pagepress

SYSTEMATICS AND PHYLOGENY

The mitochondrial genome of the antarctic springtail *Folsomotoma octooculata* (Hexapoda; Collembola), and an update on the phylogeny of collembolan lineages based on mitogenomic data

A. Carapelli,[1] P. Convey,[2] F. Nardi,[1] F. Frati[1]

[1]Department of Life Sciences, University of Siena, Italy; [2]British Antarctic Survey, Natural Environment Research Council, Cambridge, UK

Abstract

This study provides the description of the complete mitochondrial genome of the Antarctic collembola *Folsomotoma octooculata*. Since mitogenomic data are extensively used for phylogenetic reconstruction, we also provide a phylogenetic reconstruction of the internal relationships of some collembola taxa using the complete mitochondrial DNA data available at present for this group. The complete mitochondrial genome sequence of *F. octooculata* has been obtained using standard amplification and sequencing methods for long DNA templates. Sequence data are analyzed using bayesian methods based on the analysis of nucleotide and amino acid sequences. The complete set of available mitochondrial DNA sequences for ten species of the group is also analyzed for the presence of lineage-specific gene order of the 37 encoded genes. The results of the phylogenetic analysis are compared with previous morphological and molecular analyses, showing some intriguing outcomes, such as the paraphyly of Entomobryomorpha, the derived position of *Podura aquatica* within Poduromorpha, and the sister group relationships between the single Symphypleona species available for comparison and the entomobryid *Orchesella villosa*. The *F. octooculata* mitochondrial genome generally conforms with those observed in other basal hexapod species and displays the plesiomorphic gene order observed for Pancrustacea. Mitogenomic data are useful to reconstruct a molecular phylogeny of major collembolan taxa. In spite of the limited number of species for which entire mitochondrial genomes are known, the coverage of many of the major lineages allows some preliminary conclusions concerning the relationships among families and suborders.

Correspondence: Antonio Carapelli, Department of Life Sciences, University of Siena, Via Aldo Moro 2, 53100 Siena, Italy.
Tel.: +39.0577.234410 - Fax: +39.0577.234476.
E-mail: antonio.carapelli@unisi.it

Key words: Collembola, molecular phylogenetics, mitochondrial DNA, molecular evolution, Antarctica, nucleotide bias.

Acknowledgments: the logistic support of BAS, during the collection of the material, is acknowledged. Partial support was also provided by the University of Siena. We also thank Dr. Alessandra Cosenza and Dr. Alessandro Manenti for their collaboration with the analysis of data.

Funding: the Italian Program of Research in Antarctica (PNRA), the British Antarctic Survey (BAS) and the Italian MIUR (PRIN).

Contributions: AC collected the specimen, sequenced the new mitochondrial genome, performed the molecular and phylogenetic analyses, and drafted the manuscript; FN assisted with the phylogenetic analysis and with the draft of the manuscript; FF directed the research and produced the final manuscript; PC collaborated with the sample of the specimen and greatly improved the draft of the final manuscript.

Received for publication: 20 December 2013.
Accepted for publication: 24 December 2013.

Entomologia 2014; 2:190
doi:10.4081/entomologia.2014.190

Introduction

Collembola (=springtails) are a primitive wingless basal-hexapod group ubiquitous in every damp environment, from poles to the equator, where sufficient levels of fresh water are bioavailable. Their long evolutionary history is confirmed by the early appearance in the fossil record, since the Devonian (Hirst & Maulik, 1926), and marks the intermediate steps of land colonization from marine crustacean-like arthropods to flying insect. Recent physiological, morphological and molecular data suggest a direct transition from marine to soil environments (Little, 1990; D'Haese, 2002, 2003), probably from a benthic crustacean. Alternative hypotheses still point to a crustacean origin, but from freshwater dwelling organisms, similar to present branchiopods (Glenner *et al*., 2006). Whatever was their ancestor, Collembola were originally considered one of the five orders of the Apterygota (=Collembola, Protura, Diplura, Microcoryphia and Zygentoma). When the plesiomorphic status of the apterygote condition was recognized (Hennig, 1981), the taxon Apterygota was dismissed and the Collembola grouped with Protura and Diplura in the Entognatha (with mouth parts enclosed in the lateral folds of the head capsule), the sister taxon of the remaining insects (Ectognatha=Insecta *s. s.*; Kristensen, 1981).

While the iterordinal relationships of the basal hexapods are currently subject to detailed revaluation (Jannsens & Christiansen, 2011), even challenging the monophyly of the Hexapoda as traditionally defined (Carapelli *et al*., 2007), here we focus on the internal phy-

logeny of the Collembola, which represent the most species-rich taxon of the basal hexapods and whose phylogenetic history is relevant to understand the early evolutionary steps of ancient lineages of hexapods on the terrestrial environment. Several morphological traits shared by collembolan species are unique features among arthropods, the most relevant of which probably being the reduced number of abdominal segments (6). Some of these traits, like the presence of a jumping organ (furca) are tied with lifestyle in terrestrial environments full of potential predators. Others, as the presence of a ventrally projected osmoregulatory organ (ventral tube), are connected with fluid exchanges and electrolyte balance. From a taxonomic point of view, the ~8000 species are arranged into four major lineages: Entomobryomorpha, Neelipleona, Poduromorpha and Symphypleona.

In the present script we comply with Janssens (http://www.collembola.org) in considering Collembola as a class, the aforementioned 4 lineages as orders, and Arthropleona, including Entomobryomorpha and Poduromorpha as a super order. We avoid the use of the name Symphypleona in its traditional meaning of Symphypleona ss+Neelipleona. While the monophyly of Collembola has never been questioned, internal phylogenetic relationships have been investigated using morphological characters, such as body shape, that is still the most distinctive character used to define major lineages. For instance, an elongated body (from head to the last abdominal tergite) is sufficient to identify the Arthropleona. Within this latter, the occurrence of a reduced I thoracic tergite is typical of Entomobryomorpha, whereas well developed thoracic tergites defines the Poduromorpha. In Symphypleona and Neelipleona, the body shape is globular, with the partial fusion of the thoracic and abdominal tergites, respectively. An exhaustive synthesis of the competing phylogenetic hypothesis between collembolan lineages has been provided by D'Haese (2003). Major focus on this issue centers on the relationships between Symphypleona and Neelipleona (frequently considered sister groups) and on the position of Poduridae (Poduromorpha). This latter taxon is composed of the single species *Podura aquatica* that, due to the hypognathous (rather that prognathous) orientation of the head, has been in turn considered: the sister group of Actaletidae (forming the largely abandoned superfamiliar taxon Metaxypleona), the basalmost lineage of the Poduromorpha or the sister taxon of all remaining Collembola. Recent morphological and molecular studies roughly support the basal position of Symphypleona and/or Neelipleona and a more derived position for *P. aquatica* (D'Haese, 2003; Schneider *et al*., 2011).

In this study we are investigating the phylogenetic relationships of those collembolan lineages for which entire mitochondrial genomes (mtDNAs) are available. The mtDNA is an independent molecular marker that coexists with the nuclear DNA within the eukaryotic cells. With its peculiar (matrilinear) mode of inheritance, the putative absence of recombination and relatively high nucleotide substitution rates with respect to nuclear DNA, the mtDNA is frequently used to study the phylogenetic relationships of animal lineages at different taxonomic levels. Phylogenetic reconstruction also exploit differences in gene order (Lavrov & Lang, 2005). Over the last two decades the mitochondrial genome has been extensively used to reconstruct arthropod relationships, sometimes providing support to traditional hypotheses, and sometimes suggesting new hypotheses, conflicting with the existing ones. Areas of disagreement have also been found between phylogenetic reconstructions drawn from mitochondrial and nuclear genome data (see below).

