections known as palettes, which are oriented so that their edges are

perpendicular with respect to the surface annuli, aiding in the grip of the body against the interstitial surface [61]. These palettes have also been found in the Nanorchestidae [62] where their function is uncertain, and some superficially similar structures may be seen in *Bimichaelia* (Alycidae) (SEM pictures in [63]).

The relationships of Sarcoptiformes recovered here (endeostigmatan families Micropsammidae, Terpnacaridae and Oehserchestidae plus Oribatida, including Astigmata) are in general agreement with the morphology-based analysis by Norton [64]. It is interesting to note that most molecular studies [2, 15, 16] did not support these relationships, with *Gehyppochthonius* (Parhyposomata) assuming a more basal position. This basal position seems questionable given the presence of opisthonotal glands, a trait shared among certain derived oribatids and the Astigmata. Our topology also allows testing more specific evolutionary hypotheses, such as the position of the enigmatic mite family Pediculochelidae (*Paralycus*). Previously, this family was placed among Prostigmata, Astigmata or Endeostigmata, but our study places it among protoplophoran oribatids (Fig. 5), supporting the hypothesis that its unusual morphology resulted from paedomorphosis as suggested by Norton et al. [65]. This grouping has a very high support, and it shares many unique modifications in rDNA secondary structure.

Trombidiformes, diagnosed by the loss of primary segmentation of the anamorphic segments AN and PA, the reduction to fewer than four pairs of setae on hysterosomal segment C, and fewer than three pairs on segments D and E [12], was recovered in this study with a high support. In contrast, Prostigmata was recovered with low or no support, mainly due to instability of *Labidostomma* with respect to Paratydeidae and *Hybalicus*. Despite this, *Labidostomma* was consistently recovered in a basal position, which is congruent with the absence of secondary tracheal openings, termed neostigmata, in this taxon [66], presence of six pairs of prosomal setae (a simplesiomorphy shared by Sphaerolichida), two pairs of bothridial setae, and chelatae chelicerae bearing two setae on the fixed digit. In previous classifications, Labidostommatidae was included in Eupodina [11, 13], but its basal position was already recognized in recent classifications [67]. The same simple condition of the stigmata is known for Rhagidiidae, another possible basal taxon (not included in this analysis), while in derived Prostigmata, the stigmata can be either subchelicerous or dorsal (neostigmata).

In our analyses, Eleutherengona was recovered monophyletic, but Eupodina and Anystina were not. It is not surprising since Eleutherengona is a clade supported by many morphological apomorphies: the loss of the third nymphal stage in most taxa; the presence of a sclerotized aedagus in males; the cheliceral bases contiguous or

fused, with loss of independent movement; the chelicerae with the movable digit pointed and partly retractable; the genital and anal openings adjacent; a correlated loss of the epimeral organs and genital acetabula; and the fusion of the femoral segments [13, 68]. In this study, however, sampling was limited to Raphignathae, leaving out the Heterostigmata, a clade hypothesized to be the sister group to Raphignathae [13, 68].

A key group in interpreting the non-monophyly of Eupodina and Anystina are the marine mites (family Halacaridae). Traditionally, they were regarded as closely related to the terrestrial predacious superfamily Bdelloidea (=Bdellidae + Cunaxidae) in Eupodina, but we found them to be the sister-group to Parasitengona (includes chiggers and fresh water mites). At first glance, this appears as a major conflict among morphological and molecular data, however, now it is apparent that the traditional classification relies on superficial resemblance rather than accurate consideration of morphology. The position of Halacaridae as Parasitengona's sister group was anticipated by Witte [69] who found six synapomorphies for this relationship: (1) the palp with conspicuous spiniform setae (in most Halacaridae), which are probably homologous to the tibial spine that opposes to the palptarsus and forms the "thumb-claw" process in many Anystina and Eleutherengona; (2) the fixed digit of the chelicerae is reduced, and the movable digit is often hook-like and serrate; (3) sigmoid pieces are sclerotized structures projecting ventrally into the infracapitulum and bend anteriorly, continuing under the capitular saddle as a sclerotized supporting bar, serving as a stabilizing and protecting element for the neostigmatal processes during movement of the cheliceral bases [70]; despite the absence of peritremes in Halacaridae, the sigmoid pieces are present and the chelicerae extend posteriorly beyond them, a similar condition to Parasitengona; (4) the sigmoid pieces extend their proximal portions interiorly, and cheliceral protractor 1 originates on them; (5) a similar mechanism of chelicerae protraction, through forward rotation of the tip of the sigmoid piece; (6) the internal podocephalic canal, although the reduction of the podocephalic glands observed in Halacaridae downplays this character state.

Similarly, several other traditional groupings could not be confirmed in our analyses, suggesting that the current view on the phylogenetic relationships of Prostigmata needs to be revised. Neither Bdelloidea nor Anystidae were recovered. The AU test rejects the hypothesis of a monophyletic Bdelloidea in the three-partition analyses. Morphological support for this placement, however, needs to be investigated more closely in the future. Both Bdellidae and Anystidae have especially conflicting regions of the phylogeny as indicated by the consensus networks (Fig. 3) and IC and ICA values (Additional file 4).

In disagreement with currently accepted classification, but similarly to our results, Otto [71], recovered Anystidae as a non-monophyletic group. In fact, the hypothesis of a close relationship of Anystinae and Erythracarinae was based mainly on the overall appearance — ‘spider-like’ and ‘long legged, fast and mostly reddish’. Additional characters were: legs arranged in a radiating way, idiosoma lacking a sejugal furrow, movable digit hook-like, presence of 1–3 claw-like setae on the palp tibia and soft cuticle. Since most of these traits were convincingly dismissed as shared apomorphies in a previously published work [71], which is supported by our molecular evidence, we propose to remove the subfamily Erythracarinae from Anystidae and consider it again as a separate family, Erythracaridae Oudemans, 1936 stat. ressur., as it already was used in prior literature [72] (zoobank.org:act:041D4A94-6A5F-4D94-A018-A2C176E6BB1F).

Likelihood mapping, consensus networks and certainty indices as techniques for phylogenetic signal assessment

Many phylogenetic studies rely solely on Bayesian posterior probabilities and bootstrap proportions to give some indication on how reliable are the inferences, despite other methods of estimation of the robustness of phylogenetic signal that are available. In this study, Likelihood Mapping, Consensus networks on bootstrap trees and the *Internode Certainty* (IC) and *Internode Certainty All* (ICA) indices gave interesting insights on the strength of the data and hence, on the reliability of the analyses.

First, Likelihood Mapping was shown to be very sensitive to systematic errors due to Long Branch Attraction (LBA): exclusion of fast evolving lineages resulted in a larger proportion of quartets resolved according to the recovered topology in all tested groupings, except for Pedipalpi (Fig. 2c). This sensitivity considerably reduces the utility of quartet mapping to test data suitability for phylogenetic analyses, since an alignment may be able to resolve most quartets due to LBA. In our analyses this seems to be the case for the loop partition that has similar percentages of star-like and partially resolved quartets as the stem partition (Fig. 2b), but when the loop bipartitions are plotted in the consensus networks (Fig. 3) it is clear that there is more discordance. Despite these limitations, Likelihood Mapping, jointly with consensus networks, can be useful for LBA detection.

