THE ROYAL SOCIETY

Received 21 November 2001
Accepted 11 February 2002
Published online 9 May 2002

Were the first springtails semi-aquatic? A phylogenetic approach by means of 28S rDNA and optimization alignment

Cyrille A. D'Haese†

ESA 8043 CNRS, Laboratoire d'Entomologie, Muséum National d'Histoire Naturelle, 45 rue Buffon, F-75005 Paris, France

Emergence from an aquatic environment to the land is one of the major evolutionary transitions within the arthropods. It is often considered that the first hexapods, and in particular the first springtails, were semi-aquatic and this assumption drives evolutionary models towards particular conclusions. To address the question of the ecological origin of the springtails, phylogenetic analyses by optimization alignment were performed on D1 and D2 regions of the 28S rDNA for 55 collembolan exemplars and eight outgroups. Relationships among the orders Symphypleona, Entomobryomorpha and Poduromorpha are inferred. More specifically, a robust hypothesis is provided for the subfamilial relationships within the order Poduromorpha. Contrary to previous statements, the semi-aquatic species *Podura aquatica* is not basal or 'primitive', but well nested in the Poduromorpha. The analyses performed for the 24 different weighting schemes yielded the same conclusion: semi-aquatic ecology is not ancestral for the springtails. It is a derived condition that evolved independently several times. The adaptation for semi-aquatic life is better interpreted as a step towards independence from land, rather than indication of an aquatic origin.

Keywords: 28S rDNA; Collembola; optimization alignment; phylogenetic test; semi-aquatic; terrestrialization

1. INTRODUCTION

The ecology of the first hexapods, and in particular of the first springtails, has generated debate. It raises a question of one of the major evolutionary transitions within arthropods: when did the passage from a marine to a terrestrial environment take place? Did terrestrialization occur at the base of the Atelocerata (if one promotes the Atelocerata hypothesis over the competing Pancrustacea) or at the origin of the hexapods? Were there single or multiple transitions? Was fresh water an 'intermediate' environment for that transition? The springtails (class Collembola) are one of the key taxa to assess arthropod relationships (Giribet & Ribera 2000), particularly hexapod relationships, and therefore to assess the question of the ecological origin of terrestriality in this major animal clade.

Traditionally a semi-aquatic origin has been proposed for springtails and therefore for hexapods as a whole. Some authors, interested in the palaeozoic ecology of terrestrial arthropods, have stated that early Collembola were semi-aquatic (Shear & Kukalová-Peck 1990; Kukalová-Peck 1991). According to these authors, when springtails appeared, soil and litter habitats were not yet present (see also Kukalová-Peck 1987). Shear & Kukalová-Peck (1990) thus assumed that the first 'terrestrial' arthropods lived in algal mats and emergent vegetation. Springtails, then, would have adapted to live in saturated terrestrial substrates and only later colonized the ground. This idea has been based on the fact that springtails are the first hexapods to appear in the fossil record with the early Devonian *Rhyniella praecursor* (at the Siegenian specifically, 391 to 397 Myr ago; Westoll 1977). These fossils are associated with an aquatic environment. However, the majority of fossils are preserved in aquatic environments, so this putatively semi-aquatic fossil cannot be taken to indicate a semi-aquatic origin (Pritchard *et al.* 1993).

The reason behind the aquatic origin assumption is also embedded in the paradigmatic vision of *Podura aquatica*. This semi-aquatic species has always been regarded as 'primitive' (e.g. Yosii 1961; Uchida 1971; Dallai 1973) and thus supposed to provide a direct insight into the ancestor characteristics. Examination of all proposed phylogenetic scenarios finds *P. aquatica* placed in a basal position, regardless of the groups to which it is attached (Börner 1906; Yosii 1961; Uchida 1971; Massoud 1976; Cassagnau 1990). Therefore, semi-aquatic has been equated with primitive.

Kukalová-Peck (1983) presented another hypothesis assuming that the ancestor of the Myriapoda + Hexapoda stem group already had aquatic first instars and that the lineages became terrestrial independently. She asserted that most generalized pterygote orders retain aquatic juveniles. That aquatic scenario for the origin of hexapods is rather widespread and has been advocated for some time (e.g. Riek 1971; Kukalová-Peck 1978; Tom 1984; Thomas *et al.* 2000). It is based on the fact that the pterygote groups generally regarded as 'primitive' have aquatic larvae: mayflies (Ephemeroptera), dragonflies and damselflies (Odonata), stoneflies (Plecoptera), alderflies and dobsonflies (Megaloptera) and caddisflies (Trichoptera). Averof & Cohen (1997) also assumed that the first hexapods were aquatic and went further, considering five or six independent transitions from aquatic to terrestrial environments, and in particular, the colonization of terres-

† Present address: Division of Invertebrate Zoology, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024, USA (dhaese@amnh.org).

Table 1. Taxa examined in this study with the accession codes from GenBank.

