Available online at www.sciencedirect.com
ScienceDirect
Molecular Phylogenetics and Evolution 45 (2007) 251–260
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

The complete mitochondrial genome of *Scutigerella causeyae* (Myriapoda: Symphyla) and the phylogenetic position of Symphyla

Lars Podsiadlowski *, Heiko Kohlhagen, Markus Koch

AG Systematik und Evolution der Tiere, Institut für Biologie, Freie Universität Berlin, Germany

Received 4 December 2006; revised 10 July 2007; accepted 16 July 2007
Available online 31 July 2007

Abstract

The first complete mitochondrial genome of a representative of the Symphyla, *Scutigerella causeyae* (Arthropoda: Myriapoda), was sequenced using a PCR-based approach. Its gene order shows different positions for three tRNA genes compared to the ancestral arthropod pattern. Presence of a pseudogene with partial sequence similarity to *rrnS* favours the duplication–random loss model as an explanation for at least one of the translocations. None of the genome rearrangements hypothesized for *S. causeyae* are shared by any of the other four myriapod mitochondrial genomes sequenced so far (two from Chilopoda and two from Diplopoda). Different rearrangement events must have occurred independently in the lineages leading to *S. causeyae*, *Lithobius forficatus*, *Scutigera coleoptrata* and Diplopoda. Phylogenetic analyses could not unequivocally elucidate the position of Symphyla among myriapods. While the nucleotide dataset of eleven protein-coding genes gives weak support for an affinity to Chilopoda, this is not recovered with the corresponding amino acid dataset.

Keywords: Symphyla; Mitochondrial genome; Myriapoda; Gene rearrangement

1. Introduction

Mitochondrial genomes have been characterized from a growing number of metazoan species. Due to constant gene content (13 protein-coding, 2 rRNA and 22 tRNA genes), maternal inheritance, lack of recombination and lack of paralogous genes, mtDNA seemed to be perfectly suited for phylogenetic analyses (Avise, 2000; Moritz et al., 1987). However, recent evidence for nuclear pseudogenes and genome duplications, as well as high substitution rates, base heterogeneity and remarkable differences between phylogenetic trees derived from mitochondrial sequences and other data questioned the importance of mtDNA for phylogenetic studies (Ballard and Whitlock, 2004; Shaw, 2002). These insights force careful execution of mtDNA sequence based analyses and render it essential that analytical results are interpreted with caution (Rubinoff and Holland, 2005).

Besides sequence based analyses, other characters derived from the mitochondrial genome were used in a phylogenetic context, predominantly the mitochondrial gene order. The most prominent example is the gene translocation of *trnL*(UUR) to a position between *cox1* and *cox2*, which proved to be a common feature of Hexapoda and Crustacea, while Myriapoda, Chelicerata and Onychophora have retained the ancestral state (Boore et al., 1998, 1995). Yet, not much is known about the mechanisms underlying gene translocations. The duplication–random loss model (Moritz et al., 1987) and the duplication–non-random loss model (Lavrov et al., 2002) may account for larger rearrangements, but translocations of single tRNAs over large distance can hardly be explained by these models and may rather stem from single translocation events of small genome fragments.

* Corresponding author. Fax: +49 30 83854869.
E-mail address: lars@podsiadlowski.de (L. Podsiadlowski).

1055-7903/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2007.07.017

Symphylans are small, soil-dwelling, eyeless myriapods. Adults have a trunk consisting of 14 segments, the first 12 each carrying a pair of legs. The number of tergal plates (15–22) does not follow the segmentation of the trunk. The second maxillae are basally fused and form a complex labium. As in Diplopoda and Pauropoda the gonopores are found in the anterior body, a morphological character in favour of the Progoneata hypothesis (= a unit consisting of Symphyla, Pauropoda, Diplopoda). The Symphyla comprise about 160 described species traditionally classified into the subtaxa Scutigerellidae and Scolopendrellidae (Scheller and Adis, 2002). For our study we chose the scutigerellid species *Scutigerella causeyae* Michelbacher, 1942, a frequent soil-dweller of the Grunewald in Berlin, Germany (Haupt, 1977).

The phylogenetic position of symphylans among myriapods has not yet been determined unequivocally by molecular data: some tend to confirm the morphological view that the Symphyla form the sister group of the Dignatha (= Pauropoda + Diplopoda) within a monophyletic unit Progoneata (Edgecombe, 2004; Giribet et al., 2005), another analysis favours a closer relationship between Symphyla, Pauropoda and Chilopoda (Gai et al., 2006), while in some publications position of Symphyla could not clearly be determined (Mallatt and Giribet, 2006; Regier et al., 2005). Fast evolving rRNA genes in symphylans seem to hamper cladistic analyses. Available data on myriapod mitochondrial genomes question the monophyly of Myriapoda (Negrisolo et al., 2004b) and corroborate the current view that myriapods are closer related to Chelicerata than to Pancrustacea (Hwang et al., 2001; Mallatt et al., 2004; Pisani et al., 2004). Neuroanatomical data even indicate a polyphyletic origin of myriapods (e.g., Strausfeld et al., 2006). In none of these studies symphylans were included in the analysis. However, a sister group relationship between insects (Hexapoda) and symphylans (Ventrovesiculata-hypothesis; Willmann, 2003) or between Hexapoda and Progoneata (Labiophora-hypothesis; e.g., Kraus, 2001) was never retrieved by molecular data.

In this publication we present the first complete mitochondrial genome sequence from a symphylan species. Gene order in Symphyla and in other myriapod taxa is compared. We discuss the possible steps that led to the derived gene order of *S. causeyae* and try to infer the position of Symphyla among myriapods by a sequence based phylogenetic analysis.

2. Materials and methods

2.1. DNA extraction, PCR and sequencing

Specimens of *S. causeyae* were sampled at the Grunewald in Berlin, Germany. DNA extraction from one individual was done using the DNeasy tissue kit (Qiagen). We amplified two large fragments of the mitochondrial genome using the long-range PCR primers described by Yamauchi et al. (2004): L39-Met/H5244-CO3 (~5 kB) and L5281/H12230-16S (~7 kB). Takara LA *Taq* (Takara) was used under the following PCR conditions: 25 μl reaction volumes (13.75 μl sterile distilled water; 2.5 μl 10× LA PCR buffer; 2.5 μl $MgCl_2$, 25 mM; 4 μl dNTP, 2.5 mM; 1 μl primer mix, 10 μM each; 1 μl DNA; 0.25 μl Takara LA *Taq*) were set up and an initial denaturation step (94 °C, 1 min) was followed by 35 cycles of denaturation (98 °C, 10 s), annealing (60 °C, 1 min) and extension steps (68 °C for 12 min), and a final elongation step (72 °C, 4 min).

