SOIL ORGANISMS Volume 83 (3) 2011 pp. 383–398

ISSN: 1864-6417

Unexpected diversity in Neelipleona revealed by molecular phylogeny approach (Hexapoda, Collembola)

Clément Schneider[1, 3], Corinne Cruaud[2] and Cyrille A. D'Haese[1]

[1] *UMR7205 CNRS, Département Systématique et Évolution, Muséum National d'Histoire Naturelle, CP50 Entomology, 45 rue Buffon, 75231 Paris cedex 05, France*

[2] *Genoscope, Centre National de Sequençage, 2 rue G. Crémieux, CP5706, 91057 Evry cedex, France*

[3] *Corresponding author: Clément Schneider (email: clement.schneider@mnhn.fr)*

Abstract

Neelipleona are the smallest of the four Collembola orders in term of species number with 35 species described worldwide (out of around 8000 known Collembola). Despite this apparent poor diversity, Neelipleona have a worldwide repartition. The fact that the most commonly observed species, *Neelus murinus* Folsom, 1896 and *Megalothorax minimus* Willem, 1900, display cosmopolitan repartition is striking. A cladistic analysis based on 16S rDNA, COX1 and 28S rDNA D1 and D2 regions, for a broad collembolan sampling was performed. This analysis included 24 representatives of the Neelipleona genera *Neelus* Folsom, 1896 and *Megalothorax* Willem, 1900 from various regions. The interpretation of the phylogenetic pattern and number of transformations (branch length) indicates that Neelipleona are more diverse than previously thought, with probably many species yet to be discovered. These results buttress the rank of Neelipleona as a whole order instead of a Symphypleona family.

Keywords: Collembola, Neelidae, *Megalothorax*, *Neelus*, COX1, 16S, 28S

1. Introduction

1.1. Brief history of Neelipleona classification

The Neelidae family was established by Folsom (1896), who described *Neelus murinus* from Cambridge (USA). Neelidae were traditionally considered close relatives of Symphypleona *sensu stricto* (Börner 1906, Salmon 1964). Massoud (1971) placed Neelidae in its own sub-order, Neelipleona. Neelipleona were later elevated to the order level (Massoud 1976). Massoud considered that the specific morphological characters of the Neelipleona clearly distinguished the group from the three pre-recognised orders of Collembola (Poduromorpha, Entomobryomorpha and Symphypleona). However, Bretfeld proposed a classification of Symphypleona *sensu lato* where Neelidae would be the sister family of Symphypleona s. str. (Bretfeld 1986). While Bretfeld did not conduct a formal phylogenetic analysis, he inscribed his vision of Symphypleona in the cladistic paradigm, providing characters assumed to be apomorphic to support the classification. While the classifications proposed by Hopkin (1997) and Bretfeld (1999) assumed Neelidae as a Symphypleona family, Christian (1987) agreed

with Massoud to consider Neelidae as the monotypic family of Neelipleona. Neelipleona are now classified as an order in several recent taxonomic and phylogenetic publications (e.g. D'Haese 2003a, Deharveng 2004, Janssens 2009, Xiong et al. 2008).

The close relationship between Neelipleona (i.e. Neelidae) and Symphypleona was supported by many specialists, based on morphological similarities: globular shape, neosminthuroid setae, mucro gutter-like (Bretfeld 1986). This hypothesis was supported by D'Haese's (2003a) morphological cladistic analysis, which uncovered Neelipleona as sister group of Symphypleona. Nevertheless, Janssens (2009) proposed Neelipleona to be closer to Entomobryomorpha, based on the interpretation of some characters such as the form of the retinaculum, the number of setae in anterior setal row of the clypeus, some aspects of the postembryonal development and the structure of the genital plate.

The phylogenetic position of Neelipleona was first inferred from morphological data by D'Haese (2003a, 2003b) and then from molecular data in Xiong et al. (2008) and Gao et al. (2008), based on 28S rDNA and 18S rDNA loci. Neelipleona were also represented in Dell'Ampio et al. (2009) molecular phylogeny of Hexapoda, based on 28S rDNA and in von Reumont et al. (2009) molecular phylogeny of Arthropoda, based on 18S and 28S rDNA.

1.2. Neelipleona diversity

To date, Neelipleona are represented by 35 species included in 5 genera: *Acanthothorax* Bretfeld and Griegel, 1999 (1 sp), *Megalothorax* Willem, 1900 (22 spp), *Neelides* Caroli, 1912 (5 spp), *Neelus* Folsom, 1896 (6 spp) and *Zelandothorax* Delamare Deboutteville and Massoud, 1963 (1 sp). This is very poor in regards to the other Collembola orders, each including thousands of species.

The two first species described, *Neelus murinus* Folsom, 1896 and *Megalothorax minimus* Willem, 1900 present a cosmopolitan distribution (Bretfeld 1999). *Megalothorax minimus* is especially common. For example, the presence of *Megalothorax minimus* has been reported from France (Loranger et al. 2001), Latvia (Jucevič a 2003), Israel (Bretfeld et al. 2000), Cuba (Azpiazu et al. 2004), Canada (Chagnon et al. 2000), Mexico (García-Gómez et al. 2009), among other places. Bretfeld (1999) suggested that a re-examination of different populations of *M. minimus* would be required. Cryptic diversification is known to occur within Collembola even in parthenogenetic species, like *Folsomia candida* Willem, 1902 (Deharveng 2004). It is possible that complexes of cryptic species are hiding in *Megalothorax minimus*, and more generally in Neelipleona.

In the present contribution, we intend to investigate the phylogenetic position of Neelipleona among Collembola and to explore Neelipleona specific diversity using four different loci: D1 and D2 regions of the nuclear 28S rDNA gene and the mitochondrial 16S rDNA and COX1 genes for a broad taxon sampling.

