Some Unusual Small-Subunit Ribosomal RNA Sequences of Metazoans

Authors: GIRIBET, GONZALO, and WHEELER, WARD C.

Source: American Museum Novitates, 2001(3337) : 1-16

Published By: American Museum of Natural History

URL: https://doi.org/10.1206/0003-0082(2001)337<0001:SUSSRR>2.0.CO;2

BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/terms-of-use.

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

Downloaded From: https://complete.bioone.org/journals/American-Museum-Novitates on 12 Sep 2024
Terms of Use: https://complete.bioone.org/terms-of-use

AMERICAN MUSEUM
Novitates

PUBLISHED BY THE AMERICAN MUSEUM OF NATURAL HISTORY
CENTRAL PARK WEST AT 79TH STREET, NEW YORK, NY 10024
Number 3337, 14 pp., 5 figures, 2 tables June 22, 2001

Some Unusual Small-Subunit Ribosomal RNA Sequences of Metazoans

GONZALO GIRIBET[1] AND WARD C. WHEELER[2]

ABSTRACT

The SSU rRNA gene is one of the most widely utilized loci for phylogenetic inference among eukaryotic organisms. Although they have an average length of 1800 to 1900 bp, several unusually large 18S rDNA sequences have been reported. After examining GenBank sequences and 180 new 18S rRNA sequences from several metazoan groups, we report many other extraordinary sequences ranging between ca. 1350 bp (in symphylan myriapods) to ca. 3300 bp (in some strepsipteran insects). Myriapods are particularly interesting, having independently evolved extraordinary sequences in the four classes (Chilopoda, Diplopoda, Symphyla, and Pauropoda). An insertion event of ca. 300 bp has been detected in all but the most basal family of geophilomorphan centipedes. Other major insertions are also found in other arthropod groups, in onychophorans, molluscs, chaetognaths, echinoderms, and parasitic platyhelminths. The use of information derived from secondary structure predictions combined with a new method to analyze DNA sequence data without multiple sequence alignments is proposed as a solution for analyzing sequence data that possess alternatively conservative and variable regions, such as ribosomal genes.

INTRODUCTION

The small-subunit ribosomal RNA gene is one of the most widely utilized loci in phylogenetic inference among eukaryotic organisms (e.g., van de Peer and De Wachter, 1997) especially for the examination of phylogenetic relationships among metazoans. Ribosomal genes play a fundamental role in the synthesis of proteins in eukaryotic and prokaryotic cells. The SSU rRNA locus (in particular the 18S rRNA gene) has been widely used to infer phylogenetic relationships for

[1] Assistant Professor and Assistant Curator, Department of Organismic and Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, MA 02138, USA, e-mail: ggiribet@oeb.harvard.edu
[2] Curator, Division of Invertebrate Zoology, American Museum of Natural History. e-mail: wheeler@amnh.org

Copyright © American Museum of Natural History 2001 ISSN 0003-0082 / Price $1.30

Downloaded From: https://complete.bioone.org/journals/American-Museum-Novitates on 12 Sep 2024
Terms of Use: https://complete.bioone.org/terms-of-use

reasons that have been elaborated elsewhere (e.g., Sogin, 1991; Adoutte and Philippe, 1993). The main rationale is (1) that this gene is long enough to provide phylogenetic information, (2) it is among the slowest evolving sequences found in all living organisms, and (3) different regions of the molecule evolve at different rates. For these reasons the 18S rRNA gene allows the inference of phylogenetic history across a broad taxonomic range. Moreover, the presence of many copies per genome and their homogenization through concerted evolution (Dover, 1982; Hillis and Dixon, 1991) greatly reduce intraspecific variation and facilitate DNA amplification via PCR.

The SSU rRNA gene of most organisms is between 1800 and 1900 bp in total length. However, exceptional cases of long SSU rRNA genes are not uncommon (see Crease and Colbourne, 1998 for some examples), with certain regions being particularly prone to sequence variability (both in nucleotide substitutions and in sequence length).