Measures of nodal certainty (IC and ICA) were used to detect the presence of conflicting phylogenetic signals (as opposed to uncertainty or weak signal) in specific nodes of our phylogeny (Additional file 4). Low values of these indices indicate the presence of a conflicting signal for Myriapoda, Pancrustacea, Chelicerata (except for Tetrapulmonata and Solifugae), and basal portions of “Endeostigmata”, Sarcoptiformes, and Trombidiformes. Interestingly, for the grouping Solifugae and Acariformes,

the certainty indices are either a unit (all partitions) or approach a unit (stem and loop regions), indicating the absence of conflict for this node.

Conclusions

Our analyses show that ribosomal genes unequivocally support a sister group relationship between Solifugae and Acariformes (Poecilophysidea). In contrast to existing morphological hypotheses, we found that most Endeostigmata (rare deep-soil mites known from the Devonian) represent a major basal divergence that occurred prior to the split between the two hyperdiverse acariform lineages, Sarcoptiformes and Trombidiformes. Morphology suggests a basal split between Sarcoptiformes and Trombidiformes and places endeostigmatans at the root of Sarcoptiformes. Thus, our findings may have a substantial impact on higher-level classification of acariform mites and indicate that using endeostigmatans may greatly improve the accuracy of time estimations on fossil-calibrated phylogenies. We inferred Astigmata (an unranked hyperdiverse group including many medically and economically important species) as the sister group of desmonomatan oribatids, which is consistent with some, but not all morphological and molecular hypotheses. For a long time the Astigmata was treated separately from oribatids, following the idea of the influential mite morphologist, François Grandjean, who suggested independent origins of opisthonotal glands in Astigmata and derived oribatids [73, 74]. Interestingly, a topology similar to Grandjean's idea was recovered previously based on 18S rDNA and EF1- α sequences [21]. Here we show that this grouping may be due to non-phylogenetic signal present in a single rDNA partition.

Trombidiformes (Sphaerolichida + Prostigmata) was recovered with a high support. Previously Sphaerolichida was treated within the endeostigmatan lineages [75], but later they were moved to Trombidiformes based on the absence of the rutella, primary segmentation, anamorphic segments AN and PA, and the presence of fewer than four pairs of setae on segment C and fewer than three pairs on segments D and E [12]. This was a revolutionary idea at this time, and our analysis strongly supports this grouping. Labidostommatina was recovered as a basal Prostigmata. Among other major lineages of Prostigmata, only Eleutherengona and Parasitengona were recovered. Marine mites (Halacaridae), a globally distributed group with controversial phylogenetic affinities, were recovered as the sister group of Parasitengona, with high support. This is in agreement with a previous detailed morphological study [69], indicating that the current view of placing marine mites close to the terrestrial predacious Bdelloidea should be abandoned and the superfamily Halacaroidea moved to the infraorder Anystina. Our results and previously published morphological data indicate

the need for major rearrangement in the family Anystidae (large, fast moving mites): Erythracarinae is elevated to family rank, Erythracaridae stat. ressur., and treated separately from the Anystidae.

Molecular evidence presented in this paper calls for further investigation of possible shared apomorphies of the early derivative Endeostigmata clade (Alycidae, Nanorchestidae, Nematalycidae, Alicorhagiidae) and the clade comprising the remaining Acariformes; including a greater gene and taxon sampling for clades traditionally placed in Eupodina and Anystina (Trombidiformes) and Palaeosomata (Sarcoptiformes), since consensus networks and internode certainty measures (IC, ICA) show that these tree regions are especially conflicting. Finally, our results demonstrate that regions of ambiguous sequence-similarity alignment, when aligned using secondary structure information, may provide useful phylogenetic signal.

Methods

Taxonomic sampling and sequencing

We sequenced the small and large subunit nuclear rDNA genes for a total of 118 taxa. Another 110 sequences were retrieved from GenBank. GenBank accession numbers are listed in Additional file 1 along with details on vouchering. The most distant out-group was Priapulida (2 spp). Distant outgroups included all main panarthropod lineages: 2 Onychophora, 32 Hexapoda, 25 Crustacea and 12 Myriapoda. The Chelicerata ingroup comprised 150 taxa (newly sequenced taxa are indicated in parenthesis): 5 Pantopoda, 2 Xiphosura, 9 (3) Opiliones, 24 (21) Araneae, 4 (2) Amblypygi, 3 (2) Uropygi, 7 (5) Scorpiones, 2 (2) Palpigradi, 2 (1) Ricinulei, 2 (1) Solifugae, 6 (4) Pseudoscorpiones, 16 (16) Parasitiformes, and 73 (60) Acariformes.

Molecular work was conducted at the Universidade de São Paulo by ARP following protocols described in [4] and published primers for SSU [76, 77] and for LSU in [78–82] and at the University of Michigan by PBK using previously described protocols and primers [83].

Ribosomal DNA and secondary structure alignment

Due to its ease of amplification and its among-region evolutionary rate heterogeneity, ribosomal DNA enjoyed a pioneering role in molecular phylogenetics, leading to the so-called “new animal phylogeny” [84, 85], and being – even in the genomic era – the sole source of molecular data for several large clades (*e.g.*, [6]).

Yet serious difficulties exist in proposing reliable hypotheses of nucleotide homology for regions of ambiguous alignment (RAAs) using standard multiple sequence alignment procedures. Automated alignment based on cost regimes and maximizing sequence identity proved to be a failure in this respect, mainly due the among-region evolutionary rates of substitution and the incidence of insertions and deletions (indels) [47]. Congruence-based

shortcuts, like direct optimization implemented in POY [86, 87], add the artifact of over-optimization or epistemological character non-independence [88, 4].

By providing a causal framework for proposing molecular homology, secondary structure consensus guided alignment contrasts with the methods that rely on simple sequence similarity or congruence optimization. We refer to this framework as ‘causal’ to highlight: (i) that it considers as the main evidence for homology the compensatory changes driven by stabilizing natural selection on secondary structure; (ii) the way in which it contrasts with methods that do not assume any cause for among-sequence divergence and rely only on identity (inheritance); and (iii) the differences from an agnostic standpoint on evolutionary process, as embodied by the congruence methodologies.

Employing secondary structure for aligning sequences has, however, at least two potential drawbacks. First, since automated methods, like *RNASalsa* [89], remain largely underappreciated, secondary structure guided alignment still relies on extremely time-consuming manual alignment. Second, consensus secondary structure alignment cannot be used if there is no common secondary structure among taxa, leaving out some variable regions that may contain useful phylogenetic information.

For the first problem, one cannot offer any easy solution, except for providing a carefully annotated alignment that, alongside other works [17, 34, 90], can give a useful template for adding new sequences. Fasta files containing our alignments labeled with the SSU and LSU rRNA secondary structure are given in Additional file 2. Here, matching parentheses and dots were used to indicate stem and loop regions, respectively; structural helix numbering is given after [91], except for the SSU variable region 4 (V4) for which notations of [92, 93] were used.