class	order	family	subfamily	genus species	GenBank accession 28S D1	GenBank accession 28S D2	locality
Malacostraca	Decapoda	Astacidae		*Euastacus bispinosus*	AF235981	AF235981	
				Geocharax gracilis	AF235982	AF235982	
Insecta	Hymenoptera	Vespidae	Vespinae	*Vespa crabro*	AF067145	AF067145	
		Formicidae	Myrmeciinae	*Myrmecia croslandi*	AB052895	AB052895	
	Dictyoptera	Blattidae		*Periplaneta americana*	AF321248	AF321248	
	Zygentoma	Lepismatidae		*Thermobia domestica*	AF483404	AF483462	France: Ile-de-France, Paris
	Archaeognatha	Machilidae	Petrobiinae	*Petrobius brevistylis*	AF483389	AF483447	France: Manche, Carteret
Protura		Acerentomidae		*Acerella muscorum*	AF483354	AF483412	France: Manche, Rozel
Collembola	Entomobryomorpha	Entomobryidae	Entomobryinae	*Entomobrya lanuginosa*	AF483365	AF483423	France: Ile-de-France, Paris
				Pseudosinella sp.	AF483393	AF483451	France: Ile-de-France, Paris
				Sinella curviseta	AF483396	AF483454	France: Ile-de-France, Paris
			Orchesellinae	*Orchesella cincta*	AF483385	AF483443	France: Ile-de-France, Paris
				Orchesella villosa	AF483386	AF483444	France: Ile-de-France, St Cyr sur Dourdan
		Isotomidae		*Cryptopygus antarcticus*	AF483363	AF483421	South Antarctic: King George Island
				Folsomia candida	AF483366	AF483424	France: Ariège, Moulis
				Parisotoma notabilis	AF483388	AF483446	France: Ile-de-France, Paris
				Isotoma viridis	AF483372	AF483430	France: Ile-de-France, Paris
		Oncopoduridae		*Oncopodura crassicornis*	AF483383	AF483441	France: Pyrénées-Atlantiques, Lescun
		Tomoceridae		*Tomocerus minor*	AF483406	AF483464	France: Pyrénées-Atlantiques, Lescun
	Symphypleona	Dicyrtomidae		*Dicyrtoma* sp.	AF483364	AF483422	Mexico: Chiapas, Miramar
				Ptenothrix sp.	AF483394	AF483452	Mexico: Morelos, Mexicapan
		Katianninae		*Arrhopalites sericus*	AF483359	AF483417	France: Manche, Rozel
				Sminthurinus bimaculatus	AF483398	AF483456	France: Ile-de-France, Paris
		Sminthuridae		*Allacma fusca*	AF483355	AF483413	France: Oise, Labosse
				Caprainea marginata	AF483361	AF483419	France: Oise, Labosse
				Sminthurus viridis	AF483399	AF483457	France: Oise, Labosse
		Sminthurididae		*Sminthurides* sp.	AF483397	AF483455	Mexico: Manzillano, Colima
	Poduromorpha	Brachystomellidae		*Brachystomella parvula*	AF483360	AF483418	France: Corsica, Poretta
		Hypogastruridae		*Ceratophysella gibbosa*	AF483362	AF483420	France: Ile-de-France, Paris
				Hypogastrura vernalis	AF483371	AF483429	France: Haute-Garonne, Toulouse
				Microgastrura sensiliata	AF483379	AF483437	France: Ariège, Prat d'Albi
				Paraxenylla affiniformis	AF483387	AF483445	France: Languedoc-Roussillon, Leucate
				Schoettella ununguiculata	AF483395	AF483453	France: Haute-Garonne, Toulouse
				Triacanthella perfecta	AF483407	AF483465	France: Bordeaux, Arès
				Xenylla grisea	AF483409	AF483467	France: Manche, Rozel
				Xenylla tullbergi	AF483410	AF483468	France: Oise, Labosse
				Willemia denisi	AF483408	AF483466	Sweden: Branstrop

(*Continued.*)

Table 1. (*Continued.*)

class	order	family	subfamily	genus species	GenBank accession		
					28S D1	28S D2	locality
	Poduromorpha	Neanuridae	Frieseinae	*Friesea grisea*	AF483367	AF483425	South Antarctic: King George Island
				Friesea mirabilis	AF483368	AF483426	Spain: Lerida, Nargo
				Friesea truncata	AF483369	AF483427	France: Manche, Rozel
			Morulinae	*Morulina verrucosa*	AF483381	AF483439	Poland: High Tatras, Zakopane
			Neanurinae	*Monobella g. grassei*	AF483380	AF483438	France: Oise, Labosse
				Neanura muscorum	AF483382	AF483440	France: Oise, Labosse
				Thaumanura carolii	AF483403	AF483461	Poland: Gorce Mts, Foredówki
			Pseudachorutinae	*Anuridella calcarata*	AF483358	AF483416	France: Corsica, St Florent
				Anurida granaria	AF483356	AF483414	France: Manche, Morville
				Anurida maritima	AF483357	AF483415	France: Finistère, Calotte Island
				Micranurida pygmaea	AF483377	AF483435	France: Pyrénées-Orientales, Vivès
				Pseudachorutes palmiensis	AF483392	AF483450	France: Haute-Garonne, Herran
		Odontellidae		*Superodontella alpina*	AF483400	AF483458	Spain: Lerida, Vielha
				Superodontella gisini	AF483401	AF483459	France: Pyrénées-Orientales, Estagel
				Xenyllodes armatus	AF483411	AF483469	Spain: Santander, Soba
		Onychiuridae	Onychiurinae	*Micraphorura absoloni*	AF483378	AF483436	Poland: High Tatras, Zakopane
				Onychiurus ambulans	AF483384	AF483442	France: Nord, Berthen
				Kalaphorura paradoxa	AF483373	AF483431	France: Ile-de-France, Paris
				Protaphorura armata	AF483391	AF483449	France: Manche, Quettehou
			Tetrodontophorinae	*Homaloproctus sauteri*	AF483370	AF483428	Japan: Yokotasiro Oze
				Lophognathella choreutes	AF483374	AF483432	Japan: Hokkaido, Tomakomai
				Tetrodontophora bielanensis	AF483402	AF483460	Poland: High Tatras, Zakopane
			Tullbergiinae	*Mesaphorura macrocheta*	AF483375	AF483433	France: Ile-de-France, Paris
				Metaphorura denisi	AF483376	AF483434	France: Corsica, St Florent
				Tillieria penai	AF483405	AF483463	South Antarctic: King George Island
		Poduridae		*Podura aquatica*	AF483390	AF483448	France: Oise, Lalandelle

trial environments independently for each apterygote lineage (see Averof & Cohen 1997, p. 630, fig. 4).