In some cases, the SSU rRNA gene contains type I introns that are spliced out of the mature rRNA molecule (e.g., De Wachter et al., 1992; Wilcox et al., 1992; Bhattacharya et al., 1994). However, there are many long SSU rRNA genes that do not contain introns (e.g., Hinkle et al., 1994). Some of these long SSU rRNA genes have been reported for metazoans: *Acyrthosiphon pisum* 2469 bp (Kwon et al., 1991); *Xenos vesparum* 3316 bp (Chalwatzis et al., 1995); *Daphnia pulex* 2293 bp (Crease and Colbourne, 1998); *Euperipatoides leuckarti* 2206 bp (Aguinaldo et al., 1997); *Echinococcus granulosus* 2394 bp (Picón et al., 1996). To our knowledge, type I introns have not been found in any metazoan taxon.

Other unusual phenomena have been reported for some organisms. The presence of more than one type of SSU rRNA gene has been found in the protozoan *Plasmodium* (Gunderson et al., 1987; Waters et al., 1989; Qari et al., 1994). Two types of 18S rRNA genes have also been reported in a group of metazoans, the freshwater and terrestrial planarians of the family Dugesiidae (Carranza et al., 1996, 1998a, 1998b).

Because alternate conserved and nonconserved regions are present in the SSU rRNA, sequence alignments are challenging. Some authors have based alignments on information derived from secondary structure models (see Kjer, 1995). However, different models (i.e., Gutell et al., 1985; Hendriks et al., 1988a, 1988b; Neefs and De Wachter, 1990; Van de Peer et al., 1998) can lead to different phylogenetic results (Winnepenninckx and Backeljau, 1996). Moreover, the variable regions cannot be aligned reliably and thus used as phylogenetic information, even if secondary structure predicts that a certain string of nucleotides is homologous (e.g., a variable loop, situated between a conserved stem).

The existence of different models can be attributed to problems in inferring secondary (and tertiary and quaternary) structures via direct observation. This was pointed out by Gutell and collaborators, who acknowledged that "any rigorous search for a secondary structure model for 16-S rRNA would necessitate use of the comparative method" (Gutell et al., 1985: 156). These authors also mentioned that the direct approach of X-ray crystallography remains only a remote possibility, because of the difficult procedure for preparing high-quality crystals of ribosomes or their subunits, together with the added problem that tertiary structure may be directed and/or stabilized by quaternary interactions (Gutell et al., 1985).

Where direct approaches have failed to infer the secondary structure of the SSU rRNA gene, some indirect approaches have succeeded. The comparative method has been successful to the point of building a database of hundreds of secondary structures of SSU rRNA. This database, the SSU ribosomal subunit RNA database (http://rrna.uia.ac.be/rrna/ssu [Van de Peer et al., 1998]) is maintained and updated constantly, and incorporates probably the largest comparative molecular data set.

In the present study, we report some extraordinary examples showing enormous variation in length among 18S rRNA sequences of different metazoan taxa. We also comment on the regions of the molecule that concentrate the largest variation. Finally we explore a novel method that facilitates the use of variable regions of the molecule, which cannot

Downloaded From: https://complete.bioone.org/journals/American-Museum-Novitates on 12 Sep 2024
Terms of Use: https://complete.bioone.org/terms-of-use

be accommodated into standard phylogenetic analyses using sequence alignments.

Materials and Methods

Genomic DNA samples were obtained from fresh, frozen, or ethanol-preserved tissues in a solution of Guanidinium thiocyanate homogenization buffer following a modified protocol for RNA extraction (Chirgwin et al., 1979). The 18S rDNA locus was PCR-amplified in two or three overlapping fragments of about 950, 900, and 850 bp each, using primer pairs 1F–5R, 3F–18Sbi, and 5F–9R, respectively. Primers used in amplification and sequencing were described in Giribet et al. (1996, 1999). Amplification was carried out in a 50 μL volume reaction, with 1.25 units of AmpliTaq DNA Polymerase (Perkin Elmer), 200 μM of d-NTPs, and 1 μM of each primer. The PCR program consisted of a initial denaturing step at 94°C for 60 seconds, 35 amplification cycles (94°C for 15 sec, 49°C for 15 sec, 72°C for 15 sec), and a final step at 72°C for 6 minutes in a GeneAmp PCR System 9700 (Perkin Elmer).