Our secondary structure alignment procedure follows [94], except for employing the program *BioEdit* 7.2.1 [95] for sequence editing. Reference rRNA structures [4, 34, 96] were used for this alignment. When regions were too variable, potential pairings based on thermodynamics were explored in *mfold* [97] and consensus secondary structures were inferred in *RNAalifold* [98]. Nucleotides in paired regions whose secondary structure was sustained by compensatory mutations across the entire data set were considered as homologous. Regions inferred to be ambiguously aligned were classified into regions of expansion and contraction (REC), non-pairing regions of ambiguous alignment (RAA) and regions of slipped-strand compensation (RSC) [99].

We addressed the site homology for RAAs without violating the positional homology inferred for the structural aligned regions. Briefly, regions of ambiguous

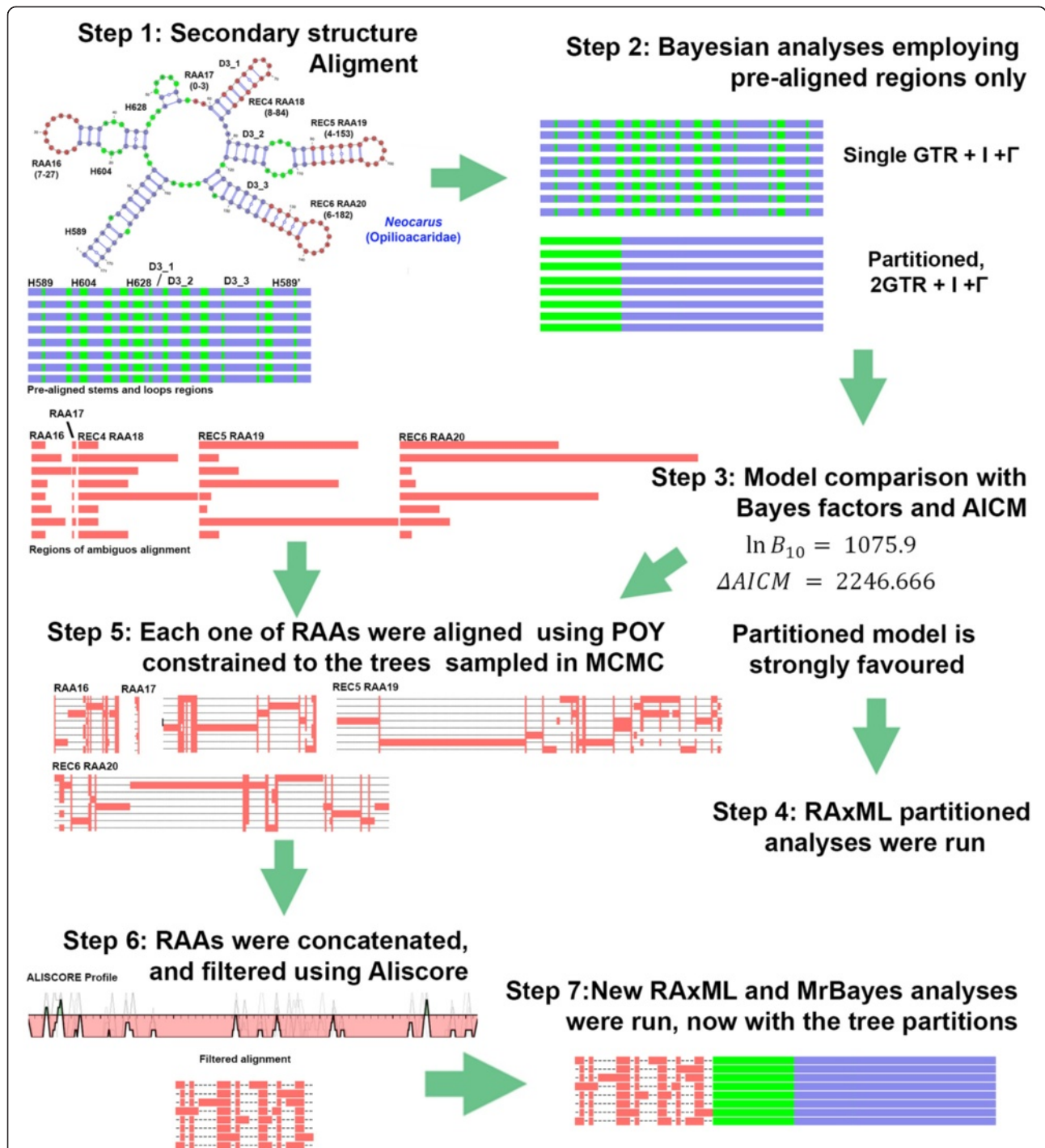


Fig. 6 Flowchart of rDNA consensus secondary structure sequence alignment and analyses. Step 1. Multiple alignment guided by secondary structure information detects paired, unpaired and regions of ambiguous alignment (RAAs). Step 2. The RAAs are excluded from the alignment for the first set of Bayesian and ML inferences. Two Bayesian analyses were run, considering either a single model for the concatenated loop and stems or separate substitution models for each of these partitions. Step 3. Bayes factors estimated by Harmonic Mean and AICM strongly favor the partitioned model. Step 4. Partitioned *RAxML* analyses were run. Step 5. Regions of ambiguous alignment (RAAs) were aligned individually using a gap opening cost of four and equally weighted gap extension and substitutions [48] and Markov chain stationary trees from Bayesian analyses as guide trees. Step 6. The resulting concatenated alignment was filtered using *Aliscore*. Step 7. Bayesian and ML analyses employing RAAs alignment as a third partition resulted in topologies largely congruent to those of the two partition analyses, but their nodal support was higher in most cases

alignment were initially excluded from preliminary analyses (Step 1–4, flowchart for our methodology on Fig. 6). Then we adapted an existing methodological framework [48] to extract phylogenetic signal from RAAs. Usually, in direct optimization analyses, the choice of weighting values for gaps (opening and extension) and substitutions is driven by maximizing the among-data congruence [100]. When combining conflicting data partitions, this approach can result in permissive gap costs, leading to an artifact where the optimization algorithm overweighs the signal from static partitions relative to those that are dynamically aligned. This phenomenon was discovered independently in two studies and, referred to as ‘over optimization’ [88] or ‘epistemological character non-independence’ [4]. Both studies showed that in order to maintain hypothesis test severity, partitions must be aligned independently prior to concatenation analyses. Liu et al. [48] suggested a cost regime that includes a four-fold gap opening cost relative to equally weighted gap extension and substitution costs and a guide tree as a combination leading to the best nucleotide homology estimates when employing direct optimization. We applied this cost regime for each of the RAAs independently and, instead of a single guide tree, we used MCMC stationary trees sampled from the posterior (Fig. 6, step 5). Then all regions were concatenated, masked in *Aliscore v2.02.2* (with the window size $w7$: Fig. 6, step 6) and further processed in *Alicut*. *Aliscore* identifies random similarity within multiple sequence alignments based on Monte Carlo resampling within a sliding window. The method infers similarity profiles from pairwise sequence comparisons and subsequently calculates a consensus profile. Thus, consensus profiles identify dominating patterns of non-random similarity or randomness within sections of multiple sequence alignments. *Alicut* simply slices sites in the alignment according to *Aliscore* output [49, 101]. The concatenated alignment was very long (28,589 nucleotide positions) due a large number of autapomorphic indels. After *Aliscore* filtering, 1,703 positions were included in the analyses (the RAA partition).