A phylogenetic framework is crucial to understand the evolution and the ecology of the first Collembola. To date, to our knowledge, no phylogenetic study of the higher relationships among Collembola has been conducted with the exception of those of Lee *et al.* (1995*a*,*b*). However, their work included only four and seven taxa, respectively, which does not allow for general phylogenetic conclusions.

The goal of this study is to test hypotheses about the terrestrial versus semi-aquatic origin of springtails, thus providing an insight into the possible aquatic or semi-aquatic origin of hexapods. It is necessary to resolve the phylogeny of the Collembola, particularly the relationships between the three orders Symphypleona, Entomobryomorpha and Poduromorpha. Within the Poduromorpha, detailed examination, down to the subfamily, was performed in order to properly address the problem of *P. aquatica* and other semi-aquatic species.

2. MATERIAL AND METHODS

(a) *Taxonomic sampling*

The species studied are listed in table 1 with GenBank accession numbers. Two crustaceans were selected from GenBank as remote outgroups, cladograms were actually rooted on *Euastacus bispinosus*. Three pterygotes (*Vespa crabro*, *Myrmecia croslandi* and *Periplaneta americana*) were also selected from GenBank. Three so-called 'apterygotes' were sequenced as close outgroups: representatives of Archaeognatha (*Petrobius brevistylis*), Zygentoma (*Thermobia domestica*) and Protura (*Acerella muscorum*). The latter is usually considered to be the sister group of Collembola. The diversity of species in the outgroup allows robust polarization of the cladograms and a strong test of the monophyly of the Collembola.

Three collembolan orders Symphypleona, Entomobryomorpha and Poduromorpha were sampled. When possible, each family or subfamily was represented by several exemplars to test its monophyly. The four most speciose families of Symphypleona were represented out of the eight: Dicyrtomidae, Katiannidae, Sminthuridae and Sminthurididae. The remaining families comprise from one to four species and are difficult to sample, with the exception of the Bourletiellidae (*ca*. 180 species but not included here). In the same way, four families of the more speciose Entomobryomorpha were sampled out of the nine of the order (Entomobryidae, Isotomidae, Oncopoduridae and Tomoceridae) with the exception of Paronellidae (*ca*. 350 species) and Cyphoderidae (*ca*. 130 species). In Poduromorpha, all families but Gulgastruridae (one species only: *Gulgastrura reticulosa*) were represented: Brachystomellidae, Hypogastruridae, Neanuridae, Odontellidae, Onychiuridae and Poduridae. Within poduromorphan families, all subfamilies except the Caputanurinae (Neanuridae) were sampled.

Efforts were made to include as many species associated with semi-aquatic environments as possible. *Podura aquatica* (Poduromorpha, Poduridae) and *Sminthurides* sp. (Symphypleona, Sminthurididae) live on the surface of ponds, allegedly the habitat of the first springtails. *Anurida maritima* (Poduromorpha, Neanuridae) is common in the intertidal zone and *Anuridella calcarata* (Poduromorpha, Neanuridae) in littoral sands.

(b) *DNA extraction, PCR and sequencing*

Total DNA was extracted by using a protocol modified from Winnepenninckx *et al.* (1993). Specimens were ground in 50 °C CTAB buffer (2% cetyltrimethylammonium bromide, 1.4 M NaCl, 20 mM EDTA, 100 mM Tris–HCl, pH 8.0, 0.2% β-mercaptoethanol) with Proteinase-K (100 mg ml^{-1}). After incubation at 50 °C for 4 h, equal-volume chloroform–isoamylalcohol (96 : 4) and two-thirds volume isopropanol were added then allowed to precipitate overnight at 4 °C. The sample was spun for 15 min at 14 000 rpm, the supernatant removed, and the pellet washed with 76% ethanol and air-dried. Finally, DNA was resuspended in bi-distilled water. Amplification of D1 and D2 regions of the 28S rDNA (800 bp) was carried out via polymerase chain reaction (PCR). Amplifications were performed for 35 cycles in 50 µl volume. In each cycle, DNA was denatured at 94 °C for 30 s, primers annealed between 52 °C and 58 °C for 30 s, and extension was at 72 °C for 30 s. Primers used for amplification were C1′ (5′-ACCCGCTGAATTTAAGCAT-3′), D2 (5′-TCCGTGTTTCAAGACGGG-3′), D3 (5′-GCATAGTTCACCATCTTTC-3′), C2′ (5′-GAAAAGAACTTTGRARAGAGAGT-3′), C2 (5′-TGAACTCTCTCTTCAAAGTTCTTTTC-3′), C2′coll (5′-GAGACCGATAGCGAACAAGTACCGTGA-3′), C2coll (5′-ATGCACTTTGGAGAGCCCCCATTCGCCTGTCTT-3′) and D2coll (5′-ACCACGCATGCWTTAGATTG-3′). The three latter (C2′coll, C2coll and D2coll) were specifically designed for this study.

PCR products were sequenced with either amplification or internal primers. Sequencing was performed manually with a Thermo Sequenase Cycle Sequencing kit (Amersham), based on the Sanger method (Sanger *et al.* 1977).

(c) *Data analysis by optimization alignment*

Sequences were analysed by the means of optimization alignment (Wheeler 1996) implemented in the software Poy (Gladstein & Wheeler 1996–2000) available at ftp://ftp.amnh.org.pub.molecular.poy. This parsimony-based method avoids intermediate alignment steps by directly assessing the number of evolutionary events, i.e. DNA sequence transformations. This is accomplished through the generalization of existing character optimization procedures to insertion and deletion events (indels) and base substitutions. The crux of the model is the treatment of indels as processes as opposed to the patterns (i.e. gaps) implied by multiple sequence alignment. This method generates more efficient explanations of sequence variation than multiple alignments and produces more congruent results (shorter trees) (Wheeler & Hayashi 1998).