All PCRs were performed with Mastercycler or Mastercycler Gradient thermocyclers (Eppendorf). PCR products were inspected on a 1% agarose gel. Long PCR products were subsequently used as a template for secondary PCR of shorter fragments using primer pairs S1 to S42 from Yamauchi et al. (2004). In PCR of short fragments 25 μl reaction volumes were used (18.3 μl distilled sterile water; 2.5 μl 10× Eppendorf PCR buffer; 1.0 μl dNTP, 10 mM; 1 μl of each primer, 10 μM; 1 μl DNA; 1:50 dilution of primary PCR products; 0.2 μl Eppendorf Hotmaster *Taq*); temperature profile: initial denaturation (94 °C, 2 min), 30 cycles of denaturation (94 °C, 30 s), annealing (45 °C, 30 s) and elongation (68 °C, 90 s) were followed by a final elongation step (68 °C, 2 min). PCR was successful with the following primer pairs: S2, S3, S8–S15, S18, S23, S24, S28, S29, S32, S34–S37, S39–S41. Gaps between these fragments were closed with species specific primer pairs designed on the basis of the primary sequence information from these PCR products. Primer pairs and corresponding annealing temperatures used in this step are listed in Table 1.

PCR cleanup of short PCR products was done using the BlueMatrix PCR/DNA clean-up kit (EURx, Gdansk, Poland); in cases where gel purification was necessary the GeneMatrix agarose-out kit was used (EURx, Gdansk, Poland). All sequencing was done on a CEQ 8000 capillary sequencer (Beckman-Coulter). Sequencing was performed using the DCTS quickstart kit (Beckman-Coulter) with 10 μl reaction volumes (4 μl sterile distilled water; 1 μl sequencing primer, 10 mM; 1 μl of purified PCR product; 4 μl DCTS master mix). The complete mitochondrial genome sequence was covered at least by two sequencing runs. If these differed from each other at least two more sequencing runs were performed for this region.

2.2. Sequence annotation and phylogenetic analysis

Sequences were primarily edited by CEQ software (Beckman-Coulter), and assemblage of the complete genome sequence was done using Bioedit 7.01 (Hall, 1999). Transfer RNA genes were identified using tRNA-scan SE 1.21 (Lowe and Eddy, 1997) or by visual inspection of otherwise non-coding regions looking for stem–loop structures similar to anticodon stem–loops. The mitochondrial genome sequence of *S. causeyae* was deposited at GenBank, Accession No. DQ666065. Phylogenetic analyses were performed with concatenated alignments

Table 1
Species specific PCR primer pairs used for amplification of fragments from the mitochondrial genome of *Scutigerella causeyae*

Primer name	Nucleotide sequence (5′–3′)	Annealing temperature (°C)
Scu-2-9	ATTTCTATATAACTTACAAAGGGC	48
Scu-9-2	AAAGCTATATCTGGAGCACC	48
Scu-15-18	GACTTTCTTTACATTGTTTTTGG	50
Scu-18-15	ACAAAAAATAGTACTTCTGAGGC	50
Scu-18-23	TTTGCCATTTCTGACTCATC	50
Scu-23-18	ATAATTTATTTTTGGCTTGTTAG	50
Scu-24-28	AATTCTTATTAGATTTTTTCC	45
Scu-28-24	TTTCAACTTTAGTAACTGCTGG	45
Scu-28-36	GTGCCCTTTTAGTTATAGCC	48
Scu-36-28	GGGAATTTATAGTTTTATTAGTTG	48
Scu-37-40	AAGACTTATTTGGCTATATTTTAG	48
Scu-40-37	TTGAATATTCGGGGTCG	48
Scu-2r-40	AAAGAATAGGCTGGAGATAATAGTTGG	55
Scu-40f-2	CCCCTCCGCAAAATCAAAAC	55

of eleven protein-coding genes (*atp8* and *nad4l* were excluded due to their shortness and strong heterogeneity even between close related species). Gene sequences from 12 arthropod taxa (see below) were aligned making use of ClustalW (Thompson et al., 1994) as implemented in Bioedit (default conditions: full multiple alignment making use of a neighbour joining bootstrap tree; gap opening and extension parameters optimized by the software according to the dataset). Nucleotide sequences were converted to amino acid sequences prior to the alignment and converted back afterwards. Amino acid alignments were controlled by eye inspections and modified in some cases to optimize the alignment. Ambigously aligned proportions were excluded from the final alignments using Gblocks 0.91b (Castresana, 2000) under the codons option. The default parameters settings were used: minimum number of sequences for a conserved/flanking position (7/10), maximum number of contiguous non-conserved positions (8), minimum length of a block (10), allowed gap positions (none). This procedure recovered 9633 nucleotides or 89% of the original alignment. A saturation analysis (Xia et al., 2003) was performed for subsets with first, second and third codon positions using DAMBE 4.2.13 (Xia and Xie, 2001). According to the results, third codon positions were omitted from the final nucleotide alignment which consisted of 6422 nucleotides. For amino acid alignments Gblocks was used with the same parameter set, resulting in an alignment of 2862 amino acids, or 79% of the original dataset.

For the nucleotide dataset Modeltest ver. 3.7 (Posada and Crandall, 1998) was performed: using the Akaike information criterion GTR + I + Γ was chosen as model for the likelihood and Bayesian analyses. Model selection for the amino acid dataset was done with Prottest ver. 1.3 (Abascal et al., 2005). Due to the Akaike information criterion model MtArt + Γ + I performed best with our dataset. As the MtArt model is a very recent addition to the models commonly used (Abascal et al., 2007), we could not implement it in Bayesian analysis, where we used the best scoring alternative, MtRev + Γ + I. For a likelihood analysis, we implemented the MtArt matrix in RAxML (Stamatakis, 2006).

Maximum parsimony analyses with both datasets (nucleotides and amino acids) were performed using PAUP* 4.0b10 (Swofford, 2002). One thousand of bootstrap replicates were analysed, each with 10 replicates with random taxon order.