2. Materials and methods

2.1. Taxonomic sampling

The taxon sampling is based on 27 collembolan taxa selected from D'Haese (2002) completed with 2 Symphypleona from GenBank (*Ptenothrix monochroma* and *Bourletiella hortensis*) and 24 Neelipleona specimens sequenced for the present study (Tab. 1). The Decapoda *Euastacus bispinosus* Clark, 1936 and *Geocharax gracilis* Clark, 1936 (Parastacidae) were selected from GenBank as remote outgroups. *Thermobia domestica* (Packard, 1837) (Zygentoma: Lepismatidae) and *Petrobius brevistylis* Carpenter, 1913 (Archaeognatha: Machilidae) were

selected as closer outgroups to Collembola. All four collembolan orders Entomobryomorpha, Poduromorpha, Symphypleona and Neelipleona were sampled. Symphypleona are represented by the families Katiannidae, Arrhopalitidae, Bourletiellidae, Sminthuridae and Dicyrtomidae (9 operational taxonomic units [OTUs] in total). Poduromorpha are represented by Odontellidae, Onychiuridae, Hypogastruridae, Poduridae and Neanuridae (11 OTUs in total). Entomobryomorpha are represented by Entomobryidae and Isotomidae families (7 OTUs in total) (Tab. 1). We added to this sample 24 Neelipleona specimens collected from various regions (Tab. 1). Samples were collected from soil or leaf litter samples, extracted with Berlese-Tullgren funnels and stored in 95% ethanol at 5 °C until genetic and morphological analyses could be carried on. Due to the especially minute size and the absence of colour of the specimens, the voucher (whole cuticle) could not be recovered after DNA extraction. In order to keep a link between morphological specimen and molecular data, we assume that specimens from the same sample (e.g. collected from a given soil or leaf litter sample) and morphologically homogeneous are members of the same population. While a picture of the specimen is taken before DNA extraction, a couple of 'sister' specimens are mounted in Marc André II mounting medium for identification under the compound microscope. Each time, DNA from several specimens was extracted and amplified separately to check the samples' genetic homogeneity (since Neelipleona are parthenogenetic, specimens from the same population should be genetically identical). Specimens were determined to genus level using the identification keys of Bretfeld (1999) and Kováč & Papáč (2010). We included in our dataset three *Neelus* specimens and 21 *Megalothorax* specimens, from which we could obtain full data for the wanted loci. *Neelus koseli* Kováč & Papáč 2010 (Slovakia) was provided by the authors from the type locality, while the two other *Neelus* species, respectively collected in a garden in Chelles (France) and in a underground carrier in Paris, appear to be undescribed species.

Megalothorax specimens were collected from France, Argentina, French Guiana, Belgium, USA (Ohio) and Chile. Morphological scrutiny shows strong affinity either with *Megalothorax minimus* either with *Megalothorax incertus* with slight variations.

Unfortunately, we lacked representatives of the three other Neelipleona genera. Gao et al. (2008) used an overlapping 28S rDNA sequence of *Neelides minutus* Caroli, 1912 (GenBank number: EF422366), but our preliminary analysis seems to indicate a misidentification, it would rather be the sequence of an undetermined *Megalothorax* species.

2.2. Molecular data

In addition to the data available from D'Haese (2002), we gathered D1 and D2 for all new OTUs (respectively ~375-bp and ~425-bp) (Tab. 1). A ~512-bp region of 16S rDNA and a 658-bp region of mitochondrial COX1 (barcode) were sequenced (from the original DNA material for taxa from D'Haese (2002)). Additional 16S rDNA and COX1 sequences were added from Greenslade et al. (2011) (Tab.1). Genomic DNA was extracted with the DNeasy© kit (Quiagen S.A.S., Courtaboeuf, France), using the standard protocol. The 28S D1 and D2 regions were PCR-amplified in one single fragment using the primer pairs C1p 5'-CCCGCTGAATTTAAGCAT-3' / D2 5'-TCCGTGTTTCAAGACGGG-3' (D'Haese 2002). The primer pair for 16S was specifically designed for Collembola for this work: 16SAcoll 5' – MGM MTG TTT AWC AAA AAC AT – 3' / 16SBcoll 5' – CGC CGG TTT GAA CTC AAA TCA – 3'. Primers for COX1 were also specifically designed for Collembola (Greenslade et al. 2011, Palacios-Vargas et al. 2011): LCO1490COL 5' – WYT CDA CWA AYC RYA ARG AYA TYG G – 3' / HCO2198COL 5' – TAN ACY TCN GGR TGN CCR AAR AAT CA – 3'.

Tab. 1 Taxonomic sample. Geographical origin of Neelipleona specimens is provided. GenBank accession number provided for all sequences available for the analysis. Missing data are indicated with a '-'. [(1)] Data from GenBank, used in D'Haese (2002). [(2)] Data from GenBank, used in Greenslade et al. (2011). * Data from GenBank, other origin. All other data are new for the present study. A '+' before the specimen ID indicates that this specimen is also represented in the reduced analysis.

	Specimen ID	Taxonomy	GenBank accession				Geographical origin
			COX1	16S	28S D1	28S D2	
		Decapoda (Malacostraca)					
+	*Euastacus bispinosus*	*Euastacus bispinosus*	DQ006317*	AF492813*	AF235981*	AF235981*	
+	*Geocharax gracilis*	*Geocharax gracilis*	EU921145*	AF235992*	AF235982*	AF235982*	
		Zygentoma (Insecta)					
+	*Thermobia domestica*	*Thermobia domestica*	JN970940	JN970984	AF483404 [(1)]	AF483462 [(1)]	
		Archaeognatha (Insecta)					
+	*Petrobius brevistylis*	*Petrobius brevistylis*	JN970936	JN970976	AF483389 [(1)]	AF483447 [(1)]	
		Entomobryomorpha (Collembola)					
		Entomobryidae					
+	*Entomobrya lanuginosa*	*Entomobrya lanuginosa*	JN970907	JN970943	AF483365 [(1)]	AF483423 [(1)]	
+	*Pseudosinella sp*	*Pseudosinella sp.*	JN970937	JN970979	AF483393 [(1)]	AF483451 [(1)]	
+	*Orchesella cincta*	*Orchesella cincta*	JN970933	JN970973	AF483385 [(1)]	AF483443 [(1)]	
+	*Orchesella villosa*	*Orchesella villosa*	JN970934	JN970974	AF483386 [(1)]	AF483444 [(1)]	
		Isotomidae					
+	*Folsomia candida*	*Folsomia candida*	HQ732049 [(2)]	JN970944	AF483366 [(1)]	AF483424 [(1)]	
+	*Isotoma viridis*	*Isotoma viridis*	JN970908	JN970946	AF483372 [(1)]	AF483430 [(1)]	
	Parisotoma notabilis	*Parisotoma notabilis*	JN970935	JN970975	AF483388 [(1)]	AF483446 [(1)]	
		Poduromorpha (Collembola)					
		Odontellidae					
	Superodontella gisini	*Superodontella gisini*	-	-	AF483401 [(1)]	AF483459 [(1)]	
	Superodontella alpina	*Superodontella alpina*	-	-	AF483400 [(1)]	AF483458 [(1)]	