The influence of gap, transition and transversion costs was studied through sensitivity analysis (Wheeler 1995) to avoid an arbitrary choice of parameters. Twenty-four parameter sets were specified to explore the sensitivity of the results to parameter variation. The ratio of weights between indels and transversions ranged from 0.5 to 8 and the ratio between transversions and transitions ranged from 0.5 to 4. In addition, the data were analysed with a transversion/transition ratio (Tv : Ts) of ∞ (transitions set to 0). The results are reported in non-interpolated Cartesian graphs of areas of the parameter space. The areas of the graphs show the strict consensus results whether the groups are monophyletic, paraphyletic or polyphyletic (Wheeler 1995; Janies 2001). Here, the graphs report areas in which the analysis recover a monophyletic, paraphyletic or polyphyletic group. Additionally, a graph reports areas in which the hypothesis of semi-aquatic origin is or is not recovered.

Table 2. Cladogram lengths and incongruence values for analyses of the 24 parameter sets.

gap cost ratio	transversion cost ratio	length D1	length D2	length combined	ILD
0.5	1	1414	5011	6486	0.009 404 872
0.5	2	1067	4037	5178	0.014 291 232
0.5	4	1680	6522	8314	0.013 471 253
0.5	∞	564	2302	2920	0.018 493 151
1	0.5	1177	4325	5573	0.012 739 996
1	1	872	3499	4402	0.007 042 254
1	2	1329	5777	7205	0.013 740 458
1	4	2017	8872	11 051	0.014 659 307
1	∞	363	1724	2125	0.017 882 353
2	0.5	1368	5402	6848	0.011 390 187
2	1	966	4162	5181	0.010 229 685
2	2	1444	6574	8098	0.009 878 982
2	4	2349	11 150	13 668	0.012 364 647
2	∞	438	2272	2785	0.026 929 982
4	0.5	1527	6355	8008	0.015 734 266
4	1	1108	5051	6273	0.018 173 123
4	2	1697	8264	10 232	0.026 485 536
4	4	2834	14 512	17 906	0.031 274 433
4	∞	558	3055	3779	0.043 926 965
8	0.5	1753	7896	9957	0.030 933 012
8	1	1314	6501	8128	0.038 508 858
8	2	2099	11 014	13 831	0.051 912 371
8	4	3691	19 857	24 812	0.050 943 092
8	∞	752	4361	5460	0.063 553 114

Congruence is used as an optimality criterion to choose the parameter set that minimizes incongruence among the regions of the amplified 28S rDNA portion. Congruence was measured by the ILD metrics (Mickevich & Farris 1981). This value is calculated by dividing the difference between overall tree length and the sum of its data components:

$$\mathrm{ILD} = (\mathrm{length}_{\mathrm{combined}} - \sum \mathrm{length}_{\mathrm{individual\ sets}})/\mathrm{length}_{\mathrm{combined}}.$$

The tree from the analysis that minimizes character conflict among all the data is taken as the best estimate of the phylogeny. An example of a complete Poy command line used for a given stepmatrix is: 'poy 1 -noleading -norandomizeoutgroup -molecularmatrix [stepmatrix] -build 10 -buildspr -buildmaxtrees 1 -random 20 -sprmaxtrees 1 -tbrmaxtrees 2 -ratchetspr 50 -ratchetpercent 15 -ratchetseverity 3 -ratchettrees 2 -fitchtrees -maxtrees 5 -slop 5 -checkslop 10 28sd1.flat 28sd2.flat > 28Sd1d2-111.out.' Relative support for each node was assessed with bremer support indices (Bremer 1994). Analyses were performed on the AMNH computer cluster (280 nodes, 1024 Mb RAM per node, 560 CPUs from 500 MHz PIII to 1 GHz PIII, 100 Mb s^{-1} Ethernet/10.4 Gb switch). Ecological features (semi-aquatic versus terrestrial life in soils) were parsimoniously optimized on the cladogram obtained by the analyses of the 28S rDNA (e.g. Coddington 1988; Carpenter 1989; Miller & Wenzel 1995; Grandcolas 1997; D'Haese 2000).

3. RESULTS

Sequences of 58 taxa, including 55 Collembola, were obtained for D1 and D2 regions of 28S rDNA (table 1), covering the major families (58%) and subfamilies (64%) of springtails.

(a) *Phylogeny*

Analyses were performed for 24 distinct parameter sets. The equal weighting scheme yielded the lowest ILD value (0.007 042 254) and therefore minimized the incongruence among the datasets (table 2). The equal weighting analysis resulted in one most parsimonious cladogram of 4402 steps (figure 1). In this tree, the monophyly of Collembola is highly supported (bremer 47). Symphypleona are paraphyletic with Sminthuridae and Dicyrtomidae paraphyletic, whereas Katiannidae are monophyletic. Entomobryomorpha are paraphyletic with Isotomidae monophyletic and Entomobryidae paraphyletic; Tomoceridae and Oncopoduridae cluster together. Poduromorpha are monophyletic with two major clades: one including *Triacanthella* (Hypogastruridae), Odontellidae and Onychiuridae, the other including the remaining Hypogastruridae, Poduridae, Brachystomellidae and Neanuridae.

Podura aquatica is the sister group of *Microgastrura* + Brachystomellidae + Neanuridae, nested within Poduromorpha.