Maximum likelihood analysis of the nucleotide alignment was done using PHYML, ver. 2.4.4 (Guindon and Gascuel, 2003). The model setting was GTR + Γ + I, with four rate categories; base frequencies, the proportion of invariable sites and alpha shape parameter were empirically estimated from the dataset. Maximum likelihood analysis of the amino acid alignment was performed using RAxML, ver. 2.2.3 (Stamatakis, 2006). We implemented the MtArt matrix (Abascal et al., 2007) and performed an analysis with 100 bootstrap replicates, with the options "–f d" for the rapid hill-climbing algorithm and the model MtArt + Γ using empirical base frequencies (RaxML has no option for a model with invariant sites).

Bayesian analyses of nucleotide and amino acid datasets were performed with MrBayes 3.1.2 (Huelsenbeck and Ronquist, 2001), using the GTR + Γ + I and MtRev + Γ + I model, respectively. Eight chains ran in parallel for 1,000,000 generations, sampling trees every 1000 generations. Bayesian analyses were repeated two times, always retrieving the same topology. According to the likelihood plots, the first 100 trees were discarded as burn-in (both datasets), while Bayesian posterior probabilities were calculated according to the remaining set of trees.

Full species names and accession numbers of the other mitochondrial genome sequences used in the analysis were as follows: *Epiperipatus biolleyi* DQ666064; *Narceus annularus*, NC_003343; *Thyropygus* sp. DVL-2001, NC_003344; *Scutigera coleoptrata*, NC_005870; *Lithobius forficatus*, *NC_002629*; *Limulus polyphemus*, NC_003057; *Japyx solifugus*, NC_007214; *Petrobius brevistylis*, NC_007689; *Daphnia pulex*, NC_000844; *Penaeus monodon*, NC_002184; *Tetraclita japonica*, NC_008974.

3. Results

3.1. Genome organization, gene order and non-coding parts

The complete mitochondrial genome of *S. causeyae* is a circular double helix, 14,637 bp long, containing all 37 genes typically found in triploblastic animals (Table 2). As usual in animal mtDNA, there are only small non-coding regions between the genes, with a few exceptions (see below). In some cases genes overlap, predominantly when both genes are encoded on different strands (*trnM–trnQ*, *trnW–trnC*, *trnE–trnF*; Table 2).

In most animal mitochondrial genomes there is one large AT-rich non-coding region, referred to as the mitochondrial control region (CR). It bears regulatory elements important for replication and transcription (Wolstenholme, 1992). As in most other arthropods, the mitochondrial control region of *S. causeyae* is located between *rrnS* and *trnI* (Table 2). It consists of 260 bp and has an AT-content of 73.5%, which is only slightly higher than that for the complete genome (72.6%) and far below that of the rRNA genes (*rrnL*: 78.0%; *rrnS*: 77.0%). Another comparably large non-coding region (NCR, 185 bp) is found between *trnV* and *trnM*. This sequence shows high sequence similarity (95%) to one third from the 5′-end of *rrnS*. Smaller non-coding regions are present between *trnQ* and *nad2* (42 bp) and between *trnS* and *nad1* (17 bp). Sometimes non-coding regions are interpreted as remnants of functionless gene duplicates, but in this case BLAST search gave no hint of sequence homology to other genes and no suspicious secondary structures (hairpins etc., as remnants of tRNA genes) were found.

The mitochondrial genome of *S. causeyae* shows three differences in gene order compared to the inferred ground pattern of Euarthropoda (Staton et al., 1997): *trnP* and *trnT* have interchanged their positions, as well as *trnM*

Table 2
Mitochondrial genome organisation of *Scutigerella causeyae*

Gene	Strand	Position	Length (nuc.)	CG-skew	Start-codon	Stop-codon	Intergenic nucleotides
cox1	+	1–1537	1537	0.127	ATT	T–	0
cox2	+	1538–2210	673	0.167	ATG	T–	0
trnK	+	2211–2273	63				−5
trnD	+	2269–2331	63				+10
atp8	+	2342–2500	159	0.579	ATT	TAA	−7
atp6	+	2494–3165	672	0.278	ATG	TAA	−1
cox3	+	3165–3950	786	0.146	ATG	TAA	+1
trnG	+	3952–4005	54				0
nad3	+	4006–4357	352	0.372	ATT	T–	0
trnA	+	4358–4418	61				+1
trnR	+	4420–4477	58				0
trnN	+	4478–4533	56				−3
trnS-AGN	+	4531–4581	51				0
trnE	+	4582–4641	60				−13
trnF	−	4629–4699	71				−6
nad5	−	4694–6363	1676	−0.439	ATA	TA–	−10
trnH	−	6354–6423	70				0
nad4	−	6424–7746	1323	−0.378	ATG	TAA	−7
nad4L	−	7740–8021	282	−0.439	ATG	TAA	+1
trnP	−	8023–8089	67				+3
trnT	+	8087–8150	64				+10
nad6	+	8161–8613	453	0.411	ATT	TAA	−4
cob	+	8610–9744	1135	0.233	ATA	T–	0
trnS-UCN	+	9745–9808	64				+17
nad1	−	9826–10725	900	−0.380	ATT	TAA	+7
trnL-UUR	−	10733–10798	66				−3
trnL-CUN	−	10796–10857	62				—
rrnL	−	10858–12023	1166	−0.390			—
rrnS	−	12024–12748	725	−0.341			—
(CR)		12749–13008	260				—
trnI	+	13009–13074	66				0
trnV	−	13075–13138	64				—
(NCR)	(−)	13139–13323	185				—
trnM	+	13324–13390	67				−12
trnQ	−	13379–13440	62				42
nad2	+	13483–14472	990	0.406	ATT	TAA	+7
trnW	+	14470–14540	61				−17
trnC	−	14524–14586	63				−1
trnY	−	14586–8	60				−8

Complete circular mtDNA has a length of 14637 bp. Non-coding regions in brackets (CR: control region; NCR: Pseudogene resembling partial *rrnS*).

and *trnQ* (Fig. 1). In addition *trnV* is found in a new position between *trnI* and *trnM*. None of these changes is found in any of the other four myriapod mitochondrial genomes sequenced so far (Fig. 1).