The mt-genome of animals is usually a circular double-stranded molecule of 11-20 kb in size. It contains the genes encoding for 13 protein products involved in the oxidative phosphorylation (OXPHOS) process, 22 transfer RNA (*trnX*), and two ribosomal rRNAs (*rrnS* and *rrnL*). Generally, a single large, noncoding region (known as AT-rich region in arthropods) includes the signals for the initiation of replication and transcription of the mtDNA.

In this study, we also describe the features of the mtDNA of the Antarctic species *Folsomotoma octooculata* (Arthropleona, Isotomidae), providing additional data to increase the size of the collembolan mitogenomic sample.

Materials and Methods

Sequencing and analysis of the mitochondrial genome of *F. octooculata*

One specimen of *F. octooculata* from the Antarctic Peninsula (Maritime Antarctica) was sampled in Devils Point, Livingston Island (62°40'S, 61°11'W) during the 2002-03 expedition of the British Antarctic Survey (BAS), in collaboration with the Italian National Antarctic Program (PNRA).

The specimen was identified, frozen and preserved at -80°C for molecular analyses. Total genomic DNA was purified using the Wizard SV genomic DNA purification system (Promega) and used for all amplifications. Partial *cox3* and *rrnL* sequences were initially produced using universal primer pairs (Simon *et al*., 1994) and standard amplification procedures. Four specific primers were then designed on these sequences to amplify the rest of the genome in two segments each of about 7.5 Kb: primer pairs FOC-rrnL-215J (5'-CGTCCCCTAACTAATTAAAGCATTTTC-3')/FOC-cox3-1773N (5'-GCCCCTAAAATTGAAGATACTCCCGC-3') and FOC-cox3-726J (5'-GGGTTTGAAGCAGCAGCATGATAC-3')/FOC-rrnL-215N (5'-GAAAATGCTTTAATTAGTTAGGGGACG-3'). Long PCR conditions were the following for both primer pairs: 94°C for 1 min, 60°C for 1 min and 68°C for 10 min, 35 cycles. Amplifications were performed in 25 μL reaction volume composed of: 10.75 μL of sterilized distilled water, 2.5 μL of LA PCR Buffer II (Takara), 2.5 μL of 25 mM $MgCl_2$, 4 μL of dNTPs mix, 1.25 μL of each primer (10 μM), 2.5 μL of DNA template and 0.25 μL (1.25 U) of Takara LA Taq polymerase (Takara). Long-PCR fragments were purified using the Kit Wizard SV Gel and PCR Clean-up System (Promega) and sequenced on both strands with a primer walking approach using specific primers (list available upon request). The consensus sequence was produced using Sequencher 4.4.2 (Gene Codes) and deposited in GenBank (accession number KC862316).

The location of protein-encoding (PCGs), rRNA and tRNA genes was identified based on their amino acid translation or secondary structure features, respectively. The largest non-coding region between *rrnS* and *trnI* was assigned to the presumed AT-rich region. Gene boundaries and overlaps, as well as start and stop codons of PCGs were identified. Nucleotide composition and codon usage statistics were calculated using PAUP* 4b10 and codonw (Peden, 1999), respectively. Strand asymmetry was calculated using the formulas AT-skew=[A%-T%]/[A%+T%] and CG-skew=[C%-G%]/[C%+G%] (Perna & Kocher, 1995; Hassanin *et al*., 2005). The presence of repeated sequences within the largest intergenic region was assessed using the mreps software (Kolpakov *et al*., 2003).

Ten specimens from different collembolan species (Table 1) have been analyzed for the segment *cob/nad1*, using a pair of primers specifically designed for collembolan taxa: Coll-cob-748J (5'-AGAAAAYTTTATCCCWGCYAACC-3')/Coll-nad1-454N (5'-GGRGTWGCYCARACYATTTCTTATG-3'). PCR amplifications have been performed using the enzyme Go-Taq Flexi (Promega), applying the protocol provided by the manufacturer. Thirty-five PCR thermal cycles were accomplished with the following series of temperature steps: 95°C for 1 min, 50°C for 1 min and 72°C for 10 min, in a GeneAmp PCR System 2700 (Applied Biosystems). Inspection for the presence of tRNA-encoding genes within gene junction *cob/nad1* was performed using the program ARWEN (Laslett & Canbäck, 2008).

Phylogenetic analyses

For the phylogenetic analysis we have downloaded nine of the ten complete (or almost complete, in the cases of *Onychiurus orientalis* and *Podura aquatica*) collembolan mtDNAs available in GenBank. For the species *F. grisea* we have arbitrarily selected only one of the two available mtDNA sequences (Torricelli *et al*., 2010). We then added the new mtDNA sequence of *F. octooculata* (see complete list of specimens in Table 1), and the mtDNA sequences from the dipluran *Japyx solifugus* (accession number NC_007214) and the microcoryphian *Trigoniophthalmus alternatus* (accession number NC_010532) which were used as outgroups. Each of the 13 protein-encoding genes was independently aligned based on its aminoacid sequence and the corresponding retroalignment produced using the Rev Trans 1.4 server (http://www.cbs.dtu.dk/services/RevTrans) (Wernersson & Pedersen, 2003). Final alignments were concatenated into a single data set of 11,985 nucleotide positions. After the exclusion of 2,337 sites of dubious homology, the final data set included 9648 nucleotide sites. Saturation was assessed at the level of codon positions using the test by Xia *et al*. (2003) as implemented in Dambe 5.3.19 (Xia & Xie, 2001). The best partitioning strategy and evolutionary model for each partition was determined using PartitionFinder 1.0.1 (Lanfear *et al*., 2012). Twenty-six initial blocks were defined by dividing the alignment according to codon positions (first and second, as third were removed, see below) and gene. Blocks were joined using the greedy algorithm, limiting models to those implemented in MrBayes (Ronquist & Huelsenbeck, 2003), under a BIC criterion. Three additional *user* partitioning schemes were also evaluated, *i.e*., partitioning according to strand and codon position, partitioning according to strand with *atp8* and *nad6* separated from the rest and codon position, partitioning according to strand and codon position with 1st positions of *cox1* as a separate partition. The selected partitioning scheme and evolutionary models were applied in a bayesian analysis using MrBayes 3.2.1 on the Cyberinfrastructure for Phylogenetic Research (CIPRES) server (Miller *et al*., 2010). Two parallel runs, each consisting of four chains, were run for 50 million generations sampling every 500th state and removing 10% as burnin. The best partitioning strategy and evolutionary model for each partition in the aminoacid alignment was initially determined using PartitionFinderProtein 1.0.1 (Lanfear *et al*., 2012). Due to software limitations that do not allow to implement the suggested model in an efficient way, we reverted to a simpler, yet robust, analysis using MrBayes. Two parallel runs, each consisting of four chains, were run for 5 million generations sampling every 500th state and removing 10% as burnin. The aminoacid *mixed* model, *i.e*., a procedure where all different available protein matrices are applied and each contributes to the final outcome according to its posterior probability, was applied with no partitioning.

Results

Gene composition and other molecular features

The mitochondrial genome of *F. octooculata* is a closed circular molecule of 15,338 bp in length. It contains the usual set of 37 genes typical

Table 1. List of species (and their taxonomic classification and accession number) analyzed for this phylogenetic study, with the reported gene order type deduced for complete mt-genomes (°) and the presumed gene order hypothesized from the analysis of the *cob/nad1* gene junction (rest of the species).