Data sets were merged in *FASconCAT v1.0*. This script was also used to read secondary structure mask (in dot-bracket format) and generate a list of pairing and non-pairing positions for downstream phylogenetic programs, like *RAxML* or *MrBayes* [102].

Model selection and phylogenetic analyses

The GTR + G + I model was selected for all three partitions based on their lowest AIC scores calculated in *jModeltest v. 0.1.1* [103]. We then analyzed secondary structure alignable regions (Fig. 6, step 2) in *MrBayes v.3.2.2* [104], employing four chains, with 2×10^7 generations, and sample frequency of 1:1000. Convergence was evaluated in *Tracer v.1.6* [105] for continuous

parameters; results are given in Additional file 3. Because *MrBayes* default branch length priors and starting tree branch length priors may lead to a gross overestimation of tree length [106, 107], we conducted additional analyses with the starting tree branch length prior set to 0.001 and the prior on the branch length set to 0.01. Results were identical with those run with the default priors. Using stems and loops as two separate partitions improved the harmonic mean of lnL over unpartitioned analysis by 1075.9, and the more accurate AICM estimated using *Tracer* of 309858.6 (SD = 0.928) for unpartitioned and 307611.9 (SD = 0.738) for partitioned, leading to a Δ AICM of 2246.7 for the unpartitioned analysis, and hence showing a very strong support for the partitioned analyses [108] (Fig. 6, step 3).

For secondary structure informed analyses, we tested 6- and 7-state rDNA-specific models [109] in a Maximum Likelihood context along with the traditional nucleotide (4x4) model in *RAxML 8.1* [110] run on the CIPRES Portal [111]. Employing these biologically sound models, nevertheless, resulted in disappointing outcomes. Letsch & Kjer [22] attributed their also disappointing results, to the poor fitting to loop regions that are subject to saturation. Our tests of saturation in *DAMBE 5.5.2* [19, 20, 112], however, could not detect saturation in any partition (see the Results section above), and an alternative explanation for the poor performance of rRNA models is needed. We suspect it is due to the strong compositional bias toward adenines (see above). Trees inferred with the rDNA-specific models are given in Additional file 4. Topologies discussed in the results sections were based on the simple 4x4 model. For ML analyses 1000 bootstrap replicates were ran.

Data exploration

The nucleotide composition of our dataset was visualized by the tetrahedric plot function of the R package *Compositions* [113]. The program *TREE-PUZZLE 5.2* [114] was employed to explore whether the base composition of each sequence was identical to the average base composition (chi-squared test at a 5 % level).

For comparing rates of molecular evolution, a Bayesian Relative Rates test [115] was conducted using a custom R script (Additional file 5).

Two graphical approaches were employed to explore the phylogenetic signal in each data partition and combined datasets, the Likelihood Mapping (LM) [116] as implemented in *TREE-PUZZLE*, and the consensus networks [117] using trees generated by bootstrap re-sampling in *SplitsTree4 V4.13.1* [118]. A maximum of 1000 bootstrap replicates for the combined dataset and each partition (loops, stems and concatenated RAAs after filtering in *Aliscore/Alicut*) were run in *RAxML*, with a threshold proportion of 0.1.

For the LM, we first assessed the GTR parameter values from the above mentioned unpartitioned ML analysis and then ran partitioned analyses without assigning any grouping scheme (to verify the proportions of further resolved quartets and hence the phylogenetic signal for each partition). Then different groupings were employed to verify the proportion of quartets that support different hypotheses.

Hypothesis testing

Many long-standing hypotheses on chelicerate, and more specifically, mite relationships (Table 1) were evaluated using the phylogenetically explicit Approximately Unbiased (AU) test in *Consel* [119]. To generate input trees for this test, the best tree representing each hypothesis was calculated in *RAxML* using the constraint command. In the case of a negative constraint, not implemented in *RAxML*, we used the best Bayesian tree (*MrBayes*, two independent runs, 2×10^7 generations each).

Quantifying incongruence

Although bootstrap proportions were long considered as measures of robustness of phylogenetic signal, they can be extremely misleading in the presence of phylogenetic conflict among different genes. For example, a node with a bootstrap support of 100 % may appear to be well-supported, although there are a large proportion of gene trees that conflict with that node [120]. Similarly, when applied to bootstrap trees, a well-supported node can represent one overwhelmingly prevalent node over an array of different reconstructions (each having a negligible frequency) or a mixture of two conflicting reconstructions (most frequent and less frequent). In the latter case, bootstrap support can be misleading. To account for this situation, Internode Certainty (IC) and Internode Certainty All (ICA) indices were introduced using Information Theory [121]. IC calculates the degree of certainty for an internode by considering the frequency of the bipartition defined by the internode jointly with that of the most prevalent conflicting bipartition. ICA calculates the degree of certainty for a given internode by considering the frequency of the bipartition defined by this internode *versus* all conflicting bipartitions. Internode certainty values near zero indicate the presence of an almost equally supported bipartition that conflicts with the inferred internode, whereas values close to one indicate the absence of conflict. IC and ICA indices were calculated in *RAxML*.

Availability of supporting data

All the supporting data are included as additional files.

Additional files

Additional file 1: Sampling data and taxonomy. Systematic classification of terminals and GenBank accession numbers for sequences and collection information for newly sequenced species and their accession numbers in the CT-UFMG (Coleção taxonômica da Universidade Federal de Minas Gerais) and MZSP (Museu de Zoologia da Universidade de São Paulo). (RAR 93 kb)

Additional file 2: Secondary structure alignments. A-B: Two FASTA files containing consensus secondary structure alignments of the 18S and 28S genes. Secondary structure notations masks: brackets “()”, “{}”, and “< >” indicate paired sites; dots “.” indicate unpaired sites, and asterisks “*” indicate regions of ambiguous alignment (RAAs). Notations for stems and variable regions follow [91] and [92, 93] for the SSU variable region 4 (V4). Length variation for each RAAs is also given separately. (RAR 211 kb)

Additional file 3: Assessing convergence of MCMC analyses. Plots of LnL of the stationary phase of the Bayesian analyses and ESS values for parameter estimates. (RAR 600 kb)

Additional file 4: Trees obtained in analyses using different partitions and nucleotide models; internode certainty measures for 3- and 2-partition phylogenetic analyses. A-F: Tree topologies for analyses combining stem and loop regions, all data partitions, and assigning 6A and 7A models for stem and GTR + I + G models for loops. Bootstrap proportion values are given for each node. G: For each node, Internode Certainty (IC) and Internode Certainty All (ICA) indices are shown on ML trees inferred from stem and loop (A) or all regions (B). IC is followed by ICA. (RAR 368 kb)

Additional file 5: R script for the Bayesian Relative Rates test. This script implements a Bayesian Relative Rates Test [116], *i.e.* sums of the branch lengths from a node ancestral to the node of interest and its tips for postburnin trees sampled from the posterior. It provides an estimate of the mean, 95 % credible interval, maximum and minimum values for each terminal. (R 6 kb)

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

ARP designed this study, performed specimen collecting, rDNA sequencing, secondary structure alignment and most analyses, and wrote the manuscript. PBK participated in the study design, conducted field collecting, sequencing and a few earlier exploratory analyses, wrote Perl scripts, and discussed results. Both authors read and approved the final version of the manuscript.