The present analysis shows without ambiguity that *P. aquatica* is not placed near the base of springtail phylogeny, contrary to the widely held opinion of its 'primitivity'.

(b) *Sensitivity analysis*

The results of the sensitivity analysis are summarized in table 2 and figure 2. Collembola are recovered in all analyses except gap : Tv : Ts = 16 : 2 : 1 and 8 : 1 : 0, where they are paraphyletic. Most often, Poduromorpha are monophyletic; the analyses do not retrieve this group for 'extreme' parameter sets (i.e. gap ratio $\log_2 > 2$ or change ratio $\log_2 = \infty$).

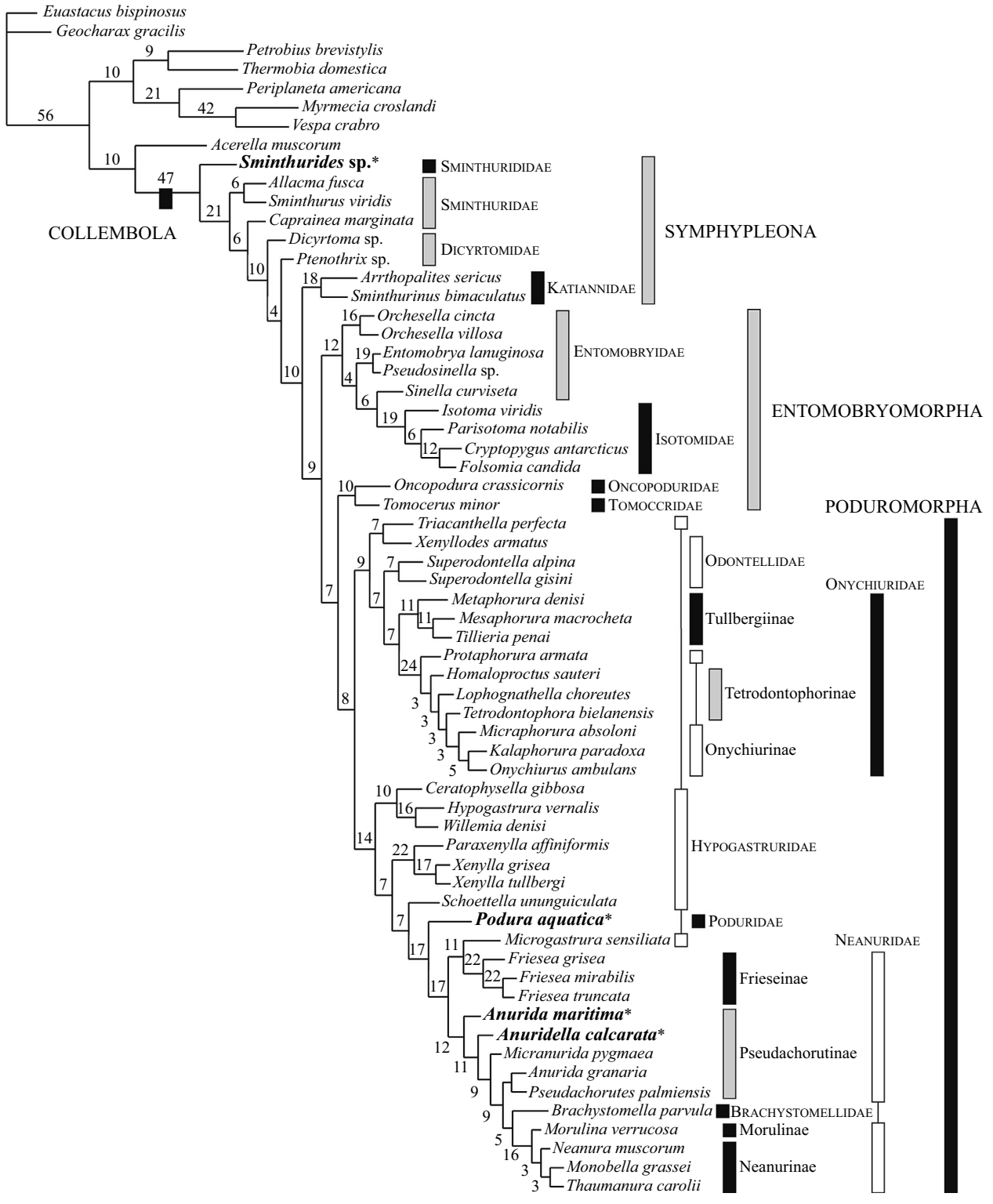

Figure 1. Most parsimonious cladogram obtained by optimization alignment for combined D1 and D2 regions of 28S rDNA with equal weighting (length, 4402; consistency index, 0.50; and retention index, 0.69). Orders, families and subfamilies are indicated by black, grey and white bars for monophyletic, paraphyletic and polyphyletic respectively. Bremer support values are indicated below branches. Semi-aquatic species are in bold type and marked by an asterisk.

Entomobryomorpha are polyphyletic most of the time, except $-1 <$ gap ratio $\log_2 < 2$, when they are generally paraphyletic, and occasionally monophyletic. Under change ratio $\log_2 = 0$ and gap ratio $\log_2 > 3$, the Symphypleona are not recovered; otherwise the group is generally monophyletic. Paraphyly or polyphyly of Entomobryomorpha is not surprising, since it is a poorly defined group. Any springtail with an elongate body shape (character present in Poduromorpha and outside Collembola) and a reduced prothorax (character present in Symphypleona) is included in this group. Conversely, the non-monophyly of Symphypleona is very unlikely given the number of derived morphological characters they possess. However, even if Symphypleona is not reco-