Order	Family	Species	Gene order	Acc. number
Entomobryomorpha				
	Entomobryidae	*Orchesella villosa*°	GO1	NC010534
	Isotomidae	*Folsomota octooculata*°	GO1	KC862316
		Gressittacantha terranova	GO1	KF982829
		Cryptopygus antarcticus°	GO1	NC010533
	Paronellidae	*Cyphoderis albinus*	GO1	KF982827
	Tomoceridae	*Pogonognathellus longicornis*	GO1	KF982831
Poduromorpha				
	Hypogastruridae	*Gomphiocephalus hodgsoni*°	GO1	NC005438
	Neanuridae	*Anurida maritima*	GO1	KF982826
		Bilobella aurantiaca°	GO1	NC011195
		Friesea grisea°	GO1	NC010535
	Onychiuridae	*Onychiurus orientalis*°	GO2*	NC006074
		Tetrodontophora bielanensis°	GO2	NC002735
		Tullbergia mixta	GO2	KF982833
	Poduridae	*Podura aquatica*°	GO1*	NC006075
Symphypleona				
	Dicyrtomidae	*Dicyrtomina saundersi*	GO1	KF982828
	Sminthuridae	*Allacma fusca*	GO3	KF982824
		Allacma gallica	GO3	KF982825
		Sminthurus viridis°	GO3	NC010536
	Sminthurididae	*Sminthurides aquaticus*	GO1	KF982832
Neelipleona				
	Neelidae	*Neelus murinus*	GO1	KF982830

°The ten species are those included in the phylogenetic analyses based on sequence data, reported in Figures 3 and 4. *Non-complete mtDNA sequences.

of metazoan mtDNAs and displays the gene order considered ancestral for Pancrustacea (Figure 1) (Carapelli *et al*., 2008) and most frequently observed in Collembola. The majority (23/37) of genes are oriented on one strand, that is therefore named J (from maJority) strand. Consequently, the specular strand is called N (from miNority). Non-coding regions are reduced at minimum with only three exceptions for the so-called AT-rich region (496 bp: putatively hosting the sequences controlling replication and transcription of the genome), and two smaller intergenic spacers, at the *trnSuga/nad1* and *trnE/trnF* gene junctions, of 218- and 99-bp, respectively. Overlaps between genes occur in three instances for one single nucleotide (*trnI/trnQ*; *trnR/trnN* and *trnP/nad6*) and in one case for 7 bp (*atp8/atp6*) (Table 2). Notably the overlaps between *atp8/atp6* and *trnR/trnN* occur among genes oriented on the same strand. While sharing of the same nucleotide(s) is often observed in mitochondrial genes encoding for ATPase subunits, overlaps of tRNA-encoding genes is less common, and it would imply that posttranscriptional mechanism(s) are necessary to restore the complete functionality of gene products. Eleven out of the 13 PCGs start with canonical ATA or ATG codons (encoding methionine) whereas *cox2* and *nad6* apparently use ATC and ATT (encoding Isoleucine), respectively (Table 2). Three genes (*nad3-5*) have incomplete termination codons (Table 2). Within the AT-rich region, three short repeats [(TA)7 at positions 14,941-14,954; (TAAATGT)2 at positions 15,148-15,161 and (TA)5 at positions 15,208-15,217] are found, whereas a longer (AT)22 repeat is located at the *trnE/trnF* gene junction (positions 6250-6293).

Nucleotide composition

It is well known that the mtDNA of hexapods is AT-rich and that the J-strands have positive values for AT- and CG-skew parameters (Hassanin *et al*., 2005).

In *F. octooculata* the total A+T-content of the mtDNA is 75%, that is the highest value observed among the 10 complete collembolan mitochondrial genomes sequenced so far (the lowest one – 69% – being found in *Sminthurus viridis*).

Although slightly positive skew values (AT-skew=0.04; CG-skew=0.15) are observed for the complete J-strand of *F. octooculata*, the calculation of the skew parameters for the two different sets of nucleotide positions of PCGs grouped according to the coding polarity (all measures calculated on the J-strand) show negative scores only for the AT-skew of the J-oriented genes (Figure 2). Dividing the PCGs sites into I, II and III codon positions, positive skew values are always observed with the only exceptions of AT-skew and CG-skew of J-oriented II and I positions, respectively. The highest values of all comparisons are the AT-skews calculated for II codon positions of PCGs having J- and N-strand polarity (-0.75 and 0.47, respectively) (Figure 2). This latter result is due to the high percentage of Ts observed at II codon positions of protein-encoding genes oriented on both strands. In J-oriented PCGs, this nucleotide trend opposites to the expected (J-) strand bias. High T frequency at second codon positions has a relevant impact also in the most frequently used codons, which are AUU (202, encoding Ile) for the J-stand oriented PCGs, and UUU (151, encoding Phe) for the N-oriented ones. As expected, the most frequently used codons (either at two- or four-fold degenerate families) are those with A or T at third position (Supplementary Tables).

Gene order

Ten complete, or almost complete, collembolan mtDNAs are currently available from GeneBank. Obviously, being Collembola a species-rich group with more than 8000 species described worldwide (http://www.collembola.org), the present taxon sampling does not represent all major lineages of the group, with a coverage of only 7 out of 33 families (Hypogastruridae, Isotomidae, Entomobryidae, Neanuridae, Onychiuridae, Puduridae and Sminthuridae), all from

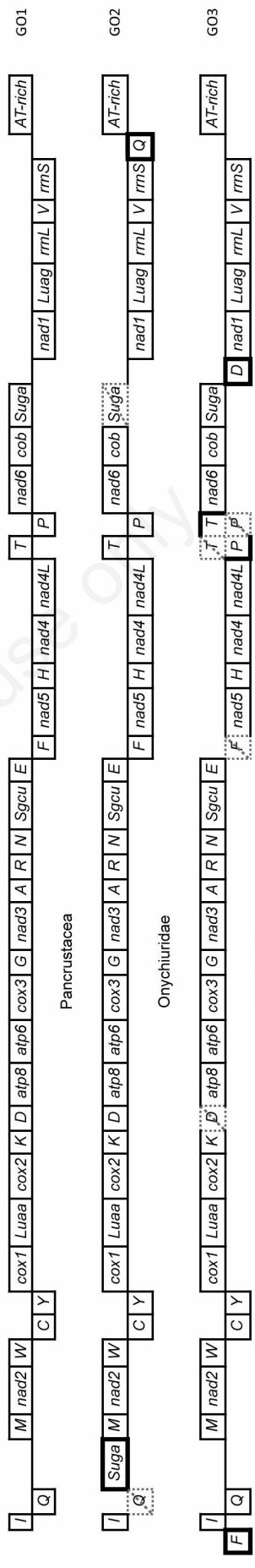

Figure 1. All known gene arrangements for the collembolan mtDNAs: GO1) Pancrustacea (ancestral) model; GO2) Onychiuridae model; GO3) Sminthuridae model. Genes, intergenic regions and overlaps are not drawn to scale. Genes are abbreviated as in the text except for the tRNA designations, which are identified with a single capital letter. Upper bar of each model identifies genes oriented on the J-strand; lower bars displays those with opposite (N-strand) polarity. Thick and dotted boxes highlight *new* and *ancestral* positions for translocated genes.

the three major orders, whereas no data are yet available for Neelipleona. Besides that, some preliminary phylogenetic results may be deduced from the analysis of the arrangement of the 37 genes along the organelle's genome. Collembolan mtDNAs display 3 gene orders (GO1-3; Figure 1). Model GO1 (Figure 1), presumably the ancestral gene order for pancrustaceans, is present in *Bilobella aurantiaca, Cryptopygus antarcticus, F. octooculata, Friesea grisea* (both specimens), *Orchesella villosa* and *P. aquatica*. So far, alternative arrangements have been found only in the two onychiurid species *O. orientalis* and *Tetrodontophora bielanensis* (GO2), and in the symphypleonan *S. viridis* (GO3). GO2 differs from GO1 for the translocation of two tRNA genes (*trnQ* and *trnSuga*) (Figure 1). In GO3, four tRNA genes (*trnD, trnF, trnT* and *trnP*) have changed positions with respect to GO1, resulting in three translocations (*trnD*, *trnT* and *trnP*, the latter two genes having switched positions) and one translocation *plus* inversion for *trnF* (Figure 1). Close examination of the fragment encompassed by *cob* and *nad1* (Figure 1) in some collembolan species (Table 1), combined with data obtained from complete mtDNA sequences, suggests that all species with GO2 and GO3 differ from GO1 for one tRNA gene insertion (*trnD* in GO3) or deletion (*trnSuga* in GO2) in this location (Figure 1). The three onychiurid species *O. orientalis*, *T. bielanensis* and *Tullbergia mixta* do

Table 2. Annotation, nucleotide composition, and other molecular features of the mtDNA of *F. octooculata*.