Acknowledgements

This article is part of ARP Post-doctoral research developed at the Universidade de São Paulo supported by FAPESP (2010/00114-8) under supervision of Dr. Ricardo Pinto da Rocha. PBK is supported by United States Department of Agriculture APHIS-S&T Cooperative Agreement (14-8130-0466-CA), and grants from the Ministry of Education and Science of Russia (№ 6.1933.2014/K) and the Russian Foundation for Basic Research (15-04-05185 A). The final revision of the manuscript was greatly facilitated by a visit of PBK to Brazil sponsored by the Science Without Borders Program (CAPES PVE 88881.064989/2014-01). We thank Drs Michael Zuker (Rensselaer Polytechnic Institute, Troy, New York) for explaining algorithms implemented in *mfold* and *unafold*, Roy Norton (State University of New York, Syracuse, New York, USA) for identifying several oribatids, Dong Liu (Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences) for identifying *Euphthiracarus pulchrus*, Katarzyna Jesionowska (University of Szczecin, Poland) for identifying the eupodid specimen, Pedro Henrique Martins for the help in identifying most of non-Acari terminals, and Barry M. OConnor, University of Michigan, Ann Arbor, Michigan, USA, for comments on the manuscript. The authors also thank FAPEMIG for partially supporting the publication costs.

Author details

¹Departamento de Zoologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Av. Antonio Carlos, 6627, 31270-901 Belo Horizonte, Brazil. ²Department of Ecology and Evolutionary Biology, University of

Michigan, 1109 Geddes Ave, Ann Arbor, MI 48109-1079, USA. ³Tyumen State University, 10 Semakova St, Tyumen 625003, Russia.