		Neanuridae				
+	*Anurida maritima*	*Anurida maritima*	HQ732028 (2)	JN970941	AF483357 (1)	AF483415 (1)
+	*Neanura muscorum*	*Neanura muscorum*	HQ732074 (2)	JN970969	AF483382 (1)	AF483440 (1)
		Poduridae				
+	*Podura aquatica*	*Podura aquatica*	HQ732077 (2)	JN970977	AF483390 (1)	AF483448 (1)
		Hypogastruridae				
+	*Hypogastrura vernalis*	*Hypogastrura vernalis*	HQ732065 (2)	JN970945	AF483371 (1)	AF483429 (1)
+	*Ceratophysella gibbosa*	*Ceratophysella gibbosa*	HQ732039 (2)	JN970942	AF483362 (1)	AF483420 (1)
+	*Schoettella ununguiculata*	*Schoettella ununguiculata*	HQ732079 (2)	JN970980	AF483395 (1)	AF483453 (1)
		Onychiuridae				
+	*Tetrodontophora bielanensis*	*Tetrodontophora bielanensis*	HQ732082 (2)	JN970983	AF483402 (1)	AF483460 (1)
+	*Kalaphorura paradoxa*	*Kalaphorura paradoxa*	HQ732071 (2)	JN970947	AF483373 (1)	AF483431 (1)
+	*Protaphorura armata*	*Protaphorura armata*	HQ732078 (2)	JN970978	AF483391 (1)	AF483449 (1)
		Symphypleona (Collembola)				
		Katiannidae				
	Arrhopalites sericus	*Arrhopalites sericus*	-	-	AF483359 (1)	AF483417 (1)
+	*Sminthurinus bimaculatus*	*Sminthurinus bimaculatus*	JN970938	JN970981	AF483398 (1)	AF483456 (1)
		Dicyrtomidae				
	Ptenothrix sp	*Ptenothrix* sp.	-	-	AF483394 (1)	AF483452 (1)
	Ptenothrix monochroma	*Ptenothrix monochroma*	-	-	FJ411425*	FJ411425*
	Dicyrtoma sp	*Dicyrtoma* sp.	-	-	AF483364 (1)	AF483422 (1)
		Bourletiellidae				
	Bourletiella hortensis	*Bourletiella hortensis*	-	-	FJ411426*	FJ411426*
		Sminthuridae				
	Allacma fusca	*Allacma fusca*	-	-	AF483355 (1)	AF483413 (1)
+	*Sminthurus viridis*	*Sminthurus viridis*	JN970939	JN970982	AF483399 (1)	AF483457 (1)
	Caprainea marginata	*Caprainea marginata*	-	-	AF483361 (1)	AF483419 (1)

Tab. 1 Taxonomic sample. (Continued previos page.)

	Specimen ID	Taxonomy	GenBank accession				Geographical origin
			COX1	16S	28S D1	28S D2	
		Neelipleona (Collembola)					
		Megalothorax					
+	*Megalothorax* GUF 1	*Megalothorax* sp.	JN970929	JN970968	JN971005	JN971005	French Guiana
	Megalothorax GUF 2	*Megalothorax* sp.	JN970928	JN970967	JN971004	JN971004	French Guiana
	Megalothorax GUF 3	*Megalothorax* sp.	JN970910	JN970949	JN970986	JN970986	French Guiana
+	*Megalothorax* ARG 1	*Megalothorax* sp.	JN970916	JN970955	JN970992	JN970992	Argentina
	Megalothorax ARG 2	*Megalothorax* sp.	JN970926	JN970965	JN971002	JN971002	Argentina
+	*Megalothorax* ARG 3	*Megalothorax* sp.	JN970918	JN970957	JN970994	JN970994	Argentina
	Megalothorax ARG 4	*Megalothorax* sp.	JN970919	JN970958	JN970995	JN970995	Argentina
+	*Megalothorax* CHL 1	*Megalothorax* sp.	JN970927	JN970966	JN971003	JN971003	Chile
	Megalothorax FRA 1	*Megalothorax* sp.	JN970912	JN970951	JN970988	JN970988	France (Brunoy)
	Megalothorax FRA 2	*Megalothorax* sp.	JN970917	JN970956	JN970993	JN970993	France (Brunoy)
	Megalothorax FRA 3	*Megalothorax* sp.	JN970911	JN970950	JN970987	JN970987	France (Brunoy)
+	*Megalothorax* FRA 4	*Megalothorax* sp.	JN970920	JN970959	JN970996	JN970996	France (Brunoy)
	Megalothorax FRA 5	*Megalothorax* sp.	JN970921	JN970960	JN970997	JN970997	France (Brunoy)
	Megalothorax FRA 6	*Megalothorax* sp.	JN970922	JN970961	JN970998	JN970998	France (Brunoy)
	Megalothorax FRA 7	*Megalothorax* sp.	JN970915	JN970954	JN970991	JN970991	France (Brunoy)
	Megalothorax FRA 8	*Megalothorax* sp.	JN970913	JN970952	JN970989	JN970989	France (Chelles)
	Megalothorax FRA 9	*Megalothorax* sp.	JN970924	JN970963	JN971000	JN971000	France (Labosse)
+	*Megalothorax* FRA 10	*Megalothorax* sp.	JN970923	JN970962	JN970999	JN970999	France (Labosse)
	Megalothorax FRA 11	*Megalothorax minimus*	JN970914	JN970953	JN970990	JN970990	France (Rozel)
+	*Megalothorax* BEL 1	*Megalothorax* sp.	JN970925	JN970964	JN971001	JN971001	Belgium
+	*Megalothorax* USA 1	*Megalothorax minimus*	JN970909	JN970948	JN970985	JN970985	USA (Ohio)
		Neelus					
+	*Neelus koseli*	*Neelus koseli*	JN970932	JN970972	JN971008	JN971008	Slovakia
+	*Neelus* FRA 1	*Neelus* sp.	JN970930	JN970970	JN971006	JN971006	France (Chelles)
	Neelus FRA 2	*Neelus* sp.	JN970931	JN970971	JN971007	JN971007	France (Paris)

The PCR reactions were performed in 25 µl volume: 5 µl of TAQ&Load, 2µl of total DNA extract for Collembola or 5 µl for Neelipleona (because the minute specimens contain little DNA), 0.5 µl of each primer at 25 µM and complement of H2O. The PCR cycles for 16 S and 28 S loci consisted of an initial denaturing step at 94 °C for 1 min, 40 amplification cycles (94 °C for 30 s, 52 °C for 45 s, 72 °C for 1.5 min), and a final step at 72 °C for 5 min. The PCR cycles for COX1 consisted of an initial denaturing step at 94 °C for 1 min, 5 amplification cycles (94 °C for 40 s, 45 °C for 40 s, 72 °C for 1 min), followed by 35 cycles with an annealing temperature of 51 °C and a final step at 72 °C for 5 min.