	A	C	G	T	Length bp	Strand	Positions	Start/stop codon
trnI	0.30	0.16	0.19	0.34	67	J	1-67	
trnQ	0.37	0.03	0.18	0.43	68	N	135-67	
trnM	0.37	0.21	0.12	0.31	68	J	147-214	
nad2	0.35	0.14	0.09	0.43	999	J	215-1213	ATG(M)/TAA
trnW	0.48	0.09	0.09	0.34	65	J	1222-1286	
trnC	0.33	0.11	0.21	0.34	61	N	1348-1288	
trnY	0.34	0.12	0.16	0.37	67	N	1421-1355	
cox1	0.28	0.17	0.17	0.38	1536	J	1428-2963	ATG(M)/TAA
trnLuaa	0.36	0.07	0.12	0.45	69	J	2973-3041	
cox2	0.33	0.16	0.14	0.36	684	J	3042-3725	ATC(I)/TAA
trnK	0.41	0.14	0.11	0.34	71	J	3727-3797	
trnD	0.43	0.03	0.08	0.46	65	J	3799-3863	
atp8	0.42	0.10	0.07	0.41	153	J	3873-4025	ATA(M)/TAA
atp6	0.34	0.16	0.11	0.38	683	J	4019-4702	ATG(M)/TAA
cox3	0.33	0.16	0.15	0.36	789	J	4705-5490	ATG(M)/TAA
trnG	0.48	0.10	0.07	0.36	61	J	5527-5587	
nad3	0.35	0.15	0.13	0.37	343	J	5588-5930	ATA(M)/T—
trnA	0.38	0.10	0.17	0.35	60	J	5931-5990	
trnR	0.38	0.17	0.14	0.32	66	J	5995-6060	
trnN	0.34	0.19	0.16	0.31	64	J	6059-6122	
trnSgcu	0.35	0.14	0.14	0.38	66	J	6127-6192	
trnE	0.41	0.06	0.08	0.45	64	J	6196-6259	
trnF	0.38	0.05	0.17	0.41	64	N	6359-6296	
nad5	0.30	0.10	0.15	0.45	1699	N	8058-6360	ATA(M)/T—
trnH	0.42	0.05	0.15	0.39	62	N	8123-8062	
nad4	0.28	0.10	0.16	0.47	1354	N	9477-8124	ATG(M)/T—
nad4L	0.27	0.09	0.16	0.48	267	N	9489-9755	ATA(M)/TAA
trnT	0.42	0.05	0.06	0.48	65	J	9756-9820	
trnP	0.34	0.08	0.18	0.39	61	N	9882-9822	
nad6	0.38	0.11	0.10	0.41	492	J	9890-10372	ATT(I)/TAA
cob	0.31	0.17	0.14	0.38	1143	J	10377-11519	ATA(M)/TAA
trnSuga	0.35	0.10	0.09	0.46	69	J	11522-11590	
nad1	0.27	0.09	0.17	0.46	933	N	12741-11809	ATA(M)/TAA
trnLuag	0.39	0.06	0.14	0.41	64	N	12807-12745	
rrnL	0.37	0.07	0.12	0.44	1260	N	14068-12809	
trnV	0.38	0.08	0.14	0.40	63	N	14131-14069	
rrnS	0.36	0.09	0.13	0.42	713	N	14842-14132	
A+T-rich	0.41	0.07	0.07	0.45	496		14843-15338	
Total	0.39	0.15	0.11	0.36	15338			

not have *trnSuga* between *nad1* and *cob* 3'ends. In *T. bielanensis* (the only onychiurid with the complete mtDNA sequenced) *trnSuga* and *trnQ* have changed position (Figure 1). In *B. aurantiaca trnSuga* is embedded between two non-coding (nc) regions of 332 and 585 bp, respectively. Of these latter, the largest non coding (nc) region is mostly composed of 5 complete and one partial (63/88) repeated units of 88-bp in length (identity consensus=97.7%).

Within Symphypleona, only the Sminthuridae species (*S. viridis*, *Allacma fusca* and *A. gallica*) share the same translocation of *trnD*, from *trnK-trnD-atp8* to *trnSuga-trnD-nad1* (Figure 1). No rearrangement occurs in this mtDNA fragment in the neelipleonan *Neelus murinus* (Table 1).

Phylogenetic analysis

The final dataset thus consists of 9648 aligned nucleotide positions corresponding to 3216 codons. The saturation test produced values of I_{SS} of 0.61, 0.28, 0.79 for first, second and third codon positions, respectively, while $I_{SS.C}$ were 0.82 for a symmetrical and 0.69 for an asymmetrical tree, indicating saturation has occurred at third, but not at first and second codon positions. Accordingly, third codon positions were removed from the analysis. PartitionFinder suggested the use of a 7-partition model as follows: 1) ATP6_1st+COB_1st+COX2_1st+COX3_1st+NAD2_1st NAD3J_ 1st (model: GTR+I+Γ); 2) ATP6_2nd+COB_2nd+COX1_ 2nd+COX2_2nd+ COX3_2nd+NAD2_2nd+NAD3_2nd (model: GTR+I+ Γ); 3) ATP8_1st+NAD6_ 1st (model: HKY+I+Γ); 4) ATP8_2nd+ NAD6_2nd (model: HKY+I); 5) COX1_1st (model: GTR+I+Γ); 6) NAD1_1st+NAD4_1st+NAD4L_1st+ NAD5_1st (model: GTR+I+Γ); 6) NAD1_2nd+NAD4_2nd+NAD4L_2nd+ NAD5_2nd (model: GTR+I+Γ). The MrBayes analysis converged well regardless of the fairly complex model employed, with average deviation of split frequencies below 0.0001, ESS values all above 17000 and PSRF all 1.000. The resulting family of trees were summarized to obtain the nucleotide-based tree (nu-t), shown in Figure 3, with associated posterior probabilities at nodes.

The final aminoacid dataset consists of 3,216 aligned positions. The MrBayes analysis converged well with average deviation of split frequencies below 0.001, ESS values all above 17,000 and PSRF all 1000. A single aminoacid model (MtRev; Adachi & Hasegawa, 1996) resulted in a posterior probability of 1, and as such contributed totally to the analysis. The resulting family of trees were summarized to obtain the tree (aa-t), shown in Figure 4, with associated posterior probabilities at nodes.

The monophyly of the class and of the 3 families represented by more than one species, Isotomidae (*C. antarcticus+F. octooculata*), Neanuridae (*B. aurantiaca+F. grisea*) and Onychiuridae (*O. orientalis+T. bielanensis*), is fully supported (Figures 3 and 4). The order Poduromorpha is also retrieved as a monophyletic group (both trees), with Onychiuridae basal to the assemblage composed by Hypogastruridae (*Gomphiocephalus hodgsoni*) *plus* Neanuridae (*P. aquatica*,) in the nucleotide based analysis (Figure 3). Internal relationships within Poduromorpha are slightly different in the aa-t, due to the reciprocal exchange of position between *G. hodgsoni* and *P. aquatica*.