Received: 19 June 2015 Accepted: 19 August 2015

Published online: 02 September 2015

References

- Dunlop JA, Alberti G. The affinities of mites and ticks: a review. *J Zool Syst Evol Res.* 2007;46:1–18.
- Dabert M, Witalinski W, Kazmierski A, Olszanowski Z, Dabert J. Molecular phylogeny of acariform mites (Acari, Arachnida): strong conflict between phylogenetic signal and long-branch attraction artifacts. *Mol Phylogenet Evol.* 2010;56(1):222–41.
- Arabi J, Judson ML, Deharveng L, Lourenço WR, Cruaud C, Hassanin A. Nucleotide composition of CO1 sequences in Chelicerata (Arthropoda): detecting new mitogenomic rearrangements. *J Mol Evol.* 2012;74(1–2):81–95.
- Pepato AR, Rocha CEF, Dunlop JA. Phylogenetic position of the acariform mites: sensitivity to homology assessment under total evidence. *BMC Evol Biol.* 2010;10:235.
- Dunlop J, Borner J, Burmester T. Phylogeny of the Chelicerates: morphological and molecular evidence. In: Wägele JW, Bartolomaeus T, editors. *Deep Metazoan Phylogeny: The Backbone of the Tree of Life.* Berlin: Walter de Gruyter; 2014. p. 399–412.
- Rota-Stabelli O, Daley AC, Pisani D. Molecular timetrees reveal a Cambrian colonization of land and a New scenario for ecdysozoan evolution. *Curr Biol.* 2013;23(5):392–8.
- Sharma PP, Kaluziak ST, Pérez-Porro AR, González VL, Hormiga G, Wheeler WC, et al. Phylogenomic interrogation of Arachnida reveals systemic conflicts in phylogenetic signal. *Mol Biol Evol.* 2014;31:2963–84.
- Klompfen H, Lekveishvili M, Black IV WC. Phylogeny of parasitiform mites (Acari) based on rRNA. *Mol Phylogenet Evol.* 2007;43:936–51.
- Wheeler WC, Hayashi CY. The phylogeny of the extant chelicerate orders. *Cladistics.* 1998;14:173–92.
- Giribet G, Edgecombe GD, Wheeler WC, Babbitt C. Phylogeny and systematic position of Opiliones: a combined analysis of chelicerate relationships using morphological and molecular data. *Cladistics.* 2002;18:5–70.
- Norton RA, Kethley JB, Johnston DE, O'Connor BM. Phylogenetic perspectives on genetic systems and reproductive modes of mites. In: Wrensch DL, Ebbert MA, editors. *Evolution and diversity of sex ratio in insects and mites.* New York: Chapman & Hall; 1993. p. 8–99.
- O'Connor BM. Phylogenetic relationships among higher taxa in the Acariformes, with particular reference to the Astigmata. In: Griffiths DA, Bowman CE, editors. *Acarology VI, vol. I.* Chichester: Ellis-Horwood Ltd; 1984. p. 19–27.
- Lindquist EE. Phylogenetic relationships. In: Lindquist EE, Sabelis MW, Bruin J, editors. *Eriophyoid mites - their biology, natural enemies and control.* Amsterdam: Elsevier Science; 1996. p. 301–27.
- Lindquist EE, Krantz GW, Walter DE. Classification. In: Krantz GW, Walter DE, editors. *A Manual of Acarology.* Lubbock: Texas Tech University Press; 2009. p. 97–103.
- Pachl P, Domes K, Schulz G, Norton RA, Scheu S, Schaefer I, et al. Convergent evolution of defense mechanisms in oribatid mites (Acari, Oribatida) shows no “ghosts of predation past”. *Mol Phylogenet Evol.* 2012;65:412–20.
- Schaefer I, Norton RA, Scheu S, Maraun M. Arthropod colonization of land – Linking molecules and fossils in oribatid mites (Acari, Oribatida). *Mol Phylogenet Evol.* 2010;57:113–21.
- Mallatt J, Craig CW, Yoder MJ. Nearly complete rRNA genes assembled from across the metazoan animals: effects of more taxa, a structure-based alignment, and paired-sites evolutionary models on phylogeny reconstruction. *Mol Phylogenet Evol.* 2010;55:1–17.
- Keller A, Förster F, Müller T, Dandekar T, Schultz J, Wolf M. Including RNA secondary structures improves accuracy and robustness in reconstruction of phylogenetic trees. *Biol Direct.* 2010;5:4.
- Xia X, Xie Z, Salemi M, Chen L, Wang Y. An index of substitution saturation and its application. *Mol Phylogenet Evol.* 2003;26:1–7.
- Xia X, Lemey P. Assessing substitution saturation with DAMBE. In: Lemey P, Salemi M, Vandamme A-M, editors. *The Phylogenetic Handbook: A Practical Approach to DNA and Protein Phylogeny.* New York: Cambridge University Press; 2009. p. 615–30.
- Domes K, Althammer M, Norton RA, Scheu S, Maraun M. The phylogenetic relationship between Astigmata and Oribatida (Acari) as indicated by molecular markers. *Exp Appl Acarol.* 2007;42:159–71.
- Letsch HO, Kjer KM. Potential pitfalls of modelling ribosomal RNA data in phylogenetic tree reconstruction: Evidence from case studies in the Metazoa. *BMC Evol Biol.* 2011;11:146.
- Zrzavý J, Hypša V, Vlášková M. Arthropod phylogeny: taxonomic congruence, total evidence and conditional combination approaches to morphological and molecular data sets. In: Fortey RA, Thomas RH, editors. *Arthropod relationships.* London: Chapman & Hall; 1998. p. 97–107.
- Giribet G, Edgecombe GD, Wheeler WC. Arthropod phylogeny based on eight molecular loci and morphology. *Nature.* 2001;413:157–61.
- Rota-Stabelli O, Telford MJ. A multi criterion approach for the selection of optimal outgroups in phylogeny: recovering some support for Mandibulata over Myriochelata using mitogenomics. *Mol Phylogenet Evol.* 2008;48:103–11.
- Regier JC, Zwick A. Sources of signal in 62 protein-coding nuclear genes for higher-level phylogenetics of arthropods. *PLoS One.* 2011;6:e23408.
- Rota-Stabelli O, Campbell L, Brinkmann H, Edgecombe GD, Longhorn SJ, Peterson KJ, et al. A congruent solution to arthropod phylogeny: Phylogenomics, microRNAs and morphology support monophyletic Mandibulata. *Proc R Soc B.* 2011;278:298–306.
- Weygoldt P, Paulus HF. Untersuchungen zur Morphologie, Taxonomie und Phylogenie der Chelicerata. 2. Cladogramme und die Entfaltung der Chelicerata. *J Zoolog Syst Evol Res.* 1979;17:177–200.
- Shultz JW. A phylogenetic analysis of the arachnid orders based on morphological characters. *Zool J Linn Soc.* 2007;150:221–65.
- van der Hammen L. A new classification of the Chelicerata. *Zool Meded.* 1977;51:307–19.
- van der Hammen L. *An Introduction to Comparative Arachnology.* SPB The Hague: Academic Publishing; 1989.
- Lindquist EE. Current theories on the evolution of major groups of Acari and on their relationships with other groups of Arachnida, with consequent implications for their classification. In: Griffiths DA, Bowman CE, editors. *Acarology VI, Volume I.* Chichester: Ellis-Horwood Ltd; 1984. p. 28–62.
- Giribet G, Edgecombe GD. The Arthropoda: A Phylogenetic Framework. In: Minelli A, Boxshall G, Fusco G, editors. *Arthropod Biology and Evolution Molecules, Development, Morphology.* Berlin Heidelberg: Springer; 2013. p. 17–40.
- Mallatt J, Craig CW, Yoder MJ. Nearly complete rRNA genes from 371 Animalia: Updated structure-based alignment and detailed phylogenetic analysis. *Mol Phylogenet Evol.* 2012;64:603–17.
- Von Reumont BM, Meusemann K, Szucsich NU, Dell’Ampio E, Gowri-Shankar V, Bartel D, et al. Can comprehensive background knowledge be incorporated into substitution models to improve phylogenetic analyses? A case study on major arthropod relationships. *BMC Evol Biol.* 2009;9:119.
- Meusemann K, von Reumont BM, Simon S, Roeding F, Strauss S, Kuck P, et al. A phylogenomic approach to resolve the arthropod tree of life. *Mol Biol Evol.* 2010;27:2451–64.
- Regier JC, Shultz JW, Zwick A, Hussey A, Ball B, Wetzler R, et al. Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. *Nature.* 2010;463:1079–82.
- Sharma PP, Giribet G. A revised dated phylogeny of the arachnid order Opiliones. *Front Genet.* 2014;5:255.
- Grimaldi D, Engel MS, Nascimbene PC. Fossiliferous Cretaceous amber from Myanmar (Burma): Its rediscovery, biotic diversity, and paleontological significance. *Am Mus Novit.* 2002;3361:1–72.
- Poinar Jr GO, Brown AE. A new genus of hard ticks in Cretaceous Burmese amber (Acari: Ixodida: Ixodidae). *Syst Parasitol.* 2003;54(3):199–205.
- Poinar GO, Buckley R. *Compluriscutula vetulum* (Acari: Ixodida: Ixodidae), A New Genus and Species of Hard Tick from Lower Cretaceous Burmese Amber. *Proc Entomol Soc Wash.* 2008;110(2):445–50.
- Dunlop JA, Bernardi LFO. An opilioacarid mite in Cretaceous Burmese amber. *Naturwissenschaften.* 2014;101(9):759–63.
- Hirst S. On some arachnid remains from the Old Red Sandstone (Rhynie Chert bed, Aberdeenshire). *Ann Mag Nat Hist.* 1923;12(70):455–74.
- Dubinin VB. Class Acaromorpha: mites or gnathosomic chelicerate arthropods. In: Rodendorf BB, editor. *Fundamentals of palaeontology.* Moscow: Academy of Sciences of the USSR; 1962. p. 447–73 [In Russian].
- Dunlop JA, Selden PA. Calibrating the chelicerate clock: a paleontological reply to Jeyaprakash and Hoy. *Exp Appl Acarol.* 2009;48:183–97.