PCR products were sequenced at the Genoscope Centre National de Sequençage (Evry, France). Resulting chromatograms were interpreted using the program Sequencher® (Gene Codes Corporation, Ann Arbor, USA). Sequences were manually checked, trimmed and cleaned when necessary. All sequences were checked as being close to Collembola using GenBank BLAST algorithm. Unfortunately, COX1 and 16S sequences could not be retrieved for 9 specimens (Tab. 1).

2.3. Phylogenetic analysis

Phylogenetic analyses were performed using direct optimisation method (Wheeler 1996) using the parsimony criterion as implemented in POY version 4.1.2.1 (Varón et al. 2010). We performed the analyses for each marker independently and combined analyses. Each analysis was run for 12 different transformation cost regimes to test the stability of the results. The detailed procedure used for each analysis, using various search algorithms implemented in POY, was as follows: a starting pool of 1000 Wagner trees was generated through random addition sequence (RAS). Each replicate was explored by a combination of TBR and SPR branch swapping (POY default). The 1000 resulting topologies were then explored by tree-fusing, and 20 optimal or sub-optimal trees were retained. Those 20 topologies were submitted to ratchet and then to tree-fusing again, the resulting 20 optimal or sub-optimal being retained. A final and more thorough branch-swapping was performed and the optimal trees were retained as final results of the analysis. The authors of the cited algorithms are: Swofford (1990) for TBR and SPR, Goloboff (1999) for tree-fusing and Nixon (1999) for ratchet.

Jackknife nodes' support was calculated by resampling the aligned data (homology hypothesis implied for the optimal topology) in 10000 iterations. For each iteration, 36% of characters were randomly removed, 10 Wagner trees were built and swapped using TBR (POY default).

2.4. Sensitivity analysis

Parameter sets (sankoff matrices) consist of two transformation cost variables: gap/transversion ratio and transversion/transition ratio. Since the parameter space is too wide to be explored exhaustively in reasonable time limit, 12 parameter sets for witch incongruence was minimised among data partitions were chosen among the 24 parameter sets explored in a previous molecular phylogeny (D'Haese 2002). We used congruence as the optimality criterion to make a non-arbitrary choice of the parameter sets that better describe the data. The optimal parameter set is the one that minimises incongruence among the 4 regions analyzed independently: D1 and D2 of 28S rDNA, 16S and COX1. Congruence was measured by the ILD metrics (Mickevich & Farris 1981, Wheeler 1999). The ILD is the ratio between (1) the difference of combined data tree length and the sum of individual data tree length, and (2) the combined data tree length:

$$ILD = (length_{combined} - \sum length_{individual\ sets}) / length_{combined}$$

(1) (2)

A total of 60 analyses were run on the Museum National d'Histoire Naturelle's mainframe cluster, each analyses performed by 10 parallelised CPUs. The topology that minimises the number of transformations under the optimal parameter set is retained as our best phylogeny hypothesis.

To evaluate the bias in branch length for the over-representation of Neelipleona in our dataset, a supplementary set of analyses was conducted, with the sampling trimmed of 14 Neelipleona (Tab. 1). We also removed the taxa with missing data (Tab. 1). This analysis was run following the same script previously used, under the equal weighting scheme.

3. Results

The ILD scores for all tested parameter sets along with each calculated tree length are given in Table 2. The equal weighting scheme (indel:Tv:Ts = 1:1:1) showed to be the model minimising incongruency among the four loci. The phylogenetic analysis yielded two trees of 9118 steps (CI 0.40 , RI 0.65). Strict consensus for these trees is provided in Figure 1 along with jackknife support. The cladogram shows monophyly of Collembola, Symphypleona, Poduromorpha, Entomobryomorpha and Neelipleona. Neelipleona is the sister group of Arthropleona (Poduromorpha, Entomobryomorpha). Symphypleona is the sister group of all other Collembola. Except for Dicyrtomidae, Neanuridae, Isotomidae and Entomobridae, every collembolan family represented by more than one exemplar is monophyletic. The Neelipleona genera *Neelus* and *Megalothorax* are monophyletic.

The stability of the main clades resolved in our phylogenetic analysis throughout the 12 investigated parameter sets is displayed in Figure 1 as non-interpolated Cartesian graphs of areas of the parameter space. Monophyly of Entomobryomorpha is recovered in 8 sets out of 12 analyses. Poduromorpha are recovered in 6 sets, Symphypleona in 7 sets. Neelipleona and the genera *Megalothorax* and *Neelus* are always recovered. The clade (Neelipleona, Poduromorpha, Entomobryomorpha) is recovered in 5 sets.

Tab. 2 Total number of evolutionary steps reported for each optimal tree inferred from the different parameter sets. The ILD values calculated from those costs for each parameter set are reported in the last column. Parameter sets are named in the first column as XYZ (X = gap cost value, Y = transversion cost value, Z = transition cost value). The lowest ILD value is found for equal weighting (111).

	COX1	16S	28S d1	28S d2	All data	ILD value
111	3311	2675	599	2351	9118	**0,01996052**
112	4600	3160	826	2983	11927	**0,03001593**
121	4599	3358	754	2807	11781	**0,02232408**
122	6097	3988	1023	3631	15156	**0,02751386**
211	3323	3057	651	2708	9974	**0,02356126**
212	4882	3859	941	3644	13703	**0,02751222**
221	4881	4381	868	3540	13981	**0,02224447**
241	7376	5744	1173	4490	19264	**0,02496885**
412	4896	4348	1034	4213	14928	**0,02927385**
421	4887	5076	964	4192	15519	**0,02577486**
441	7893	7653	1376	5799	23323	**0,02581143**
841	7905	9001	1577	7022	26318	**0,03089141**