3.2. Protein-coding genes and ribosomal RNAs

Protein-coding genes show three different start codons, ATG (5×), ATT (6×) and ATA (2×). In some cases stop codons are truncated (T, TA), as it is often found in other animal mitochondrial genomes. Truncated stop codons probably are completed by polyadenylation after cleavage of the polycistronic transcript (Ojala et al., 1981). Nucleotide composition differs between genes on different strands. Similar to most other arthropods (Hassanin et al., 2005; Hassanin, 2006) *S. causeyae* has positive CG skews in protein-coding and rRNA genes encoded on (+)-strand. Genes encoded on (−)-strand have negative CG skews (Table 2).

3.3. Secondary structure of transfer RNAs

Three of the 22 tRNAs could not be folded into typical cloverleaf secondary structures (Fig. 2). tRNA-Asn and tRNA-Ser(AGN) lack the DHU-arm, while tRNA-Gly lacks the TΨC-arm. While a truncated tRNA-Ser(AGN) is found in many other animal mitochondrial genomes, the other two derivations from the cloverleaf structure are rare. Among myriapods, *L. forficatus* also has a tRNA-Asn lacking the DHU-arm (Lavrov et al., 2000), while the other three myriapod species show a "normal" tRNA-Asn. In *S. causeyae* tRNA-Ile, tRNA-Cys, tRNA-Leu(CUN), tRNA-Leu(UUR), tRNA-Met and tRNA-Thr have one to three mismatches in the acceptor stem. RNA editing of mitochondrial tRNAs is reported in some cases, e.g., from the spider *Habronattus oregonensis* (Masta and Boore, 2004) and from the myriapod *L. forficatus* (Lavrov et al., 2000). Thus, it is possible that the sequence

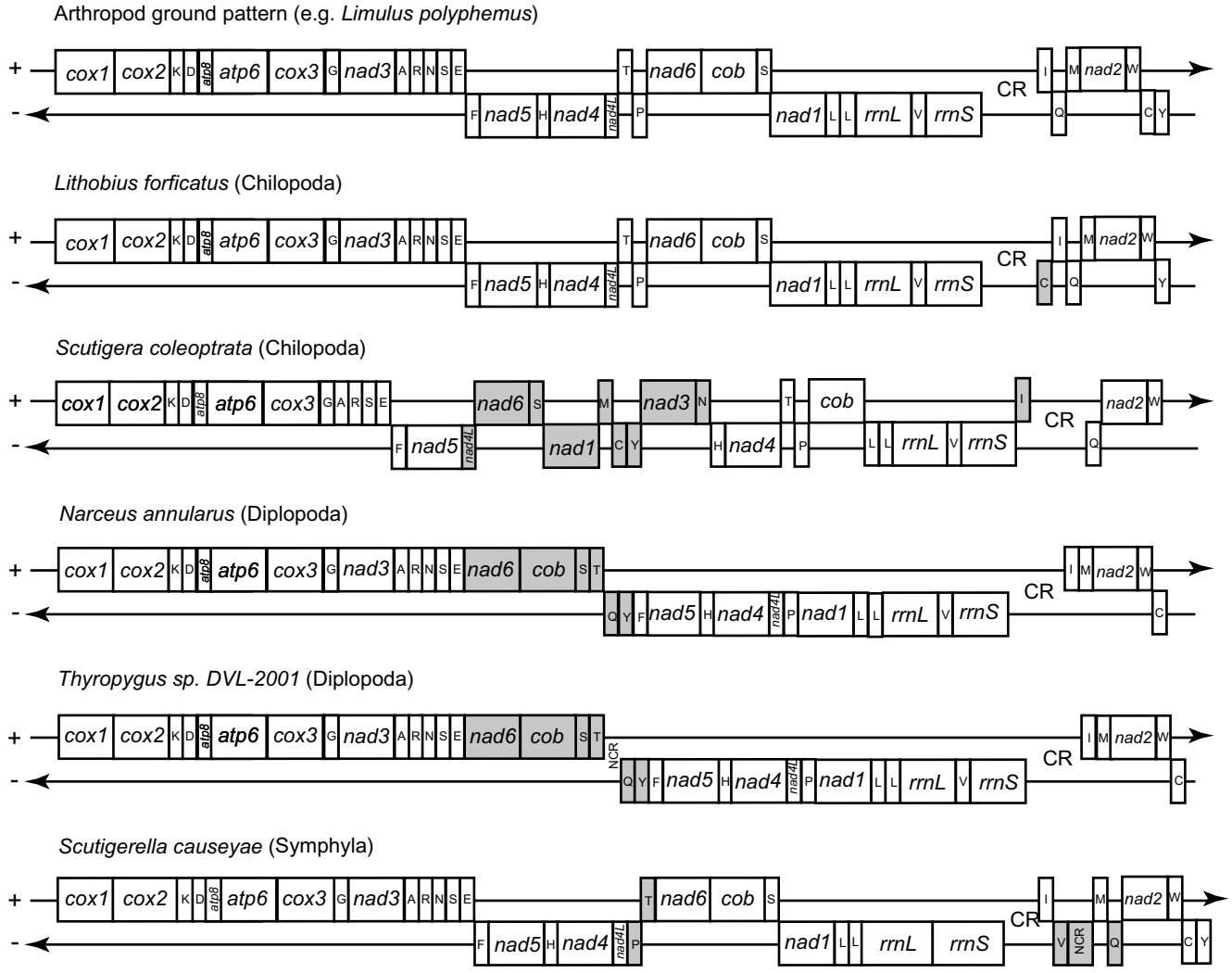

Fig. 1. Mitochondrial gene arrangements of Myriapoda compared to the arthropod ground pattern. *Narceus annularus*, *Thyropygus* sp. according to Lavrov et al. (2002); *S. coleoptrata* according to Negrisolo et al. (2004a); *L. forficatus* according to Lavrov et al. (2000). Genes shaded grey have different relative positions compared to the ground pattern. Upper line with (+)-strand genes, lower line with (−)-strand genes. CR: putative control region.