46. Dunlop JA, Penney D, Jekel D. A summary list of fossil spiders and their relatives. In: World Spider Catalog (2014). Bern: Natural History Museum; 2014. <http://www.wsc.nmbe.ch/resources/fossils/Fossils15.pdf>. Accessed in 12 June 2015.
47. Kjer KM, Roshan U, Gillespie J. Structural and evolutionary considerations for multiple sequence alignment of RNA, and challenges for algorithms that ignore them. In: Rosemberg MS, editor. Sequence alignment. Methods, Models, Concepts, and Strategies. London: University of California Press; 2009. p. 1050–149.
48. Liu K, Nelesen S, Raghavan S, Linder CR, Warnow T. Barking Up the wrong treelength: the impact of Gap penalty on alignment and tree accuracy. *IEEE/ACM Trans Comput Biol Bioinf*. 2009;6(1):7–21.
49. Misof B, Misof KA. Monte Carlo approach successfully identifies randomness in multiple sequence alignments: a more objective means of data exclusion. *Syst Biol*. 2009;58(1):21–34.
50. Schultz J, Wolf M. ITS2 sequence–structure analysis in phylogenetics: a how-to manual for molecular systematics. *Mol Phylogenet Evol*. 2009;52:520–3.
51. Wolf M, Koetschan C, Müller T. ITS2, 18S, 16S or any other RNA - simply aligning sequences and their individual secondary structures simultaneously by an automatic approach. *Gene*. 2014;546(2):145–9.
52. Wiemers M, Keller A, Wolf M. ITS2 secondary structure improves phylogeny estimation in a radiation of blue butterflies of the subgenus *Agrodiaetus* (Lepidoptera: Lycaenidae: *Polyommatus*). *BMC Evol Biol*. 2009;9:300.
53. Ahvenniemi P, Wolf M, Lehtonen MJ, Wilson P, German-Kinnari M, Valkonen JPT. Evolutionary diversification indicated by compensatory base changes in ITS2 secondary structures in a complex fungal species, *Rhizoctonia solani*. *J Mol Evol*. 2009;69(2):150–63.
54. Buchheim MA, Sutherland DM, Buchheim JA, Wolf M. The blood alga: phylogeny of *Haematococcus* (Chlorophyceae) inferred from ribosomal RNA gene sequence data. *Eur J Phycol*. 2013;48(3):318–29.
55. Salvi D, Macali A, Mariottini P. Molecular phylogenetics and systematics of the bivalve family Ostreidae based on rRNA sequence–structure models and multilocus species tree. *PLoS One*. 2014;9(9), e108696.
56. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucl Acids Res*. 2013;41(D1):D590–6.
57. Suchard MA, Redelings BD. *Bali-Phy*: Simultaneous Bayesian inference of alignment and phylogeny. *Bioinformatics*. 2006;22:2047–8.
58. McKenzie SK, Oxley PR, Kronauer DJC. Comparative genomics and transcriptomics in ants provide new insights into the evolution and function of odorant binding and chemosensory proteins. *BMC Genomics*. 2014;15:718.
59. Reuter E. Zur Morphologie und Ontogenie der acariden mit besonderer Berücksichtigung von *Pediculopsis graminum* (E. Reut.). *Acta Soc Scient Fennicae*. 1909;36(4):1–287.
60. Bolton SJ, Klompen H, Bauchan GR, Ochoa R. A new genus and species of Nematolycidae (Acari: Endeostigmata). *J Nat Hist*. 2014;48(23–24):1359–73.
61. Bolton SJ, Bauchan GR, Ochoa R, Pooley C, Klompen H. The role of the integument with respect to different modes of locomotion in the Nematolycidae (Endeostigmata). *Exp Appl Acarol*. 2015;65:149–61.
62. Alberti G, Storch V, Renner H. Über den feinstrukturellen Aufbau der Milbencuticula (Acari, Arachnida). *Zool Jahrb Abt Anat Ontog Tiere*. 1981;105:183–236.
63. Uusitalo M. Revision of the family Alycidae (Acariformes, Acari), with special reference to European species, PhD thesis. Helsinki: Helsinki University, Department of Biological and Environmental Sciences; 2010.
64. Norton RA. Morphological evidence for the evolutionary origin of Astigmata (Acari: Acariformes). *Exp Appl Acarol*. 1998;22:559–94.
65. Norton RA, O'Connor BM, Johnston DE. Systematic relationships of the Pediculochelidae (Acari: Acariformes). *Proc Entomol Soc Wash*. 1983;85:493–512.
66. Grandjean F. *Retetysdes* et les Stigmatés Mandibulaires des Acariens Prostigmatiques. *Bull Mus Nat His Nat*. 1938;10:279–86.
67. Walter DE, Lindquist EE, Smith IM, Cook DR, Krantz GW. Order Trombidiformes. In: Krantz GW, Walter DE, editors. *A Manual of Acarology*. 3rd ed. Lubbock: Texas Tech University Press; 2009. p. 233–420.
68. Lindquist EE. Transfer of the Tarsoscheilyidae to the Heterostigmata, and reassignment of the Tarsonemina and Heterostigmata to lower hierarchical status in the Prostigmata (Acari). *Can Entomol*. 1976;108:23–48.
69. Witte H. The phylogenetic relationships within the Parasitengonae. In: Dusbábek F, Bukva V, editors. *Modern Acarology*, vol. 2. The Hague: SPB Academic Publishing bv; Prague: Academia; 1991. p. 171–82.
70. Alberti G, Coons LB. Acari - Mites. In: Harrison FW, editor. *Microscopic Anatomy of Invertebrates*, vol. 8c. New York: John Wiley & Sons, Inc; 1999. p. 515–1265.
71. Otto J. A cladistic analysis of Erythracarinae (Acarina: Prostigmata: Anystidae), with the description of a new genus. *Syst Entomol*. 2000;25:447–84.
72. van der Hammen L. Acarological and Arachnological notes. *Zool Meded*. 1986;60(14):217–30.
73. Grandjean F. Sur quelques caractères des Acaridae libres. *Bull Soc Zool Fr*. 1937;62:388–98.
74. Grandjean F. Essai de classification des Oribates (Acarieus). *Bull Soc Zool Fr*. 1953;78:421–46.
75. Grandjean F. Le genre *Pachygnathus* Dugès (*Alycus* Koch) (Acarieus). Cinquième et dernière partie. *Bull Mus nat hist*. 1937;9(ser. 2):262–9.
76. Giribet G, Carranza S, Bagnà J, Riutort M, Ribera C. First molecular evidence for the existence of a Tardigrada + Arthropoda clade. *Mol Biol Evol*. 1996;13:76–84.
77. Otto J, Wilson K. Assessment of the usefulness of ribosomal 18 S and mitochondrial COI sequences in Prostigmata phylogeny. In: Halliday RB, Walter DE, Proctor HC, Norton RA, Colloff J, editors. *Acarology: Proceedings of the 10th International Congress*. Melbourne: CSIRO Publishing; 2001. p. 100–9.
78. Mallatt J, Sullivan J. 28S and 18S rDNA sequences support the monophyly of lampreys and hagfishes. *Mol Biol Evol*. 1998;15:1706–18.
79. Arango CP, Wheeler WC. Phylogeny of the sea spiders (Arthropoda, Pycnogonida) based on direct optimization of six loci and morphology. *Cladistics*. 2007;23:255–93.
80. Whiting MF, Carpenter JM, Wheeler QD, Wheeler WC. The strepsiptera problem: phylogeny of the holometabolous insect orders inferred from 18 S and 28 S ribosomal DNA sequences. *Syst Biol*. 1997;46:1–68.
81. Giribet G, Vogt L, González AP, Sharma P, Kury A. A multilocus approach to harvestman (Arachnida: Opiliones) phylogeny with emphasis on biogeography and the systematics of Laniatores. *Cladistics*. 2009;25:1–30.
82. Schwendinger PJ, Giribet G. The systematics of the south-east Asian genus *Fangensis* Rambla (Opiliones: Cyphophthalmi: Stylocellidae). *Invertebr Syst*. 2005;19:297–323.
83. Knowles L, Klimov PB. Estimating phylogenetic relationships despite discordant gene trees across loci: the species tree of a diverse species group of feather mites (Acari: Proctophylloidae). *Parasitology*. 2011;138(13):1750–9.
84. Adoutte A, Balavoine G, Lartillot N, Lespinet O, Prud'homme B, de Rosa R. The new animal phylogeny: reliability and implications. *Proc Natl Acad Sci U S A*. 2000;97:4453–6.
85. Halanych KM. The new view of animal phylogeny. *Annu Rev Ecol Evol Syst*. 2004;35:229–56.
86. Wheeler WC, Aagesen L, Arango CP, Faivovich J, Grant T, D'Haese D, et al. Dynamic homology and phylogenetic systematics: a unified approach using POY. 1. New York: American Museum of Natural History; 2006.
87. Varón A, Vinh LS, Wheeler WC. POY version 4: phylogenetic analysis using dynamic homologies. *Cladistics*. 2010;26:72–85.
88. Yoshizawa K. Direct optimization overly optimizes data. *Syst Entomol*. 2010;35:199–206.
89. Stocsits RR, Letsch H, Hertel J, Misof B, Stadler PF. Accurate and efficient reconstruction of deep phylogenies from structured RNAs. *Nucl Acids Res*. 2009;37:6184–93.
90. Zhao Y-E, Wang Z-H, Xu Y, Wu L-P, Hu L. Secondary structure prediction for complete rDNA sequences (18S, 5.8S, and 28S rDNA) of *Demodex folliculorum*, and comparison of divergent domains structures across Acari. *Exp Parasitol*. 2013;135(2):370–81.
91. Cannone JJ, Subramanian S, Schnare MN, Collett JR, D'Souza LM, Du Y, et al. The Comparative RNA Web (CRW) Site: An online database of comparative sequence and structure information for ribosomal, intron and other RNAs. *BMC Bioinf*. 2002;3:15.
92. Wuyts J, Perrière G, van der Peer Y. The European ribosomal RNA database. *Nucleic Acids Res*. 2004;32:D101–3.
93. Gillespie JJ, Yoder MJ, Wharton RA. Predicted secondary structures for 28 S and 18 S rRNA from Ichneumonidea (Insecta: Hymenoptera: Apocrita): Impact on sequence alignment and phylogeny estimation. *J Mol Evol*. 2005;61(1):114–37.
94. Kjer KM. Use of rRNA secondary structure in phylogenetic studies to identify homologous positions: An example of alignment and data presentation from the frogs. *Mol Phylogenet Evol*. 1995;4:314–30.

95. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser.* 1999;41:95–8.
96. Gillespie JJ, Johnston JS, Cannone JJ, Gutell RR. Characteristics of the nuclear (18S, 5.8S, 28S and 5S) and mitochondrial (12S and 16S) rRNA genes of *Apis mellifera* (Insecta: Hymenoptera): structure, organization, and retrotransposable elements. *Insect Mol Biol.* 2006;15(5):657–86.
97. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucl Acids Res.* 2003;31:3406–15.
98. Hofacker IL, Fekete M, Stadler PF. Secondary structure prediction for aligned RNA Sequences. *J Mol Biol.* 2002;319:1059–66.
99. Gillespie JJ. Characterizing regions of ambiguous alignment caused by the expansion and contraction of hairpin-stem loops in ribosomal RNA molecules. *Mol Phylogenet Evol.* 2004;33:936–43.
100. Wheeler WC. Sequence alignment, parameter sensitivity, and the phylogenetic analysis of molecular data. *Syst Biol.* 1995;44:321–31.
101. Kück P, Meusemann K, Dambach J, Thormann B, von Reumont BM, Wägele JW, et al. Parametric and non-parametric masking of randomness in sequence alignments can be improved and leads to better resolved trees. *Front Zool.* 2010;7:10.
102. Kück P, Meusemann K. FASconCAT: convenient handling of data matrices. *Mol Phylogenet Evol.* 2012;56:1115–8.
103. Posada D. JModelTest: phylogenetic model averaging. *Mol Biol Evol.* 2008;25:1253–6.
104. Ronquist F, Teslenko M, van der Mark P, Ayres D, Darling A, Höhna S, et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol.* 2012;61(3):539–42.
105. Rambaut A, Suchard MA, Xie D, Drummond AJ. Tracer v1.6, 2014, <http://beast.bio.ed.ac.uk/Tracer>. Accessed 12 June 2015.
106. Marshall D. Cryptic failure of partitioned Bayesian phylogenetic analyses: lost in the land of long trees. *Syst Biol.* 2010;59:108–17.
107. Brown JM, Hedtke SM, Lemmon AR, Lemmon EM. When trees grow Too long: investigating the causes of highly inaccurate Bayesian branch-length estimates. *Syst Biol.* 2010;59(2):145–61.
108. Baele G, Lemey P, Bedford T, Rambaut A, Suchard MA, Alekseyenko AV. Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Mol Biol Evol.* 2012;29(9):2157–67.
109. Jow H, Hudelot C, Rattay M, Higgs PG. Bayesian phylogenetics using an RNA substitution model applied to early mammalian evolution. *Mol Biol Evol.* 2002;19:1591–601.
110. Stamatakis A. RAxML Version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*; 2014. doi:10.1093/bioinformatics/btu033. Accessed 12 June 2015.
111. Miller MA, Pfeiffer W, Schwartz T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *Proc Gatew Comput Environ Work.* 2010;1:1–8.
112. Xia X. DAMBES: A Comprehensive Software Package for Data Analysis in Molecular Biology and Evolution. *Mol Biol Evol.* 2013;30(7):1720–8.
113. Boogaart KG, van den Tolosana R, Bren M. Package 'compositions' v. 1.30-1. 2013. <http://www.stat.boogaart.de/compositions>. Accessed 12 June 2015.
114. Schmidt HA, Strimmer K, Vingron M, von Haeseler A. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics.* 2002;18:502–4.
115. Wilcox TP, de Leon FJ G, Hendrickson DA, Hillis DM. Convergence among cave catfishes: long-branch attraction and a Bayesian relative rates test. *Mol Phylogenet Evol.* 2004;31:1101–3.
116. Nieselt-Struwe K, von Haeseler A. Quartet-mapping, a generalization of the likelihood-mapping procedure. *Mol Biol Evol.* 2001;18:1204–19.
117. Holland B, Moulton V. Consensus networks: a method for visualizing incompatibilities in collections of trees. In: Benson G, Page R, editors. *Proceedings of "Workshop on Algorithms in Bioinformatics"*. Berlin: Springer; 2003. p. 2812. 165–76.
118. Huson DH, Bryant D. Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol.* 2006;23(2):254–67.
119. Shimodaira H. An approximately unbiased test of phylogenetic tree selection. *Syst Biol.* 2002;51:492–508.
120. Salichos L, Rokas A. Inferring ancient divergences requires genes with strong phylogenetic signals. *Nature.* 2013;497:327–31.
121. Salichos L, Stamatakis A, Rokas A. Novel information theory-based measures for quantifying incongruence among phylogenetic trees. *Mol Biol Evol.* 2014;31:1261–71.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