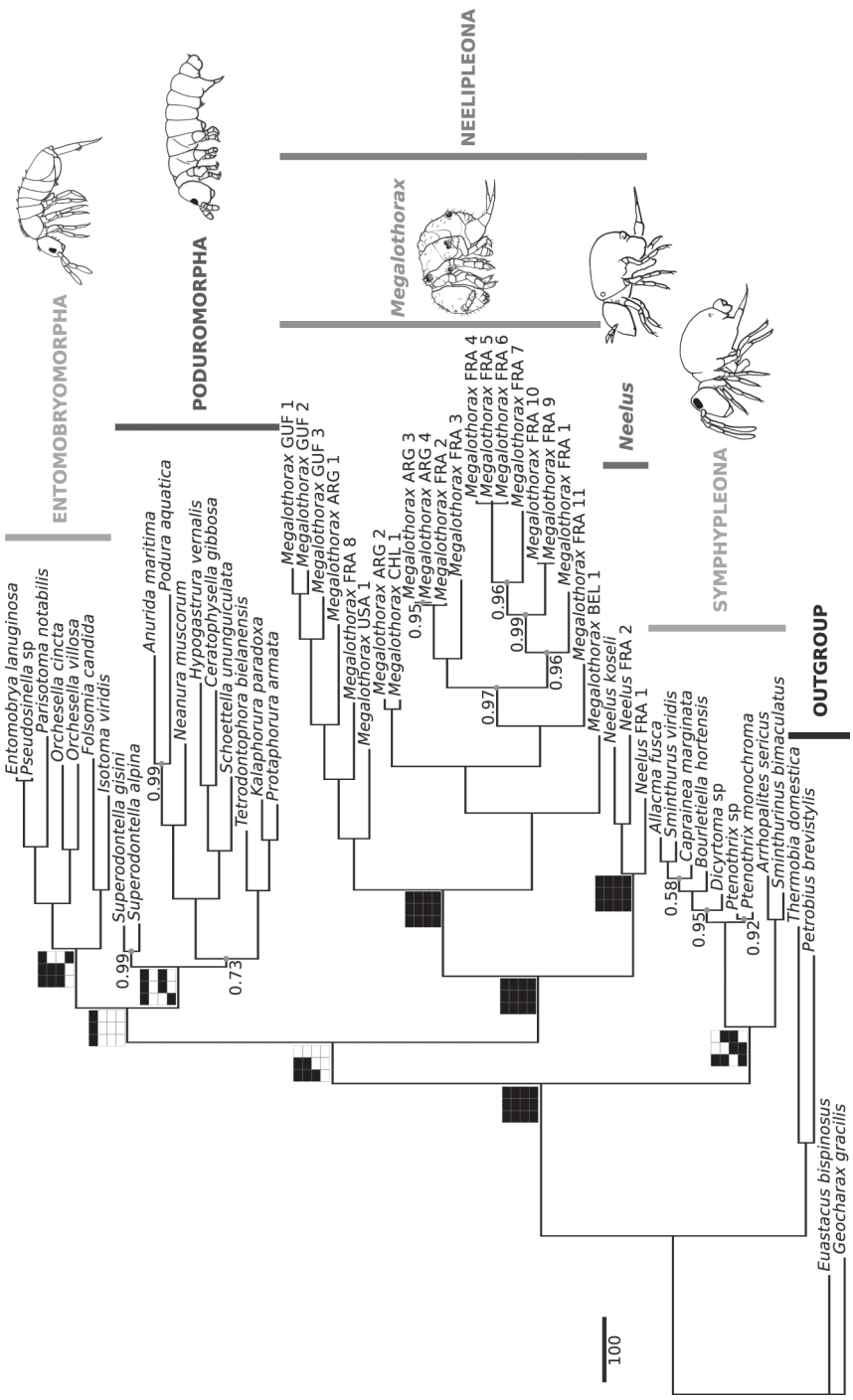

Fig. 1 Strict consensus of the two equally most parsimonious combined analysis cladograms using direct optimisation approach with equal weight of character transformation cost (indel = Ts = Tv = 1). Branches are proportional to the number of single transformation events (indel, Tv and Ts); scale represents 100 steps. Tree length = 9118 steps; consensus length = 9119 steps; RI = 0.40; CI = 0.65. Square diagrams represent the stability of the related clade throughout the different parameter sets (white = clade not recovered, black = clade recovered). Jackknife values below 1.00 are reported next to the related node (marked in grey). No value means 1.00 (i.e. recovered at 100%).

The reduced data set analysis yielded a tree of 7592 steps (CI 0.45, RI 0.55). Collembola, Neelipleona, Arthropleona, Entomobryomorpha and Poduromorpha are monophyletic, Symphypleona are paraphyletic. The schematic relationships between those taxa are: *Sminthurus viridis*, ((*Sminthurinus bimaculatus*, Neelipleona), (Poduromorpha, Entomobryomorpha)). In regards of the presented trees (Fig. 1), the relationships within Poduromorpha are unchanged, the position of the clade (*Orchesella cincta*, *O. villosa*) and (*Folsomia candida*, *Isotoma viridis*) is swapped and the positions of '*Megalothorax* FRA 8' and '*Megalothorax* USA 1' are swapped.

3.1. Geographical distribution and phylogenetic pattern within *Megalothorax* genus

The first dichotomy of the *Megalothorax* clade separates an American group (French Guiana, Chile, Argentina and USA) group from a group composed of European and some other South American specimens. The specimens from French Guiana formed a clade. The two Argentinean specimens did not group together. French specimens did not form a clade either. Two Argentinean specimens ('*Megalothorax* ARG 3' and 'ARG 4') are closer to some French specimens ('*Megalothorax* FRA 2' and 'FRA 3') than to other Argentinean or South American *Megalothorax*. Among the *Megalothorax* specimens captured in the garden of the Laboratoire d'Ecologie des Sols Tropicaux (Brunoy, France) ('*Megalothorax* FRA 1' to 'FRA 7' in Fig. 1 and Tab. 1), some showed to be nearly identical ('FRA 4', '5' and '6'), whereas some seem phylogenetically closer to Argentinean specimens ('FRA 2', 'FRA 3').

3.2. Branch length

Branch length here stands for the number of evolutionary steps separating each node and terminal in the present phylogeny. Because the equal weighting scheme has been selected, each indel, Tv and Ts event represents one evolutionary step.

Some Neelipleona OTUs are held by longer branches from the root than the other Collembola (e.g. '*Megalothorax* FRA 4', '5' and '6'). Transformations (indels, Tv and Ts) inside *Megalothorax* and *Neelus* genera are globally more numerous than in any other Collembola order. We provide here the number of transformations separating most distant OTUs in each order and in *Megalothorax* and *Neelus* genera. This distance is evaluated in priority with the phylogenetic pattern, and then with the number of single transformations separating the OTUs (the most distant relatives most transformed from a common ancestor) for ACCTRAN optimisation. *Sminthurus viridis* and *Sminthurinus bimaculatus* (Symphypleona) are separated by 489 transformations. 616 transformations separate *Parisotoma notabilis* and *Folsomia candida* (Entomobryomorpha). 801 transformations separate *Kalaphorura paradoxa* and *Podura aquatica* (Poduromorpha). 327 transformations separate *Neelus koseli* and '*Neelus* FRA 1' (Neelipleona). 1090 transformations separate '*Megalothorax* GUF 1' and *Megalothorax* 'FRA 6' (Neelipleona). Finally, 1100 transformations separate 'Neelus FRA 2' from '*Megalothorax* GUF 1' (Neelipleona). Figure 1 represents the cladogram with branches proportional to length, in ACCTRAN optimisation.