In both trees (Figures 3 and 4), Entomobryomorpha are paraphyletic: Isotomidae is the sister-group of a cluster composed by the entomobryid *O. villosa plus* the symphypleonan *S. viridis* in the nu-t, while in the aa-t relationships between Isotomidae, Poduridae and *O. villosa+S. viridis* are not resolved (Figure 4).

Discussion

The mitochondrial genome of *F. octooculata*

At least three different nucleotide biases affect the mitochondrial genome of *F. octooculata*, in agreement with those observed in other collembolan taxa: i) an A+T bias, that leads to a higher percentages of A and T bases (*vs* C and G) on both strands; ii) a strand bias, that determines the A+C *vs* G+T richness of the J-strand; iii) a site specific bias of PCGs, where II codon positions are T-rich.

The A+T richness of both strands is probably dependent from the deamination process that characterizes single and double strand mtDNAs in Hexapoda and other Arthropoda. Indeed, during both replication and transcription of the mitochondrial genome, C>U changes are the most frequent mutation type observed in single strand N-hemihelices and double strand J-helix. This molecular feature has probably led to the striking increase of A and T nucleotide frequencies on both strands, with extreme values, close to saturation, observed in some lineages of Hymenoptera and Diptera (Simon *et al*., 1994). It is still unknown why this trend do not apply to all Metazoa. Due to the asymmetrical replication process of the mitochondrial DNA, the J-strand is more prone to accumulate G to A and T to C nucleotide deaminations, resulting in a higher percentage of A *vs* T and C *vs* G (strand nucleotide bias). Obviously, the opposite trend is observed in the N-strand. However, the effect of the trend is not uniform for all genes and all nucleotide positions. In this respect, selective pressures (*i.e.,* the hydrophobic nature of the internal phospholipids bilayers of the mitochondrion inner membrane) foster the occurrence of a high number of thymines at II codon

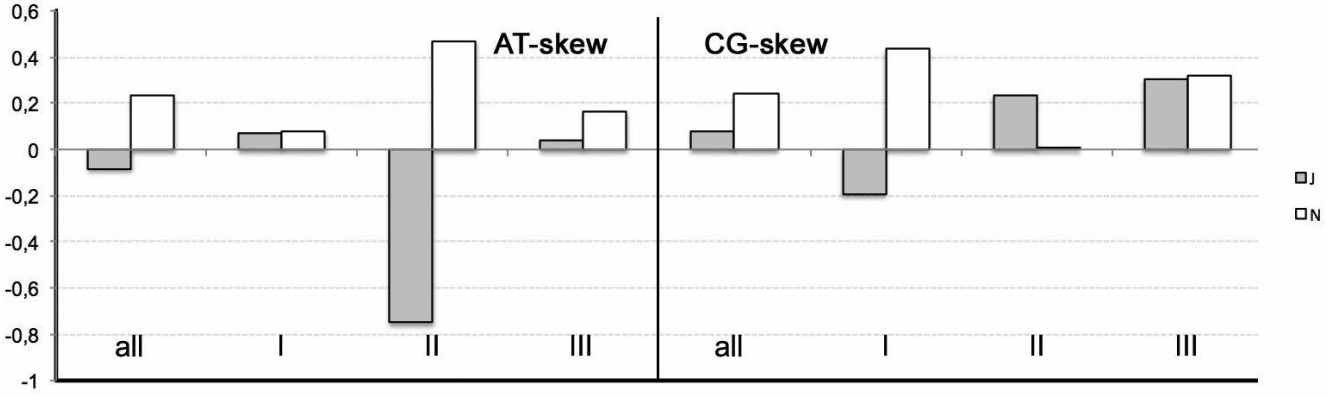

Figure 2. Histogram representation of AT- and CG-skew values in different sets of J- (grey) and N-oriented (white) protein-encoding sequences (calculated on the J-strand): all=all positions of PCGs; first, second, and third positions only.

positions of mitochondrial protein-emcoding genes, given that most hydrophobic amino acids are encoded by NTN triplets (Carapelli *et al*., 2013). This site-specific nucleotide bias applies regardless of the opposite trend apparently working on the J-strand.

The occurrence of regulatory stretches of sequence controlling the multiple transcription units of mitochondrial mRNAs has been demonstrated in *Drosophila* (Roberti *et al*., 2006) and other invertebrates (Fernandez-Silva *et al*., 2001). These intergenic sequences, that act as binding sites for attenuator/terminator regulatory molecules during mRNA synthesis, have been identified between blocks of genes with abrupt shift of transcription polarity, such as those at gene junctions *trnE/trnF* and *trnSuga/nad1*. Some of the collembolan mtDNAs so far analyzed, including that of *F. octooculata*, have intergenic spacers exactly in these locations (Carapelli *et al*., 2008) suggesting that similar regulatory mechanisms may act to control mRNA synthesis in hexapods. Although larger genomes have been observed in basal Metazoa, the mtDNA of animals is typically reduced in size (~14-20 kb), with no introns, few intergenic sequences, some overlaps between adjoining genes oriented on opposite strands and, sometimes, the lack of complete termination codons in PCGs (Boore, 1999). These molecular features may depend on the necessity to compact the genome size in order to guarantee its rapid replication. This trend (not alone) is also considered responsible for the reduction of the gene content along the evolutionary history of Metazoa (Burger *et al*., 2013), given that most of the proteins necessary for oxidative phosphorylation are located within the nuclear (rather than the mitochondrial) genome. In this respect, some genes lack a complete terminal codon, resulting in a T- or TA- 3'-end, and posttranscriptional polyadenilation restores the complete termination signals (Ojala *et al*., 1981). The occurrence of overlapping PCGs, like the *atp8/atp6* junction of *F. octooculata* and other collembolan mt-sequences (Carapelli *et al*., 2008), is a common feature of many animal mt-genes. These bicistronic units are transcribed together and post-transcriptionally cleaved into separate mRNAs before protein synthesis (Krzywinski *et al*., 2006). As for tRNA genes, overlaps that occur between genes oriented on the same strand would produce at least one incomplete acceptor arm (either in up- or downstream gene). Although posttranscriptional modification of mRNAs is frequently observed in the mitochondria of protozoans (Price & Gray, 1999) and plants (Chateigner-Boutin & Small, 2010), some evidence suggest that a similar mechanism(s) may also act in arthropods (Lavrov *et al*., 2000). In Collembola, some mt-tRNA genes show mismatches of acceptor arms and/or overlaps between genes, but post- or co-transcriptionally mechanisms have not been demonstrated yet (Carapelli *et al*., 2008). In animal mtDNA, the occurrence of PCGs with alternative start codons (other than ATG or ATA) is not uncommon. In Collembola, this molecular feature is not lineage specific and appears to be acquired independently in different species.