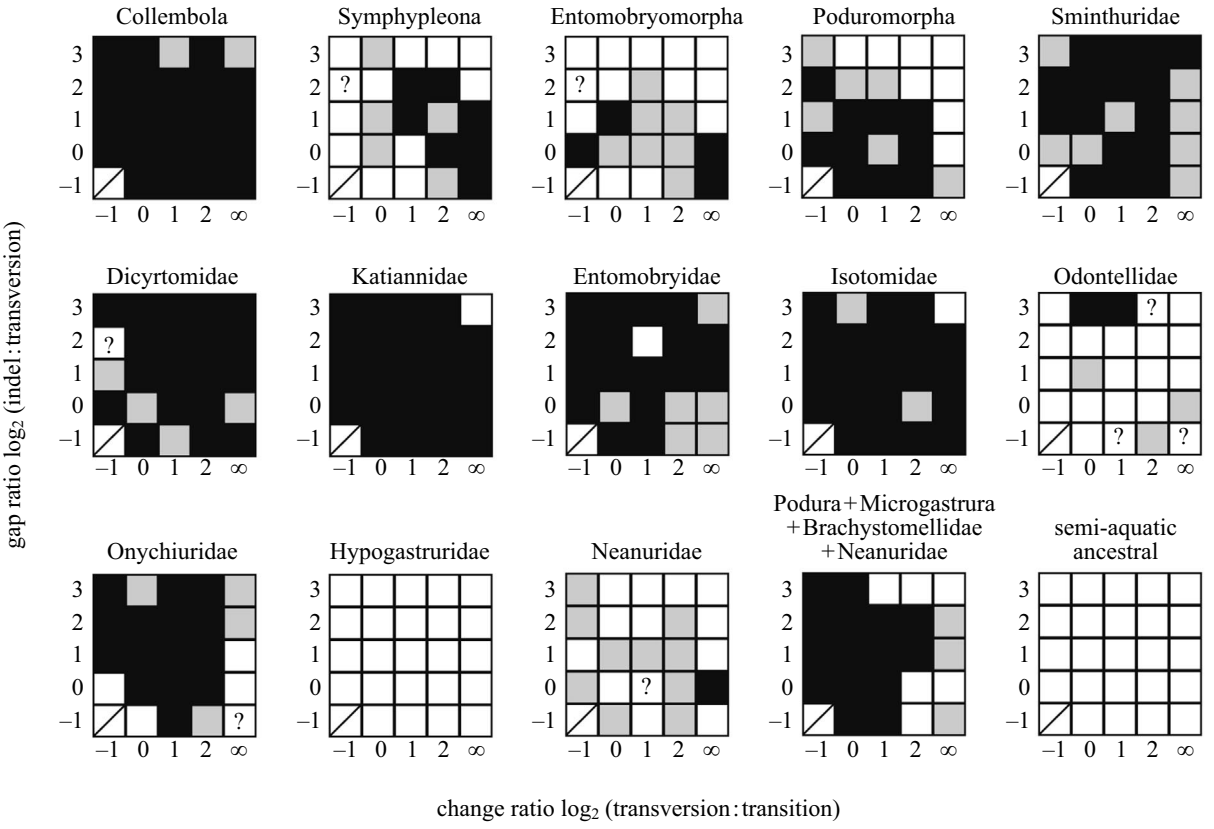

Figure 2. Summary of higher taxonomic groups recovered under the 24 analytical conditions. The recovery of monophyletic, paraphyletic and polyphyletic groups are designated by black, grey and white squares, respectively. Question marks stand for ambiguous answers (polytomies). The graph at the bottom right reports the recovery of a semi-aquatic origin (black) or not (white) under the 24 analytical conditions.

vered in the majority of the analyses, it is monophyletic for many parameter sets (figure 2).

This problem for Symphypleona is connected to the variation of the root position for Collembola depending on the parameter sets. The variation of the position of the root (eight times in Symphypleona, six in Entomobryomorpha, five in Poduromorpha and four in a basal polytomy composed of Symphypleona and Entomobryomorpha) points to a random outgroup problem (Lanyon 1988; Wheeler 1990). Collembola are morphologically rather distinct compared with the other hexapods. It is then difficult to recover enough relevant information to properly assess the position of the root, whether from molecules or from morphology. The information is saturated, and the root tends to be placed near long branches, even in groups where monophyly is not in question like Symphypleona. According to these considerations and the results of the sensitivity analysis, the relationships (Entomobryomorpha (Symphypleona Poduromorpha)) with Entomobryomorpha paraphyletic are more likely.

The traditionally recognized families Sminthuridae, Dicyrtomidae, Katiannidae, Entomobryidae, Isotomidae and Onychiuridae are largely unaffected by parameter variation and are usually recovered.

Hypogastruridae are always polyphyletic, *Triacanthella* being more closely related to Odontellidae and Onychiuridae, and *Microgastrura* more closely related to *Podura, Brachystomella* and Neanuridae. Again, this result is not a surprise as Hypogastruridae were defined on symplesiomorphies: presence of a molar plate and absence of pseudocelli. Traditionally, Odontellidae were grouped with Neanuridae and Brachystomellidae (Massoud 1967, 1976; Cassagnau 1971) on the basis of mandible reduction. In the present study, Odontellidae are closely related to Onychiuridae (closer to the Tullbergiinae, specifically). Actually, the mouthparts of the Odontellidae are different from those of Neanuridae and Brachystomellidae (e.g. direct joint stipes–fulcrum in Odontellidae as opposed to the classical stipes–cardo–fulcrum articulation; see Deharveng (1981)). Paraphyly or polyphyly of the Neanuridae is the result of the position of *Brachystomella* and *Microgastrura* in that family. Both have reduced mandibles, as do Neanuridae.

With near uniformity, *Podura* is the sister group of *Microgastrura* + Brachystomellidae + Neanuridae, or the sister group of *Microgastrura*, that clade being the sister group of Brachystomellidae + Neanuridae.