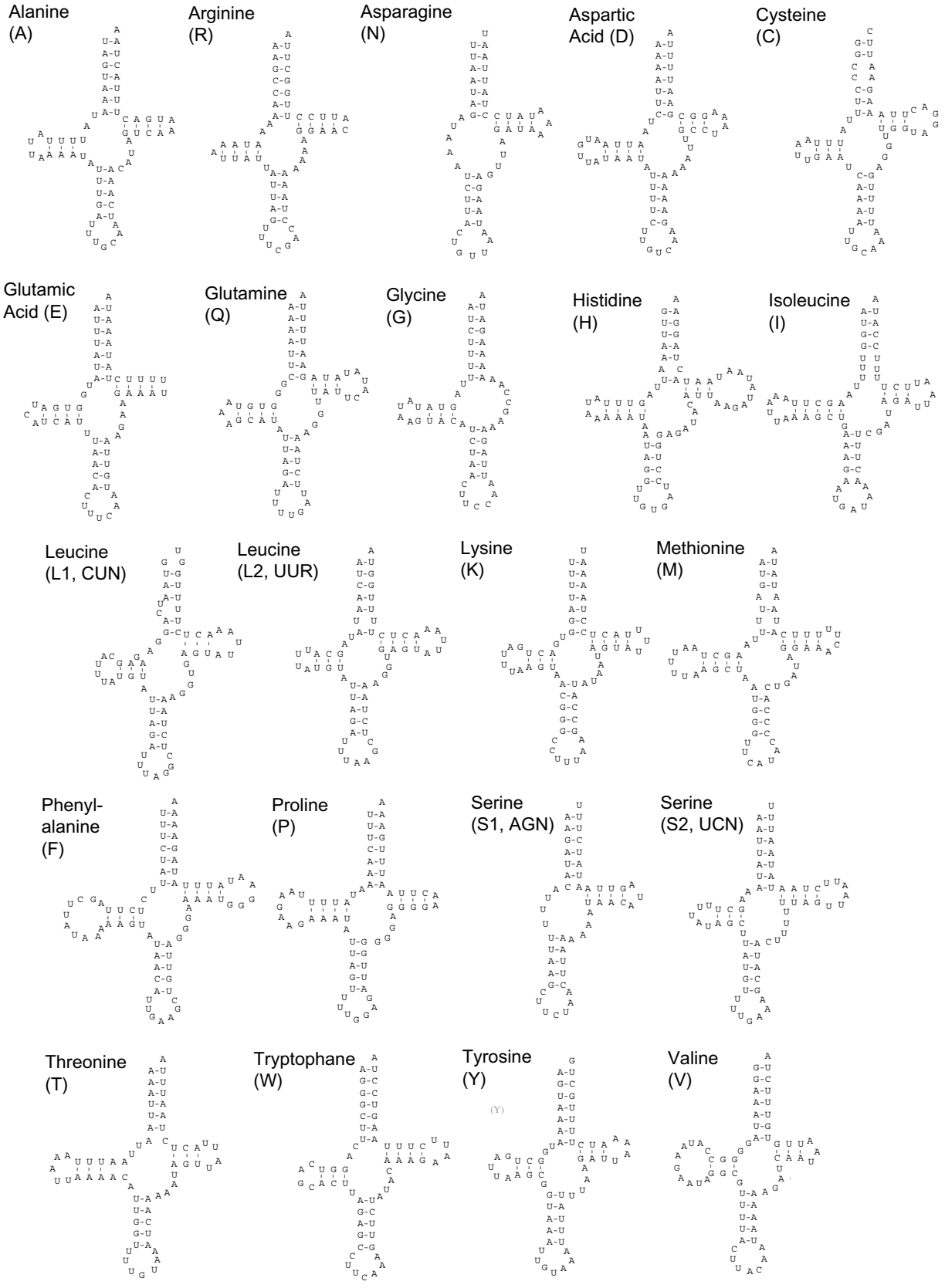

Fig. 2. Putative secondary structures of tRNAs from the mitochondrial genome of *S. causeyae*.

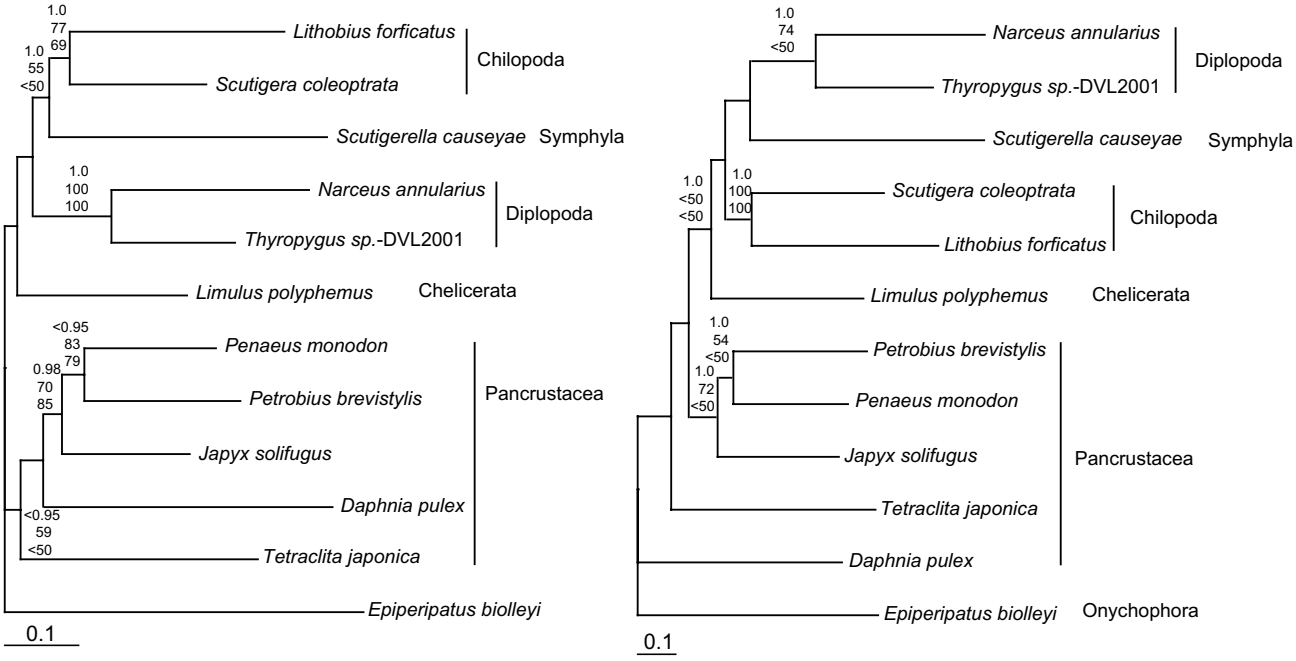

Fig. 3. Phylogenetic trees of myriapod relationships obtained with the nucleotide dataset (left) and the amino acid dataset (right) from 11 mitochondrial protein-coding genes. Branch lengths and topologies from Bayesian analyses. Numbers above branches specify posterior probabilities from Bayesian inference (BPP), bootstrap percentages from maximum likelihood (ML, 100 pseudoreplicates) and maximum parsimony analysis (MP, 1000 pseudoreplicates). No numbers above a branch indicate that BPP < 0.95, and ML and MP bootstrap percentages are <50.