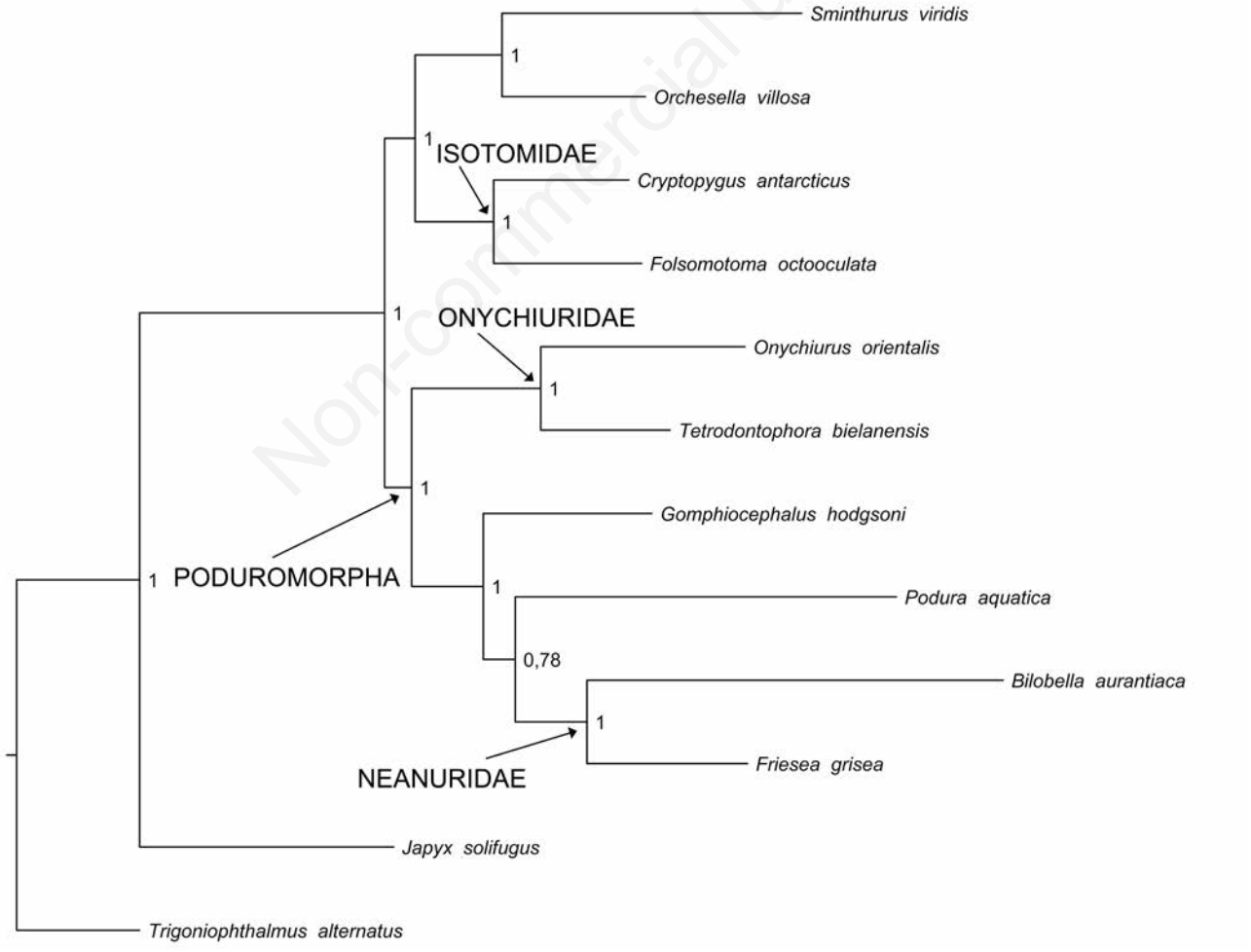

Figure 3. Phylogenetic tree of collembolan taxa, based on I and II nucleotide positions (nu-tree) of 13 mitochondrial protein-encoding genes.

Phylogenetic analysis

Gene order data

It has been shown that the comparison of the mtDNA gene order may provide relevant phylogenetic information (Boore *et al*., 1998). In fact, the arrangement of the 37 genes along the mitochondrial genome is not unique for all animal lineages and changes of the mtDNA gene order occur sporadically during the evolution of taxa, at all taxonomic levels. Due to the idiosyncratic nature of gene order changes, convergence (unrelated taxa exhibiting the same gene order due to independent homoplastic events) is unlikely; therefore, having the same gene arrangement along the mtDNA chromosome is usually evidence of common evolutionary history and (hence) phylogenetic relatedness between taxa (Lavrov & Lang, 2005).

Apparently, GO1 is the most frequent arrangement of mt-genes for the collembolan families Hypogastruridae, Isotomidae, Neanuridae and Poduridae, whereas unique translocations have occurred in members of Onychiuridae and Sminthuridae, suggesting that the underlined molecular changes (gene translocations) in the gene order have occurred in the common ancestor of these taxa.

Partial mtDNA sequencing for other collembolan species (Table 1) suggest that the fragment encompassing the 3'-end of *cob* and the 3'-end of *nad1* (these genes are usually oriented on opposite strands; Figure 1) may represent a hot spot region for gene rearrangements in this group. Within Onychiuridae, the translocation of *trnSuga* (GO2) is shared by species of the three subfamilies Onychiurinae, Tetrodontophorinae and Tullbergiinae and is probably a common molecular feature of the whole family. Differently, GO3 is not present in all Symphypleona, but appears to be unique of the family Sminthuridae.

Although these preliminary data are not conclusive to depict the whole information eventually provided by the study of gene order, the presence of alternative models of gene arrangement across different collembolan lineages suggests that additional phylogenetically useful information may be hidden in the mitochondrial genomes to investigate relationships at the family and ordinal level. In addition, despite the presumed long evolutionary history of major collembolan lineages, it appears that the presumed ancestral gene order state for the Pancrustacea (GO1) is also the most common gene order observed in Collembola, with minor changes occurring only in some taxa. Interestingly, GO1 is also the most frequently observed gene order in three of the remaining basal hexapod taxa (Diplura, Microcoryphia and Zygentoma), with the only notable exception of Protura. In this latter taxon, three species have been so far sequenced for the complete or almost complete mtDNA, with each of the showing a peculiar gene order, and none of them corresponding to GO1.

Sequence-based analysis

Although our taxon sample does not represent the overall diversity of Collembola, and is somewhat unbalanced towards the Poduromorpha,