(c) *Optimization of the semi-aquatic life history*

A semi-aquatic origin for springtails is not retrieved from the phylogeny. Instead, the data show that semi-aquatic ecology is a derived condition that evolved several times independently (figure 1). Furthermore, whatever the weighting scheme, the semi-aquatic condition is never ancestral (figure 2). Under all the weighting schemes, an edaphic origin is retrieved.

4. DISCUSSION: ECOLOGICAL ORIGIN OF COLLEMBOLA

According to the position of the semi-aquatic taxa in the phylogenetic analyses here, semi-aquatic ecology is definitely not ancestral for the Collembola. On the contrary, a semi-aquatic lifestyle is a secondary acquisition that occurred several times independently in the evolution of Collembola. An edaphic lifestyle is the ancestral state. The ancestors of Collembola were living in soils as the vast majority of extant species do.

The marine aquatic origin would be found before the common ancestor of extant hexapods or the Atelocerata stem group, depending of the position of hexapods in arthropods.

Contrary to the statement of Kukalová-Peck and Shear (Kukalová-Peck 1987, 1991; Shear & Kukalová-Peck 1990), biogenic soils were present in the Ordovician period at least, well before the likely origin of Collembola in the early Devonian. The spores and cuticles of plants of bryophyte type that were found 460 Myr ago (Edwards & Selden 1992) gave a minimum age for the colonization of the land by plants and therefore for the presence of soils and litter. Soil habitats were probably among the first terrestrial ecosystems available, and the colonization of the ground by plants provided food, substrate and protection for the first Hexapoda, and for the first Collembola.

In order to colonize the terrestrial environments, a number of physiological barriers had to be overcome. The soil is a refuge from the biotic and abiotic perturbations of the aerial terrestrial environment (Villani *et al.* 1999). Litter and soil microhabitats allow for a transition, providing a continuum of environments from fully aquatic to fully terrestrial (Ghilarov 1958; Vannier 1973, 1978). Indeed, soils are saturated with humidity, preventing desiccation while allowing respiration. In the first hexapods, gaseous exchanges were made through the thin cuticule, as in extant springtails. Secondarily, specialized organs (tracheae) were developed to overcome the thickening of the cuticule required by a real aerial life, as in pterygots. Nearly all the so-called 'primitive' or 'basal' hexapods are found in the humid surroundings of soil habitats: Diplura, Protura, Collembola, and to a lesser extent Archaeognatha and Zygentoma. Accordingly, the most parsimonious scenario points to an edaphic origin (Ghilarov 1958; Wigglesworth 1972; Little 1983, 1990).

The semi-aquatic life is better interpreted as a step towards independence from land, rather than indication of an aquatic origin.

I thank Wanda Weiner, Judith Najt, Louis Deharveng, José Palacios-Vargas and Florence Rousseau for the specimens they provided. I also thank Meredith Mahoney, Daniel Stillman, Ward Wheeler and especially Philippe Grandcolas for help and comments on the manuscript. DNA sequences were obtained at the Service de Systématique Moléculaire (FR 1541 CNRS) where I found financial and technical support: I especially thank Simon Tillier, Annie Tillier, Céline Bonillo and Josie Lambourdière.

REFERENCES

Averof, M. & Cohen, S. M. 1997 Evolutionary origin of insect wings from ancestral gills. *Nature* **385**, 627–630.

Börner, C. 1906 Das system der Collembolen, nebst beschreibung neuer Collembolen des Hamburger naturhistorischen museums. *Mitt. Nat. Mus. Hamburg* **23**, 147–188.

Bremer, K. 1994 Branch support and tree stability. *Cladistics* **10**, 295–304.

Carpenter, J. M. 1989 Testing scenarios: wasp social behaviour. *Cladistics* **5**, 131–144.

Cassagnau, P. 1971 La phylogénie des Collemboles à la lumière des structures endocrines rétrocérébrales. In *I Symposio Internacional Zoolfilogenia*, pp. 333–349. Universidad de Salamanca.

Cassagnau, P. 1990 Des Hexapodes vieux de 400 millions d'années: les Collemboles. I. Biologie et évolution. *Année Biol.* **29**, 1–37.

Coddington, J. A. 1988 Cladistic tests of adaptational hypotheses. *Cladistics* **4**, 3–22.

Dallai, R. 1973 Richerche sui Collemboli. XVIII. La cuticola e gli gnatiti di *Tetrodontophora bielanensis* (Waga) al microscopio eletronico a scansione. *Redia* **54**, 105–116.

Deharveng, L. 1981 La famille des Odontellidae: phylogénèse et taxonomie. *Trav. Lab. Ecobiol. Arthrop. Edaph. Toulouse* **3**, 1–21.

D'Haese, C. 2000 Is psammophily an evolutionary dead end? A phylogenetic test in the genus *Willemia* (Collembola: Hypogastruridae). *Cladistics* **16**, 255–273.

Edwards, D. & Selden, P. A. 1992 The development of early terrestrial ecosystems. *Bot. J. Scotland* **46**, 337–366.

Ghilarov, M. S. 1958 L'importance du sol dans l'origine et l'évolution des Insectes. In *Proc. 10th Int. Cong. of Entomology*, vol. I (ed. E. C. Becker), pp. 443–451. Ottawa, Canada: Mortimer Limited.

Giribet, G. & Ribera, C. 2000 A review of arthropod phylogeny: new data based on ribosomal DNA sequences and direct character optimization. *Cladistics* **16**, 204–231.

Gladstein, D. S. & Wheeler, W. C. 1996–2000 POY. Software for direct optimization of DNA and other data. New York: American Museum of Natural History.