(and hence secondary structure) of a native mitochondrial tRNA is modified after transcription and therefore derives from the mitogenomic sequence presented here.

3.4. Phylogenetic analysis

Phylogenetic analysis was performed using conserved blocks of nucleotide and amino acid sequences of protein-coding genes. In order to elucidate the phylogenetic position of Symphyla, all complete mitochondrial genomes from myriapods (two diplopods, two chilopods and the symphylan) and from seven outgroup species covering Chelicerata, Crustacea, Hexapoda and Onychophora were included in the analysis (Fig. 3). In the topology favoured by Bayesian inference and maximum likelihood analysis with the nucleotide dataset *S. causeyae* is sister to the two species of Chilopoda, well supported by a Bayesian posterior probability (BPP) of 1.0, but only weakly supported by a maximum likelihood (ML) bootstrap percentage of 55. In contrast, analyses with the amino acid dataset retrieved a topology where *S. causeyae* is found as sister to the two diplopod species, no matter if the mtArt (ML) or mtRev (BI) model was used, but BPP is <0.95 and ML/MP bootstrap percentages are less than 50%.

Both datasets favour a topology where *L. polyphemus*, representing the Chelicerata, clusters with Myriapoda, but only the Bayesian analysis with the amino acid dataset gives further support for this relationship (BPP = 1.0). While Diplopoda and Chilopoda are well supported by BPP and ML bootstrap percentages with both datasets, Myriapod monophyly (also recovered in the best topologies) found no support from resampling techniques (BPP <0.95, ML/MP bootstrap <50%).

4. Discussion

4.1. Events leading to changes in gene order of S. causeyae

The most parsimonious hypothesis for the ground pattern of arthropod mitochondrial gene order is the one shared by *L. polyphemus* (Xiphosura), *Ixodes* spp. (Acari), *Carios moubata* (Acari), *Heptathela hangzhouensis* (Araneae) and *Achelia bituberculata* (Pycnogonida). It differs from the pattern found in many crustaceans and hexapods in the position of tRNA-Leu(CUN). Arthropod gene orders differing from these two patterns are restricted to single species or closely related species, therefore probably represent secondarily derived conditions. The gene order of the onychophoran *E. biolleyi* (DQ666064) differs to a large extent from the arthropod ground pattern, but this is probably due to changes restricted to onychophorans or onychophoran subtaxa. In contrast the recently published mitochondrial sequence of *Priapulus caudatus* (Priapulida) differs from the arthropod ground pattern in only one large inversion covering about half of the genome (Webster et al., 2006, 2007). The relative position of tRNA-Leu (CUN) in *P. caudatus* is the same as in chelicerates and myriapods, clearly supporting the view that this represents the ground pattern condition, while that of Pancrustacea represents the derived state, as suggested before with an

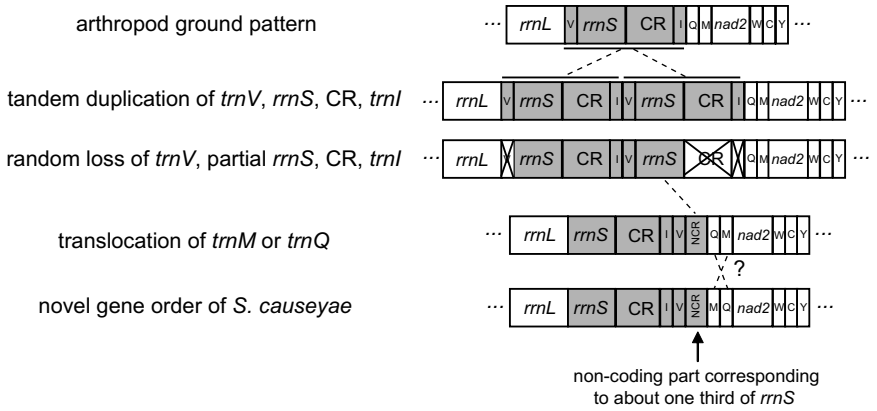

Fig. 4. Hypothesis of tandem duplication of a fragment of the mitochondrial genome and random loss of genes, leading to the derived gene order of *S. causeyae*.

incomplete dataset from an onychophoran species as outgroup (Boore et al., 1998).

Compared to the arthropod ground pattern, represented by the gene order of *L. polyphemus* (Staton et al., 1997), at least three events have to be assumed to obtain the gene order of *S. causeyae* (Fig. 4). Translocation of *trnV* is probably the result of a duplication event also involving *rrnS*. In *S. causeyae trnV* is neighbour to the non-coding region with similarity to *rrnS*, both are located between *trnI* and *trnM*. Gene duplication of *trnV* and *rrnS*, followed by a deletion of a major part of *rrnS*, would be a reasonable explanation for this gene order (tandem duplication and random loss model; Moritz et al., 1987). If a tandem duplication is assumed, then *trnI* and the control region must have been involved in that event as well, otherwise the position of *trnV* and the pseudo-*rrnS* is not easily explained. Subsequent loss of *trnV* in the first replicate and loss of partial *rrnS*, control region and *trnI* in the second replicate have led to the gene order present now in the *S. causeyae* mitochondrial genome (Fig. 4). In addition to this duplication/random loss event, *trnT* and *trnP* have swapped their position, as well as *trnM* and *trnQ*. In both cases it is not determinable which of the genes has probably moved and which maintained in its position.