The analysis with a reduced taxon sampling shows comparable results. The most distant OTUs in Neelipleona are separated by 585 transformations, in *Neelus*: 216 transformations, in *Megalothorax:* 522 transformations, in Entomobryomorpha: 420 transformations and in Poduromorpha: 486 transformations. Since Symphypleona did not form a clade, they are not discussed here.

4. Discussion

4.1. Neelipleona Monophyly

The phylogenetic analysis results strongly support the monophyly of Neelipleona although the lack of data for the three remaining genera *Neelides*, *Acanthothorax* and *Zelandothorax* is a limitation to this conclusion. *Acanthothorax* and *Zelandothorax* are monospecific genera. *Neelides* is a more important genus within Neelipleona in term of occurrence and of diversity (with five species described). Monophyly of Neelipleona was never questioned since the discovery of the group. Neelipleona are indeed characterised by a set of unique morphological characters, supposedly apomorphic: a globular shape formed largely by expansion of the thoracic segments rather than abdominal ones as it is observed in Symphypleona; coxae and subcoxae well developed, coxae being larger than trochanters; special sensilla in sensory fields on the head and abdomen; epicuticula scales off, i.e. desquamation (as opposed to the classical tegumentary granules in other Collembola) (Massoud 1971, Dallai 1979); midgut with 4 diverticula; dentes subdivided. Other characters shared by Neelipleona include the absence of eyes (also in many other Collembola groups), the special structure and ultrastructure of the antennae, the secretion of wax rods (Vannier & Massoud 1967) (also found in Dicyrtomidae, probably a convergence) and a retinaculum lacking any setae (the retinaculum always bears setae in adult Eusymphypleona). The strong morphological affinities shared by all Neelipleona leave little doubt on the monophyly of the group, but more data should be acquired if inner Neelipleona relationships are to be studied in detail.

4.2. Relationship among Collembola

The phylogenetic position of Neelipleona uncovered in the present work disagrees with previous molecular and morphological phylogenies that included representatives of this clade. D'Haese (2003a) supported Neelipleona as sister group of Symphypleona (traditional hypothesis of Symphypleona s. l.), Symphypleona s. l. as sister group of Entomobryomorpha and the whole being sister group of Poduromorpha. Gao et al. (2008) molecular phylogeny of Hexapoda and von Reumont et al. (2009) molecular phylogeny of Arthropoda supported Neelipleona as sister group of all other Collembola. Xiong et al. (2008) molecular phylogeny of Collembola supported Symphypleona as sister group of Neelipleona, the whole being sister group of Arthropleona. Finally, Dell'Ampio et al. (2009) phylogeny of Hexapoda uncovered Neelipleona either emerging in a basal trichotomy together with Entomobryomorpha and (Poduromorpha + Symphypleona), or as sister group of all other Collembola.

While each of those analyses proposed a different hypothesis it can be remarked that, at the collembolan order level, the unrooted trees recovered in D'Haese (2003a), Xiong et al. (2008) and Gao et al. (2008) are compatible. Disagreement among those multiple phylogenies are due to the root position. In Xiong et al. (2008) root is on the branch linking Neelipleona + Symphypleona to Entomobryomorpha + Poduromorpha, recovering the traditional separation of Collembola in Arthropleona and Symphypleona s. l. In the present analysis, the root occurs on the branch linking Symphypleona to other Collembola. The unrooted topologies of Collembola of Dell'Ampio et al. (2009) and von Reumont et al. (2009) differ from the unrooted topologies of D'Haese (2003a), Xiong et al.(2008) and Gao et al. (2008). While von Reumont et al. (2009) got strong Bayesian support for the root of Collembola, the Dell'Ampio et al. (2009) root is more problematic, with a basal trifurcation. The cited phylogenies are summed up at the order level in Fig. 2 with the indication of root position. It is to be suspected that the molecular divergence of Collembola with respect to any other Hexapodan lineage is so

pronounced that attempts to polarise the phylogeny are subject to a random outgroup principle (Lanyon 1988, Wheeler 1990), resulting in a random root for the tree. This is precisely what happened with the Symphypleona being 'basal' and paraphyletic in D'Haese (2002). The major challenge in resolving the phylogeny between those orders will be to overcome the problem to provide a robust root hypothesis. The close relationships between Neelipleona and Symphypleona, supported by morphology (Bretfeld 1986, D'Haese 2003a) and recovered by Xiong et al. (2008) molecular based phylogeny are not supported in the present study. Janssens' (2009) hypothesis of Neelipleona being derived from Entomobryomorpha is not supported either. However, this study does not claim to solve the problem of relationship between Collembola orders. The phylogenetic position of Neelipleona remains unclear, and thus, it is to be suspected that Neelipleona is separated from any other collembolan family. Unveiling their precise relationship either with Symphypleona or other Collembola shall require a cladistic analysis including molecular and morphological data with strong homology hypotheses. From this perspective, efforts to understand homology relations between Neelipleona original attributes and other Collembola shall prove decisive.

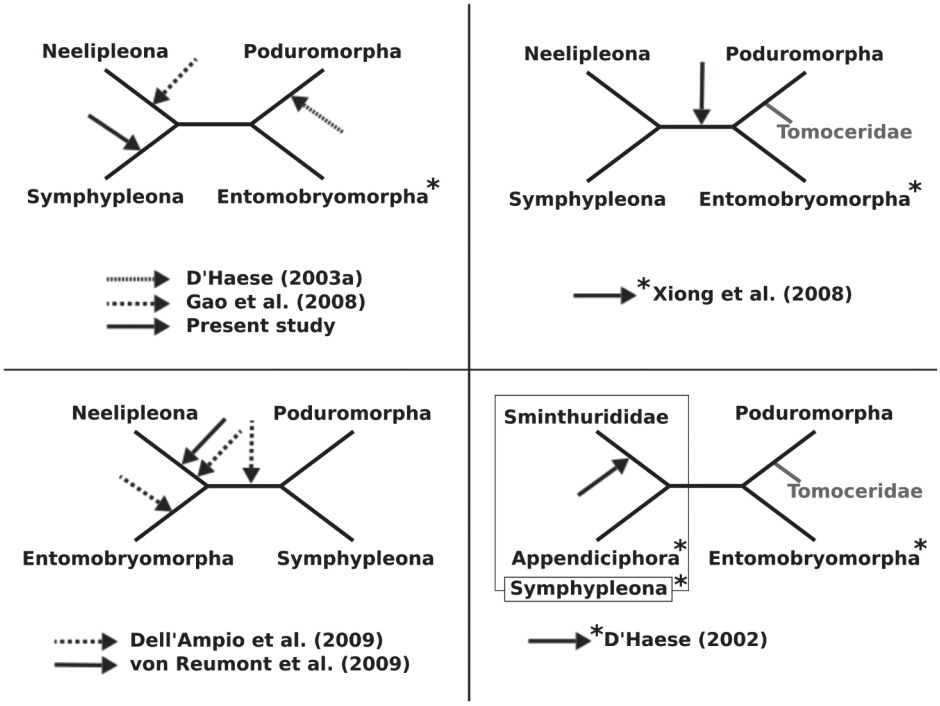

Fig. 2 Phylogenetic relationships among Collembola orders from different publications. The arrows indicate root position for the respective publications. More than one arrow indicates a root ambiguity. A '*' following a taxon name and preceding a publication citation indicates that the taxon is not monophyletic in the respective publication (when Entomobryomorpha is not monophyletic, it is always because of Tomoceridae grouping with Poduromorpha). Except for Dell'Ampio et al. (2009), previous phylogenies differ from the present study only by the root position.