Grandcolas, P. (ed.) 1997 *The origin of biodiversity in insects: phylogenetic tests of evolutionary scenarios*. Paris: Mémoires du Muséum National d'Histoire Naturelle.

Janies, D. 2001 Phylogenetic relationships of extant echinoderm classes. *Can. J. Zool.* **79**, 1232–1250.

Kukalová-Peck, J. 1978 Origin and evolution of insect wings and their relation to metamorphosis, as documented by the fossil record. *J. Morphol.* **156**, 53–126.

Kukalová-Peck, J. 1983 Origin of the insect wing and wing articulation from the arthropodan leg. *Can. J. Zool.* **61**, 1618–1669.

Kukalová-Peck, J. 1987 New carboniferous Diplura, Monura and Thysanura, the hexapod ground plan, and the role of thoracic lobes in the origin of wings (Insecta). *Can. J. Zool.* **65**, 2327–2345.

Kukalová-Peck, J. 1991 Fossil history and the evolution of hexapod structures. In *The insects of Australia*, vol. 1 (ed. I. D. Naumann), pp. 141–179. Melbourne, Australia: Melbourne University Press.

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

Lee, B.-H., Hwang, U.-W., Kim, W., Park, K.-H. & Kim, J.-T. 1995*a* Phylogenetic study of the suborder Arthropleona (Insecta: Collembola) based on morphological characters and 18S rDNA sequence analysis. *Polskie Pismo Ent.* **64**, 261–277.

Lee, B.-H., Hwang, U.-W., Kim, W., Park, K.-H. & Kim, J. T. 1995*b* Systematic position of cave Collembola, G*ulgastrura reticulosa* (Insecta) based on morphological characters and 18S rDNA nucleotide sequence analysis. *Mém. Biospéléol.* **22**, 83–90.

Little, C. 1983 *The colonisation of land. Origins and adaptations of terrestrial animals*. Cambridge University Press.

Little, C. 1990 *The terrestrial invasion. An ecophysiological approach to the origins of land animals*. Cambridge Studies in Ecology. Cambridge University Press.

Massoud, Z. 1967 Monographie des Neanuridae, Collemboles Poduromorphes à Pièces Buccales Modifiées. In *Biologie de l'Amérique Australe*, vol. 3 (ed. C. Delamare Deboutteville & E. H. Rapoport), pp. 7–399. Paris: CNRS.

Massoud, Z. 1976 Essai de synthèse sur la phylogénie des Collemboles. *Rev. Ecol. Biol. Sol.* **13**, 241–252.

Mickevich, M. F. & Farris, J. S. 1981 The implications of congruence in *Menidia*. *Syst. Zool.* **30**, 351–370.

Miller, J. S. & Wenzel, J. W. 1995 Ecological characters and phylogeny. *A. Rev. Entomol.* **40**, 389–415.

Pritchard, G., McKee, M. H., Pike, E. M., Scrimgeour, G. J. & Zloty, J. 1993 Did the first insects live in water or air? *Biol. J. Linn. Soc.* **49**, 31–44.

Riek, E. F. 1971 The origin of insects. *Proc. XIII Int. Cong. Ent.* **1**, 292–293.

Sanger, F., Nicklen, S. & Coulson, A. R. 1977 DNA sequencing with chain-terminating inhibitors. *Proc. Natl Acad. Sci. USA* **74**, 5463–5467.

Shear, W. A. & Kukalová-Peck, J. 1990 The ecology of paleozoic terrestrial arthropods: the fossil evidence. *Can. J. Zool.* **68**, 1807–1834.

Thomas, M. A., Walsh, K. A., Wolf, M. R., McPheron, B. A. & Marden, J. H. 2000 Molecular phylogenetic analysis of evolutionary trends in stonefly wing structure and locomotor behavior. *Proc. Natl Acad. Sci. USA* **97**, 13 178–13 183.

Tom, R. B. 1984 Were the first insects terrestrial or aquatic? *S. Afr. Tydskr. Wet.* **80**, 319–323.

Uchida, H. 1971 Tentative key to the Japanese genera of Collembola in relation to the world genera of this order I. *Sci. Rep. Hirosaki Univ.* **18**, 64–76.

Vannier, G. 1973 Originalité des conditions de vie dans le sol due à la présence de l'eau: importance thermodynamique et biologique de la porosphère. *Annls Soc. R. Zool. Belg.* **103**, 157–167.

Vannier, G. 1978 La résistance à la dessication chez les premiers arthropodes terrestres. *Bull. Soc. Ecophysiol.* **3**, 13–42.

Villani, M. G., Allee, L. L., Díaz, A. & Robbins, P. S. 1999 Adaptative strategies of edaphic arthropods. *A. Rev. Entomol.* **44**, 233–256.

Westoll, T. S. 1977 Northern Britain. In *A correlation of the Devonian rocks of the British Isles*, vol. 8 (ed. M. R. House, J. B. Richardson & W. G. Chaloner), pp. 66–93. London: Geol. Soc. Lond. Spec. Rep.

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

Wheeler, W. C. 1995 Sequence alignment, parameter sensitivity, and the phylogenetic analysis of molecular data. *Syst. Biol.* **44**, 321–331.

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

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

Wigglesworth, V. B. 1972 *The life of insects*. London: Weidenfeld & Nicolson.

Winnepenninckx, B., Blackeljau, T. & De Wachter, R. 1993 Extraction of high molecular weight DNA from molluscs. *Trends Genet.* **9**, 407.

Yosii, R. 1961 Phylogenetische bedeuntung der chaetotaxiee bei den Collembolen. *Contr. Biol. Lab. Kyoto Univ.* **12**, 1–37.

As this paper exceeds the maximum length normally permitted, the author has agreed to contribute to production costs.