None of these changes in gene order is found in any of the other four myriapods for which complete mtDNA sequences are known (Fig. 1). The two diplopod species show a number of shared derivations from the arthropod ground pattern, probably the result of one duplication and non-random loss events, followed by a single translocation of *trnT* (Lavrov et al., 2002). In *L. forficatus trnC* is found between the control region and *trnI* (Lavrov et al., 2000), while the mitochondrial genome of *S. coleoptrata* shows several other derivations from the arthropod ground pattern (Negrisolo et al., 2004a). Two events may explain the differences between this gene order and the putative arthropod ground pattern. A tandem duplication of the complete mitochondrial genome, followed by random loss of genes may explain the majority of gene translocations. The first copy has lost *trnM*, *trnC*, *trnY*, *nad3*, *trnN*, *trnH*, *nad4*, *trnT*, *trnP*, *cob*, both *trnL*, *trnV*, *rrnS* and *rrnL*, while the second copy has retained these genes and lost the remainder. A second event must have led to the position of *trnI* right behind *rrnS*, on the opposite side from the control region than in other arthropods. With the exception of the events shared by both diplopod species none of these characters are of phylogenetic value concerning the questions of myriapod monophyly and interrelationships among Chilopoda, Diplopoda and Symphyla (Fig. 5).

4.2. Phylogenetic position of Symphyla

An analysis of a large morphological dataset and a combined dataset of nine genes and morphology (Giribet et al., 2005) supported both the traditional view of a close relationship between Symphyla, Pauropoda and Diplopoda

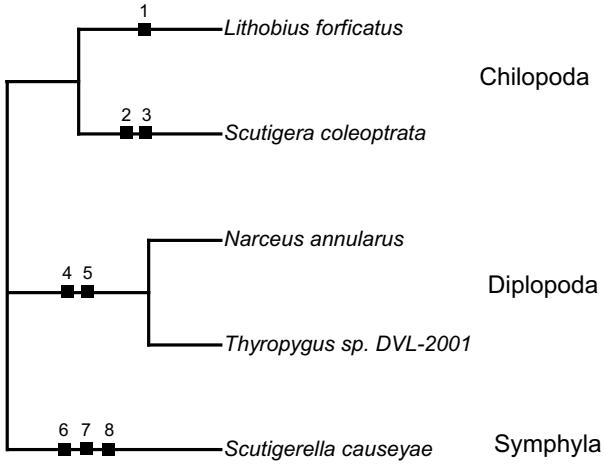

Fig. 5. Putative translocation events in the evolution of myriapod mitochondrial genomes. (1) translocation of *trnC*; (2) tandem duplication of the complete genome and random loss of genes; (3) translocation of *trnI*; (4) tandem duplication of whole genome and non-random loss of genes; (5) translocation of *trnT*; (6) tandem duplication of *trnV*-*rrnS*-CR-*trnI* and random loss of genes; (7) exchange between *trnQ* and *trnM*; (8) exchange between *trnT* and *trnP*.

(Progoneata) and placed this taxon as sister group to Chilopoda in a monophyletic Myriapoda. In contrast, molecular data from combined 18S and 28S sequences even failed to place Symphyla among Euarthropoda (Mallatt and Giribet, 2006), while monophyly of the remaining myriapods (Chilopoda and Diplopoda) was well supported. Another 18S/28S analysis with emphasis on myriapod relationships found support for a Symphyla + Pauropoda clade forming the sister group to Chilopoda (Gai et al., 2006). Nuclear rRNA sequences of Symphyla appear to be highly derived and lead to very long branches of symphylan species in the phylogenetic analyses mentioned above. In a broad analysis of myriapod relationships using three nuclear protein-coding genes (Regier et al., 2005), the monophyly of Myriapoda, Diplopoda and Chilopoda is well supported, but the positions of Symphyla and Pauropoda remain unresolved.

In our analysis, nucleotide sequences from mitochondrial protein-coding genes provide weak support for a relationship between Chilopoda and Symphyla, as proposed in one of the rRNA analyses (Gai et al., 2006). However, this result was not recovered in the analyses performed with the amino acid dataset. Different evolutionary models (ML: mtArt; BI: mtRev) for the amino acid dataset yielded identical topologies (Symphyla + Diplopoda), but without bootstrap or BPP support. It seems that either the nucleotide or the amino acid model does not properly reflect evolutionary change in these datasets. Although the substitution rate in mitochondrial sequences (e.g., reflected in branch lengths in Bayesian trees, Fig. 3) of myriapods is by far not that imbalanced as shown for 18S and 28S sequences (Gai et al., 2006), the current dataset is still insufficient to resolve interrelationships of the major myriapod subtaxa. Probably a combined analysis of nuclear sequences, mitochondrial data and morphology may help to solve this problem, given a better taxon sampling of combined data.

Acknowledgments

The authors thank Thomas Bartolomaeus (Freie Universität Berlin, Germany) for his kind support through all stages of this work. We are also grateful to Greg Edgecombe (Australian Museum) for critical comments on the manuscript and linguistic improvements and likewise acknowledge the improvement of the manuscript by an anonymous reviewer (Reviewer 2). This study was supported by the German Science Foundation, Grant DFG Ba 1520/10-1 (to L.P.).

References

Abascal, F., Posada, D., Zardoya, R., 2007. MtArt: a new model of amino acid replacement for arthropoda. Mol. Biol. Evol. 24, 1–5.

Abascal, F., Zardoya, R., Posada, D., 2005. ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21, 2104–2105.

Avise, J.C., 2000. Phylogeography—the History and Formation of Species. Harvard University Press, Cambridge, Mass.

Ballard, J.W., Whitlock, M.C., 2004. The incomplete natural history of mitochondria. Mol. Ecol. 13, 729–744.

Boore, J.L., Collins, T.M., Stanton, D., Daehler, L.L., Brown, W.M., 1995. Deducing the pattern of arthropod phylogeny from mitochondrial DNA rearrangements. Nature 376, 163–165.

Boore, J.L., Lavrov, D.V., Brown, W.M., 1998. Gene translocation links insects and crustaceans. Nature 392, 667–668.

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17, 540–552.

Edgecombe, G.D., 2004. Morphological data, extant Myriapoda, and the myriapod stem-group. Contrib. Zool. 73, 207–252.

Gai, Y.H., Song, D.X., Sun, H.Y., Zhou, K.Y., 2006. Myriapod monophyly and relationships among myriapod classes based on nearly complete 28S and 18S rDNA sequences. Zool. Sci. 23, 1101–1108.

Giribet, G., Richter, S., Edgecombe, G.D., Wheeler, W.C., 2005. The position of crustaceans within Arthropoda—evidence from nine molecular loci and morphology. Crust. Issues 16, 307–330.

Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids Symp. Ser. 41, 95–98.