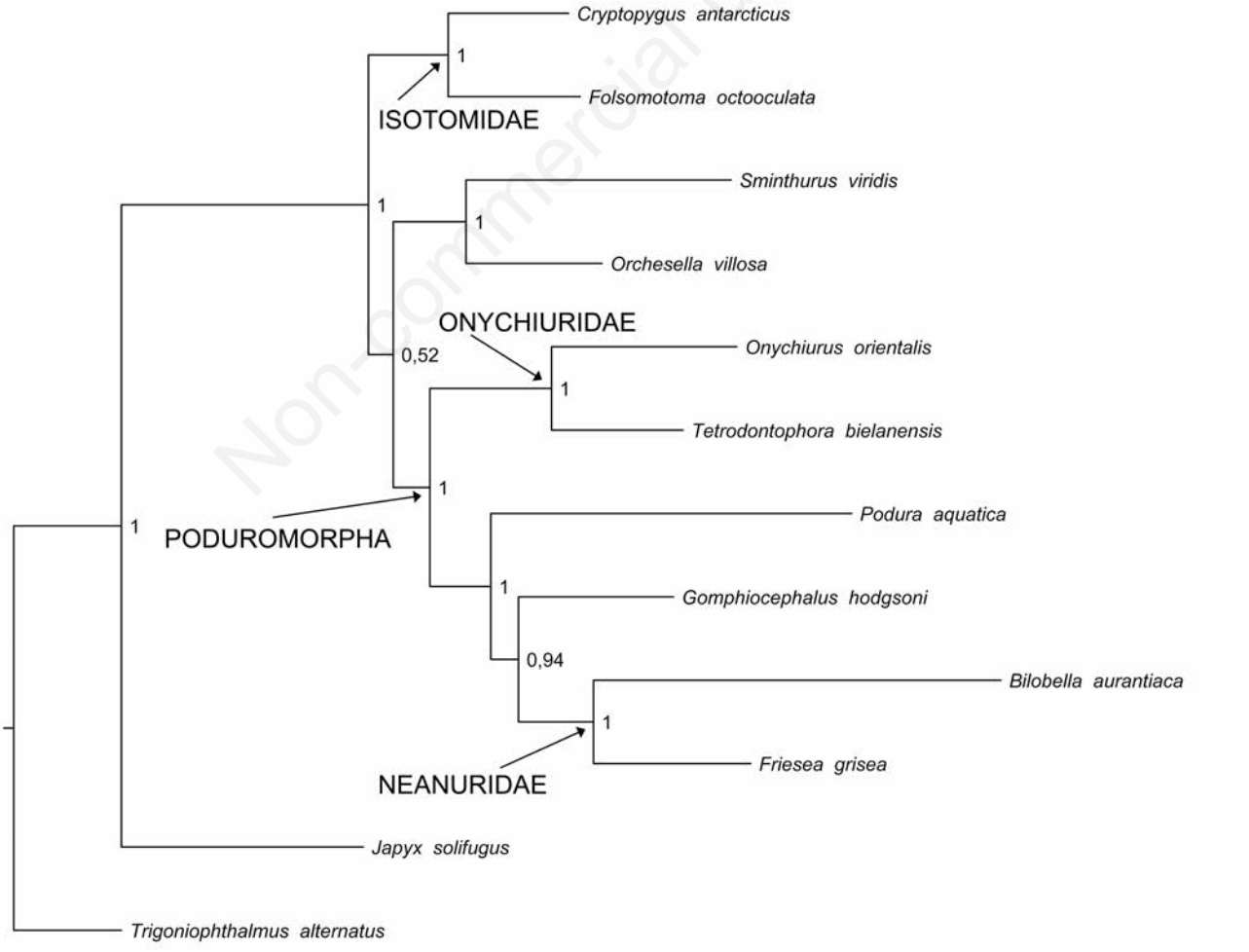

Figure 4. Phylogenetic tree of collembolan taxa, based on amino acid positions (aa-tree) of 13 mitochondrial protein-encoding genes.

our phylogenetic analysis provides some preliminary results (Figures 3 and 4), mostly centered on the phylogenetic placement of: i) Symphypleona with respect to Arthropleona; ii) *P. aquatica*; iii) Hypogastruridae, Neanuridae, Onychiuridae and Poduridae, within Poduromorpha. In this phylogenetic analysis, Arthropleona and Entomobryomorpha are paraphyletic, given that the symphypleonan *S. viridis* is the sister-group of the entomobryid *O. villosa* and therefore nested within the Arthropleona (taxa with an elongated body). This view is not in agreement with traditional morphological studies, usually suggesting a more basal placement for Symphypleona, within the collembolan tree. According to this result, entomobryids and symphypleonans would share a more recent common ancestor than do isotomids with either of them. However, despite the different shape of the whole body, either Symphypleona, Neelipleona and Entomobryomorpha have reduced prothorax without setae. According to Janssens (http://www.collembola.org) these three orders should be placed into the superordinal taxon Neocollembola also for their shared location of springing muscles, which are situated in the III and IV abdominal segments. The presumed sister-group relationship between Symphypleona and Neelipleona cannot be rejected, given that no species from this latter group is yet available for comparison. Within Poduromorpha, the position of Poduridae has long been discussed. Poduridae has been in turn proposed as: i) the sister group of all other Poduromorpha (Börner, 1906; Cassagnau, 1971); ii) a well differentiated lineage that, according to some distinctive morphological characters regarding (hypognathous) head orientation and structure of mouthparts, should be placed with Actaletidae in the Metaxypleona (Salmonm, 1964); iii) the basal taxon to all collembolans (Yosii, 1961).

In this study, the Poduromorpha are represented as a monophyletic lineage, in agreement with other molecular and/or morphological data (D'Haese, 2002, 2003; Xiong *et al*., 2008). The Onychiuridae are the basal group of Poduromorpha, whereas Poduridae is not a basal group (neither for all collembolan, nor for Poduromorpha), but it rather appears as a clade well nested within Poduromorpha (Figures 3 and 4). The observed phylogenetic grouping [Onychiuridae, (Hypogastruridae, (Poduridae, Neanuridae))], obtained in the nu-t, is in strong agreement with the morphological analysis by D'Haese (2003) (although in this latter study Hypogastruridae and Onychiuridae are retrieved as polyphyletic) and with rDNA-based analyses (D'Haese, 2002; Luan *et al*., 2005). The tree based on amino acid sequences (Figure 4) provides a different arrangement for Poduromorpha, with Onychiuridae resulting still the most basal lineage, but this time with Poduridae outside the cluster Hypogastruridae+Neanuridae (Figure 4). Discrepancies between nucleotide and amino acid based analyses may be resolved with a denser taxon sampling for both Hypogastruridae and Neanuridae.

Conclusions

In summary, our study adds to the several attempts, which have been recently performed, to elucidate internal relationships within Collembola using different and innovative sets of morphological (D'Haese, 2003) and molecular data sets (Luan *et al*., 2005; Gao *et al*., 2008; Xiong *et al*., 2008; Schneider *et al*., 2011). Despite these attempts, some internal relationships are still disputable and require additional investigation. Ribosomal-DNA-based studies suggest a basal position for the monophyletic (Xiong *et al*., 2008) or paraphyletic Symphypleona, with some species of this latter group more closely related to Entomobryomorpha+Poduromorpha, than to other symphypleonans (D'Haese, 2002; Luan *et al*., 2005). Mitogenomic data differ from nuclear ones with respect to the position of Symphypleona that appear nested within Entomobryomorpha.

Large agreement between different molecular phylogenetic studies is recovered for some internal relationships of Poduromorpha, although with Hypogastruridae or Poduridae alternatively resulting as the actual sister-group of Neanuridae. Generally speaking, multiple nuclear markers can (or have been) devised that display different variability and/or informativeness at different taxonomic levels and the use of multiple unlinked markers would allow for the use of comparative methods in phylogenetic analyses. As such, nuclear comparative genomic approaches are likely to provide more solid reconstructions in the near future. However, mitochondrial data have proven to be efficient in recovering solid information at least at some phylogenetic levels, apparently performing as well as nuclear data for the reconstruction of internal relationships of the collembolan lineages, while they appear to be less effective in resolving basalmost nodes. Furthermore, due to the long time necessary to gather new complete mtDNA sequences, the present sampling of collembolan taxa does not match, in terms of sampling density, the data sets obtained for nuclear DNA markers, although these latter are usually confined to just one or a few genes. However, the improvement of methods of sequencing mitochondrial DNAs will provide, in a relatively short time, a better taxonomic coverage for most of the families of the group. This is particularly necessary for the order Neelipleona, a key group of the collembolan tree, which is still missing in the mitogenomic data set. Relevant improvements on taxon density are therefore expected in the near future to help providing a thorough comparison with nuclear and morphological data aimed at drawing the most complete picture of the phylogenetic relationships within Collembola.

References

ADACHI J., HASEGAWA M., 1996 - MOLPHY version 2.3: programs for molecular phylogenetics based on maximum likelihood. Available from: http://www.researchgate.net/profile/Masami_Hasegawa/publication/228466208_MOLPHY_version_2.3_programs_for_molecular_phylogenetics_based_on_maximum_likelihood/file/9fcfd4ffe29b5b417a.pdf

BOORE J.L., LAVROV D.V., BROWN W.M., 1998 - Gene translocation links insects and crustaceans. - Nature 392: 667-668.

BOORE J.L., 1999 - Animal mitochondrial genomes. Nucleic Acids Res 27: 1767-1780.

BoRNER C., 1906 - Das system der Collembolen, nebst beschreibung neuer Collembolen des Hamburger naturhistorischen museums. - Mitt. Naturhist. Mus. Hamburg 23: 147-188.