4.3. Branch length and Neelipleona biodiversity

Our results show that for the Neelipleona and Collembola sampling at hand, molecular divergence (for D1, D2, 16S, COX1 regions) in Neelipleona was greater than in any other Collembola orders. Moreover, the molecular divergence in the ubiquitous *Megalothorax* genus alone is greater than in any other order, each represented by several families. This is to be compared to the fact that out of more than 8000 known species of Collembola, only 35 are Neelipleona. We could retrieve 16S and COX1 only for two Symphypleona species, *Sminthurus viridis* and *Sminthurinus bimaculatus*. Those two species are believed to be distant relatives within Symphypleona (this is also buttressed by our analysis), but the lack of data for the other taxa prevents any detection of homoplasy. The strong representation of Neelipleona in the sampling could also have led to a better detection of homoplasy within the group, in regards of the other orders. The reduced data set analyses aimed to reduce this bias. Even though the gap is smaller, the genetic divergence of Neelipleona inferred with this trimmed dataset remains greater than the divergence within Entomobryomorpha and Poduromorpha.

The less diversified order of Collembola (to date) shows a molecular divergence comparable to each other order. This gap between molecular data and taxonomy is a first argument to support the idea that Neelipleona diversity remains largely undescribed.

Another striking result is that the specimens from Brunoy (labeled '*Megalothorax* FRA 1' to 'FRA 7') all collected in the same habitat did not form a monophyletic group and showed high molecular divergence. Some of them even showed a closer relationship with specimens from other regions (Argentina, other regions of France). This means that the sample here was heterogeneous. Using a field microscope, we could not retrieve any morphological evidence for different species being present in the sample (all observed specimens from Brunoy belong to the *minimus* group). We observed the cohabitation of two species of the *minimus* group in recently acquired material (not included in the present study). The species could be distinguished only with the microscope (1000x magnification). It is possible that a similar association occurs at our Brunoy sample. However, since the sequenced specimens are lost, we cannot conclude any further. In this situation the need to save the voucher during DNA extraction to allow microscopical determination is obvious, but extremely difficult. The molecular divergence among the investigated Neelipleona specimens is not correlated to geographical distance and is greater than in any other Collembola orders. This result pleads for the idea that our sample contains several unidentified species. It is at least an optimistic indicator of the need to perform a thorough research on Neelipleona. Morphological strong affinities shared among Neelipleona probably hide the level of diversification of the group, which may be as diversified as other Collembola orders are.

5. Conclusion

A comparison of existing phylogenetic analyses shows that if the Neelipleona position within Collembola is not clearly elucidated, Neelipleona however form a well-defined monophyletic group and are never placed inside any other order. Our analyses buttress this, and reveal a great molecular diversity among a sample of Neelipleona specimens from various worldwide geographical origins; a diversity that is comparable to what is found within other Collembola orders. Neelipleona certainly deserves its taxonomic rank in this regards. This is an example of how molecular tools can bring new light on a poorly known group, justifying the need to thoroughly explore it. DNA sequences, and especially the barcode sequences, can be useful tools (among others, see Stevens et al. 2011) for taxonomists to discover overlooked diversity

(Saunders & McDonald 2010). We believe that further work on taxonomy and evolutionary processes in Neelipleona promises to reveal new insights into collembolan natural history.

6. Supplementary Material

Supplementary information, containing the general POY script used for each analysis plus support files, the tree files along with strict consensus, and the tree resulting from the reduced analysis, is available at http://www.soil-organisms.org.

7. Acknowledgements

A great thank to L. Kováč for providing fresh *Neelus koseli* specimens as well as to C. Houssin for providing *Megalothorax* specimens from Belgium. Most of the other specimens were collected during CaFoTrop field trips throughout the world (www.cafotrop.com), especially during the 'CaFoTrop-Rescapés du Gondwana' project, and partly financed by PPF 'Etat et structure phylogénétique de la biodiversité actuelle et fossile' directed by P. Janvier. DNA extraction and amplification was done in our molecular lab (BoEM). This research was partly supported by the ATM Barcode (Muséum National d'Histoire Naturelle) and the 'Consortium National de Recherche en Génomique', part of the agreement n°2005/67 between the Genoscope and the Muséum National d'Histoire Naturelle on the project 'Macrophylogeny of life' directed by Guillaume Lecointre. Computing resources were provided by the cluster of the Muséum National d'Histoire Naturelle (CNRS UMS 2700). We thank the anonymous reviewers for their helpful comments and criticisms.

8. References

Azpiazu, M. D., V. G. Cairo, J. G. Palacios-Vargas & J. L. Sánchez (2004): Clave dicotómica para la determinación de los colémbolos de Cuba (Hexapoda: Collembola). – Boletín de la Sociedad Entomológica Aragonesa **34**: 73–83.

Börner, C. (1906): Collembola Symphypleona. Fam. Neelidae. – Genera Insectorum **45**: 1–6.

Bretfeld, G. (1986): Phylogenetic systematics of the higher taxa of Symphypleona Börner, 1901 (Insecta, Enthognatha, Collembola). – In: Dallai, R. (ed): Second International Seminar on Apterygota. – University of Siena, Siena: 307–311.

Bretfeld, G. (1999): Symphypleona. – In: W. Dunger (ed.): Synopses on Palaearctic Collembola, Vol. 2. – Abhandlungen und Berichte des Naturkundemuseums Görlitz, Görlitz: **71**: 1–318.

Bretfeld, G., D. Poliakov & M. Broza (2000): Collembola Symphypleona (Insecta, Entognatha) from Israel. – Israel Journal of Zoology **46**: 313–341.