Hassanin, A., 2006. Phylogeny of Arthropoda inferred from mitochondrial sequences: strategies for limiting the misleading effects of multiple changes in pattern and rates of substitution. Mol. Phylogenet. Evol. 38, 100–116.

Hassanin, A., Leger, N., Deutsch, J., 2005. Evidence for multiple reversals of asymmetric mutational constraints during the evolution of the mitochondrial genome of Metazoa, and consequences for phylogenetic inferences. Syst. Biol. 54, 277–298.

Haupt, J., 1977. Untersuchung zur Verbreitung und Phänologie von *Scutigerella causeyae* (Symphyla), *Allopauropus vulgaris* (Pauropoda) und *Eosentomon transitorium* (Protura) im Berliner Naturschutzgebiet Langes Luch. Sber. Ges. Naturf. Freunde N. F. 17, 28–38.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755.

Hwang, U.W., Friedrich, M., Tautz, D., Park, C.J., Kim, W., 2001. Mitochondrial protein phylogeny joins myriapods with chelicerates. Nature 413, 154–157.

Kraus, O., 2001. "Myriapoda" and the ancestry of the Hexapoda. Ann. Soc. Entomol. France 37, 105–127.

Lavrov, D.V., Boore, J.L., Brown, W.M., 2002. Complete mtDNA sequences of two millipedes suggest a new model for mitochondrial gene rearrangements: duplication and nonrandom loss. Mol. Biol. Evol. 19, 163–169.

Lavrov, D.V., Brown, W.M., Boore, J.L., 2000. A novel type of RNA editing occurs in the mitochondrial tRNAs of the centipede *Lithobius forficatus*. Proc. Natl. Acad. Sci. USA 97, 13738–13742.

Lowe, T.M., Eddy, S.R., 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25, 955–964.

Mallatt, J., Giribet, G., 2006. Further use of nearly complete 28S and 18S rRNA genes to classify Ecdysozoa: 37 more arthropods and a kinorhynch. Mol. Phylogenet. Evol. 40, 772–794.

Mallatt, J.M., Garey, J.R., Shultz, J.W., 2004. Ecdysozoan phylogeny and Bayesian inference: first use of nearly complete 28S and 18S rRNA gene sequences to classify the arthropods and their kin. Mol. Phylogenet. Evol. 31, 178–191.

Masta, S.E., Boore, J.L., 2004. The complete mitochondrial genome sequence of the spider *Habronattus oregonensis* reveals rearranged and extremely truncated tRNAs. Mol. Biol. Evol. 21, 893–902.

Moritz, C., Dowling, T.E., Brown, W.M., 1987. Evolution of animal mitochondrial-DNA—relevance for population biology and systematics. Ann. Rev. Ecol. Syst. 18, 269–292.

Negrisolo, E., Minelli, A., Valle, G., 2004a. Extensive gene order rearrangement in the mitochondrial genome of the centipede *Scutigera coleoptrata*. J. Mol. Evol. 58, 413–423.

Negrisolo, E., Minelli, A., Valle, G., 2004b. The mitochondrial genome of the house centipede *Scutigera coleoptrata* and the monophyly versus paraphyly of myriapods. Mol. Biol. Evol. 21, 770–780.

Ojala, D., Montoya, J., Attardi, G., 1981. tRNA punctuation model of RNA processing in human mitochondria. Nature 290, 470–474.

Pisani, D., Poling, L.L., Lyons-Weiler, M., Hedges, S.B., 2004. The colonization of land by animals: molecular phylogeny and divergence times among arthropods. BMC Biol. 2, 1.

Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Regier, J.C., Wilson, H.M., Shultz, J.W., 2005. Phylogenetic analysis of Myriapoda using three nuclear protein-coding genes. Mol. Phylogenet. Evol. 34, 147–158.

Rubinoff, D., Holland, B.S., 2005. Between two extremes: mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. Syst. Biol. 54, 952–961.

Scheller, U., Adis, J., 2002. Symphyla. In: Adis, J. (Ed.), Amazonian Arachnida and Myriapoda. Pensoft, Sofia, pp. 547–554.

Shaw, K.L., 2002. Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: what mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. Proc. Natl. Acad. Sci. USA 99, 16122–16127.

Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690.

Staton, J.L., Daehler, L.L., Brown, W.M., 1997. Mitochondrial gene arrangement of the horseshoe crab *Limulus polyphemus* L.: conservation of major features among arthropod classes. Mol. Biol. Evol. 14, 867–874.

Strausfeld, N.J., Strausfeld, C.M., Loesel, R., Rowell, D., Stowe, S., 2006. Arthropod phylogeny: onychophoran brain organization suggests an archaic relationship with a chelicerate stem lineage. Proc. Roy. Soc. B Biol. Sci. 273, 1857–1866.

Swofford, D.L., 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4.0b10. Sinauer associates, Sunderland, MA.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. Clustal-W—improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Webster, B.L., Copley, R.R., Jenner, R.A., Kenzie-Dodds, J.A., Bourlat, S.J., Rota-Stabelli, O., Littlewood, D.T.J., Telford, M.J., 2006. Mitogenomics and phylogenomics reveal priapulid worms as extant models of the ancestral Ecdysozoan. Evol. Dev. 8, 502–510.

Webster, B.L., Kenzie-Dodds, J.A., Telford, M.J., Littlewood, D.T.J., 2007. The mitochondrial genome of *Priapulus caudatus* Lamarck (Priapulida: Priapulidae). Gene 389, 96–105.

Willmann, R., 2003. Die phylogenetischen Beziehungen der Insekten: offene Fragen und Probleme. Verh. Westdeutsch. Entomol. Tag. 2003, 1–64.

Wolstenholme, D.R., 1992. Animal mitochondrial DNA: structure and evolution. Int. Rev. Cytol. 141, 173–216.

Xia, X., Xie, Z., 2001. DAMBE: software package for data analysis in molecular biology and evolution. J. Hered. 92, 371–373.

Xia, X., Xie, Z., Salemi, M., Chen, L., Wang, Y., 2003. An index of substitution saturation and its application. Mol. Phylogenet. Evol. 26, 1–7.

Yamauchi, M.M., Miya, M.U., Machida, R.J., Nishida, M., 2004. PCR-based approach for sequencing mitochondrial genomes of decapod crustaceans, with a practical example from kuruma prawn (*Marsupenaeus japonicus*). Mar. Biotechnol. 6, 419–429.