BURGER G., GRAY M.W., FORGET L., FRANZ LANG B., 2013. Strikingly bacteria-like and gene-rich mitochondrial genomes throughout Jakobid protists. - Genome Biol. Evol. 5: 418-438.

CARAPELLI A., LIO' P., NARDI F., VAN DER WATH E., FRATI F., 2007 - Phylogenetic analysis of mitochondrial protein coding genes confirms the reciprocal paraphyly of Hexapoda and Crustacea. - BMC Evol. Biol. 7(Suppl 2):S8.

CARAPELLI A., COMANDI S., CONVEY P., NARDI F., FRATI F., 2008 - The complete mitochondrial genome of the Antarctic springtail *Cryptopygus antarcticus* (Hexapoda: Collembola). - BMC Genomics 9: 315.

CARAPELLI A., TORRICELLI G., NARDI F., FRATI F., 2013 - The complete mitochondrial genome of the Antarctic sea spider *Ammothea carolinensis* (Chelicerata; Pycnogonida). - Polar Biol. 36: 593-602.

CASSAGNAU P., 1971 - La phylogenie des Collemboles a la lumiere des structures endocrines retrocerebrales. In: Universidad de Salamanca, Facultad de Ciencias (ed). I Symposio Internacional Zoolfilogenia, Salamanca, pp 333-349.

CHATEIGNER-BOUTIN A.L., SMALL I., 2010. - Plant RNA editing. - RNA Biol. 7: 213-219.

D'HAESE C., 2002 - Were the first springtails semi-aquatic? A phylogenetic approach by means of 28S rDNA and optimization alignment. - Proc. R. Soc. B 269: 1143-1151.

D'HAESE C., 2003 - Morphological appraisal of Collembola phylogeny with special emphasis on Poduromorpha and a test of the aquatic origin hypothesis. - Zool. Scripta 32: 563-586.

FERNANDEZ-SILVA P., LOGUERCIO POLOSA P., ROBERTI M., DI PONZIO B., GADALETA M.N., MONTOYA J., CANTATORE P., 2001 - Sea urchin mtDBP is a two- faced transcription termination factor with a biased polarity depending on the RNA polymerase. - Nucleic Acids Res. 29: 4736-4743.

GAO Y., BU Y., LUAN Y.X., 2008. - Phylogenetic relationships of basal hexapods reconstructed from nearly complete 18S and 28S rRNA gene sequences. - Zool. Sci. 25: 1139-1145.

GLENNER H., THOMSEN P.F., HEBSGAARD M.B., SØRENSEN M.V., WILLERSLEV E., 2006 - The origin of insects. - Science 314: 1883-1884.

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.

HENNIG W., 1981 - Insect phylogeny. John Wiley and Sons (eds), New York.

HIRST S., MAULIK S., 1926 - On some arthropod remains from the Rhynie chert (Old Red Sandstone). - Geol. Mag. 63: 69-71.

JANSSENS F., CHRISTIANSEN K.A., 2011 - Animal biodiversity: an outline of higher-level classification and survey of taxonomic richness. - Zootaxa 3148: 192-194.

KOLPAKOV R., BANA G., KUCHEROV G., 2003 - mreps: efficient and flexible detection of tandem repeats in DNA. - Nucleic Acids Res. 31: 3672-3678-

KRISTENSEN N.P., 1981 - Phylogeny of insect orders. - Ann. Rev. Entomol. 26: 135-157.

KRISTENSEN N.P., 1998 - The groundplan and basal diversification of the hexapods. In: FORTEY R.A., THOMAS R.H., (eds). Arthropod relationships, systematic association. Vol. 55. Chapman and Hall, London, pp 281-293.

KRZYWINSKI J., GRUSHKO O.G., BESANSKY N.J., 2006 - Analysis of the complete mitochondrial DNA from *Anopheles funestus*: an improved dipteran mitochondrial genome annotation and a temporal dimension of mosquito evolution. - Mol. Phylogenet. Evol. 39: 417-423.

LANFEAR R., CALCOTT B., HO S.Y.W., GUINDON S., 2012 - PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. - Mol. Biol. Evol. 29: 1695-1701.

LASLETT D., CANBÄCK B., 2008 - ARWEN, a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. - Bioinformatics 24: 172-175.

LAVROV D.V., BOORE J.L., BROWN W.M., 2000 - The complete mitochondrial DNA sequence of the horseshoe crab *Limulus polyphemus*. - Mol. Biol. Evol. 17: 813-824.

LAVROV V.L., LANG B.F., 2005 - Poriferan mtDNA and animal phylogeny based on mitochondrial gene arragements. - System. Biol. 54: 651-659.

LITTLE C., 1990 - The terrestrial invasion: an ecophysiological approach to the origins of land animals. Cambridge University Press, Cambridge.

LUAN Y.X., MALLATT J.M., XIE R.D., YANG Y.M., YIN W.Y., 2005 - The phylogenetic positions of three basal-hexapod groups (Protura, Diplura, and Collembola) based on ribosomal RNA gene sequences. - Mol. Biol. Evol. 22: 1579-1592.

MILLER M.A., HOLDER M.T., VOS R., MIDFORD P.E., LIEBOWITZ T., CHAN L., HOOVER P., WARNOW T., 2010 - The CIPRES portals. In: CIPRES, vol. 2010.

OJALA D., MONTOYA J., ATTARDI G., 1981 - tRNA punctuation model of RNA processing in human mitochondria. - Nature 290: 470-474.

PEDEN J.F., 1999 - Analysis of codon usage. Dissertation. University of Nottingham, UK.

PERNA N.T., KOCHER T.D., 1995 - Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. J. Mol. Evol. 41: 353-358.

PRICE D.H., GRAY M.W., 1999 - A novel nucleotide incorporation activity implicated in the editing of mitochondrial transfer RNAs in *Acanthamoeba castellanii*. - RNA 5: 302-317.

ROBERTI M., BRUNI F., LOGUERCIO POLOSA P., GADALETA M.N., CANTATORE P., 2006 - The *Drosophila* termination factor DmTTF regulates in vivo mitochondrial transcription. - Nucleic Acids Res. 34: 2109-2116.

RONQUIST F., HUELSENBECK J.P., 2003 - MRBAYES 3: bayesian phylogenetic inference under mixed models. - Bioinformatics 19: 1572-1574.

SALMON J.T., 1964 - An index to the Collembola. I. - R. Soc. New Zeal. Bull. 7: 1-144.

SCHNEIDER C., CRUAUD C., D'HAESE C.A., 2011 - Unexpected diversity in Neelipleona revealed by molecular phylogeny approach (Hexapoda, Collembola). - Soil Organ. 83: 383-398.

SIMON C., FRATI F., BECKENBACH A., CRESPI B., LIU H., FLOOK P., 1994 - Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. - Ann. Entomol. Soc. Am. 87: 651-701.

TORRICELLI G., CARAPELLI A., CONVEY P., NARDI F., BOORE J.L., FRATI F., 2010 - High divergence across the whole mitochondrial genome in the pan-Antarctic springtail *Friesea grisea*: evidence for cryptic species? - Gene 449: 30-40.

WERNERSSON R., PEDERSEN A.G., 2003 - RevTrans: constructing alignments of coding DNA from aligned amino acid sequences. - Nucleic Acids Res. 31: 3537-3539.

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.

XIONG Y., GAO Y., YIN W.Y., LUAN Y.X., 2008 - Molecular phylogeny of Collembola inferred from ribosomal RNA genes. - Mol. Phylogenet. Evol. 49: 728-735.

YOSII R., 1961 - Phylogenetische bedeuntung der chaetotaxiee bei den Collembolen. Contributions from the Biological Laboratory, Kyoto University, 12: 1-37.