Chagnon, M., D. Paré, & C. Hébert (2000): Relationships between soil chemistry, microbial biomass and the collembolan fauna of southern Québec sugar maple stands. – Ecoscience **7**(3): 307–316.

Christian, E. (1987): Collembola (Springschwänze). – Catalogus Faunae Austriae, **12A**: 1–80.

D'Haese, C. A. (2002): Were the first springtails semi-aquatic? A phylogenetic approach by means of 28S rDNA and Optimization Alignment. – Proceedings of the Royal Society of London B **269**: 1143–1151.

D'Haese, C. A. (2003a): Morphological appraisal of Collembola phylogeny with special emphasis on Poduromorpha and a test of the aquatic origin hypothesis. – Zoologica Scripta **32**: 563–586.

D'Haese, C. A. (2003b): Homology and morphology in Poduromorpha (Hexapoda, Collembola). – European Journal of Entomology **101**: 385–407.

Dallai, R. (1979): Investigations on Collembola. XXIV. On the systematic of Neelidae with redescription of *Neelides folsomi* Caroli. – Animalia **6**: 271–281.

Deharveng, L. (2004): Recent advances in Collembolan systematics. 6th International Seminar on Apterygota, Siena (2002). – Pedobiologia, **48**: 415–433.

Dell'Ampio, E., N. U. Szucsich, A. Carapelli, F. Frati, G. Steiner, A. Steinacher & G. Pass (2009): Testing for misleading effects in the phylogenetic reconstruction of ancient lineages of hexapods: influence of character dependence and character choice in analyses of 28S rRNA sequences. – Zoologica Scripta **38**(2): 155–170.

Folsom, J. W. (1896): *Neelus murinus*, representing a new thysanuran family. – Psyche **7**: 391–409.

Gao, Y., Y. Bu & Y.-X. Luan (2008): Phylogenetic Relationships of Basal Hexapods Reconstructed from Nearly Complete 18S and 28S rRNA Gene Sequences. – Zoological Science **25**(11): 1139–1145.

García-Gómez, A., G. Castaño-Meneses & J. G. Palacios-Vargas (2009): Diversity of springtails (Hexapoda) according to a altitudinal gradient. – Pesquisa Agropecuária Brasileira **44**(8): 911–916.

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

Greenslade, P., M. I. Stevens, G. Torricelli & C. A. D'Haese (2011): An ancient Antarctic endemic genus restored: morphological and molecular support for *Gomphiocephalus hodgsoni* (Collembola: Hypogastruridae). – Systematic Entomology **36**: 223–240.

Hopkin, S. P. (1997): Biology of the springtails. Insecta: Collembola. – Oxford University Press, New York: 330 pp.

Janssens, F. (2009): Note on the collembolan ordinal morphogenetic relationships. – Checklist of the Collembola of the World [http://www.collembola.org/].

Juceviča, E. (2003): Nomina Collembola Latviae. – Latvijas Entomologs **40**: 16–20.

Kováč, L. & V. Papáč (2010): Revision of the genus *Neelus* Folsom, 1896 (Collembola, Neelida) with the description of two new troglobiotic species from Europe. – Zootaxa **2663**: 36–52.

Lanyon, S. M. (1988): The stochastic mode of molecular evolution: what consequences for systematic investigations? – Auk **105**: 565–573.

Loranger, G., I. Bandyopadhyaya, B. Razaka & J.-F. Ponge (2001): Does soil acidity explain altitudinal sequences in collembolan communities ? – Soil Biology & Biochemistry **33**: 381–393.

Massoud, Z. (1971): Contribution à la connaissance morphologique et systématique des Collemboles Neelidae. – Revue d'Ecologie et de Biologie du Sol **8**: 195–198.

Massoud, Z. (1976): Essai de synthèse sur la phylogénie des Collemboles. – Revue d'Ecologie et de Biologie du Sol **13:** 241–252.

Mickevich, M. F. & J. S. Farris (1981): The implications of congruence in Menidia. – Systematic Zoology **27**: 143–158.

Nixon, K. C. (1999): The Parsimony Ratchet, a new method for rapid parsimony analysis. – Cladistics **15**: 407–414.

Palacios-Vargas, J. G., L. Deharveng & C. D'Haese (2011): The genus *Pronura* (Collembola: Neanuridae) in South America, with description of two new species and a barcode sequence for one of them. – Revue Suisse de Zoologie **118**: 197–205.

von Reumont, B. M., K. Meusemann, N. U. Szucsich, E. Dell'Ampio, V. Gowri-Shankar, D. Bartel, S. Simon, H. O. Letsch, R. R. Stocsits, Y.-X. Luan, J. W. Wägele, G. Pass, H. Hadrys & B. Misof (2009): Can comprehensive background knowledge be incorporated into substitution models to improve phylogenetic analyses? A case study on major arthropod relationships. – BMC Evolutionary Biology **9**: 119.

Salmon, J. T. (1964): An index to the Collembola. Volume 1. – Royal Society of New Zealand Bulletin **7**: 1–144.

Saunders, G. W. & B. McDonald (2010): DNA barcoding reveals multiple overlooked Australian species of the red algal order Rhodymeniales (Florideophyceae), with resurrection of *Halopeltis* J. Agardh and description of *Pseudohalopeltis* gen. nov. – Botany **88**(7): 639–667.

Stevens, M. I., D. Porco, C. A. D'Haese & L. Deharveng (2011): Comment on „Taxonomy and the DNA Barcoding Enterprise" by Ebach (2011). – Zootaxa **2838**: 85–88.

Swofford, D. (1990): PAUP: Phylogenetic Analysis Using Parsimony, ver. 3.0. – Illinois Natural History Survey, Champaign, Il.

Vannier, G. & Z. Massoud (1967): Productions cireuses chez les Collemboles Neelidae. – Revue d'Ecologie et de Biologie du Sol **4**(1)**:** 123–130.

Varón, A., L. S. Vinh & W. C. Wheeler (2010): POY version 4: phylogenetic analysis using dynamic homologies. – Cladistics **26**: 72–85.

Wheeler, W. C. (1990): Nucleic acid sequence phylogeny and random outgroups. – Cladistics **6**(4): 363–367.

Wheeler, W. C. (1996): Optimization Alignment: the end of multiple sequence alignment in phylogenetics? – Cladistics **12**: 1–9.

Wheeler, W. C. (1999): Measuring Topological Congruence by Extending Character Techniques. – Cladistics **15**(2): 131–135.

Xiong, Y.,Y. Gao, W.-Y. Yin & Y.-X. Luan (2008): Molecular phylogeny of Collembola inferred from ribosomal RNA genes. – Molecular Phylogenetics and Evolution **49**(3): 728–735.

Accepted 2 November 2011