PCR samples were purified with the GENECLEAN III kit (BIO 101 Inc.) and directly sequenced using an automated ABI Prism 377 DNA sequencer. Cycle-sequencing with AmpliTaq DNA Polymerase, FS (Perkin-Elmer) using dye-labeled terminators (ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit) was performed in a GeneAmp PCR System 9700 (Perkin Elmer). Amplification was carried out in a 10 μL volume reaction: 4 μL of Terminator Ready Reaction Mix, 10–30 ng/mL of PCR product, 5 pmoles of primer and dH_20 to 10 μL. The cycle-sequencing program consisted of a step at 94°C for 3 minutes, 25 sequencing cycles (94°C for 10 sec, 50°C for 5 sec, 60°C for 4 min) and a rapid thermal ramp to 4°C and hold. The BigDye-labeled PCR products were isopropanol-precipitated following manufacturer protocol.

Some PCR products that could not be sequenced directly were purified and ligated into pUC 18 *Sma* I/BAP dephosphorylated vector using the SureClone Ligation Kit (Pharmacia P-L Biochemicals) as described in Giribet et al. (1996). Sequencing was then performed by the dideoxy termination method (Sanger et al., 1977) using T7 DNA polymerase (Sequencing Kit from Pharmacia Biotech).

All sequences have been deposited in GenBank (see taxonomy, sequence length, and accession codes in table 1). The new sequences have been compared to other published sequences available from GenBank. The terminology used for the secondary structure topology follows the nomenclature of Van de Peer et al. (1998).

RESULTS AND DISCUSSION

Extraordinary 18S rRNA Sequences of Metazoans

After studying the 18S rRNA gene of about 400 metazoan taxa (180 of these sequences collected by the authors; table 1), we have observed that several animal taxa possess large insertions at different regions of the molecule. This variation is shown in figure 1 and table 1. For example, within the phylum Mollusca, the cephalopods *Loligo pealei* (squid), *Sepia elegans* (cuttlefish), and *Nautilus scrobiculatus* present large insertions in regions V2, V4, V7, and V9. Other groups of molluscs such as the anomalodesmatan bivalves present large insertions in region V7, and some Archaeogastropoda have insertions in regions V2 and V4. But insertions are not restricted to the phylum Mollusca. Sea cucumbers (Echinodermata, Holothuroidea) and arrow-worms (Chaetognatha) present insertions in region V4; some leeches (Annelida, Hirudinea) present insertions in region V7; some parasitic planarians (Platyhelminthes) present insertions both in regions V4 and V7; and velvet worms (Onychophora) present insertions in regions V2, 11, E23–7, V7, and V9. Within the arthropods, there are many extraordinary cases among insects (see the reported strepsipteran sequences by Chalwatzis et al., 1995; 1996; Whiting et al., 1997) and certain crustacean groups. But the most bizarre case within arthropods (and perhaps for the entire Metazoa) are myriapods.

The 18S rRNA Gene of the Myriapods

Myriapods comprise four groups of terrestrial arthropods. Centipedes (class Chilopo-

Downloaded From: https://complete.bioone.org/journals/American-Museum-Novitates on 12 Sep 2024
Terms of Use: https://complete.bioone.org/terms-of-use

TABLE 1

List of the 180 Species of 18S rRNA Sequences Generated by the Authors

(with GenBank accession codes and sequence length [excluding a total of 46 bp from the external primers 1F and 9R]; asterisks refer to noncomplete sequences, and thus the length is not reported)

	GenBank	bp
Annelida (Polychaeta) (3 sp.)		
Eunice torquata	AF123304	1768
Dinophilus gyrociliatus	AF119074	1784
Myzostoma sp.	AF123305	1770
Phylum Sipuncula (3 sp.)		
Aspidosiphon misakiensis	AF119090	1766
Themiste alutacea	AF119075	1757
Phascolopsis gouldii	AF123306	1765
Phylum Echiura (2 sp.)		
Bonellia viridis	AF123307	1787
Urechis sp.	AF119076	1772
Phylum Nemertea (2 sp.)		
Prostoma eilhardi	U29494	1790
Amphiporus sp.	AF119077	1778
Phylum Mollusca (61 sp.)		
Polyplacophora (2 sp.)		
Lepidopleurus cajetanus	AF120502	1761
Acanthochitona sp.	AF120503	1763
Cephalopoda (3 sp.)		
Nautilus scrobiculatus	AF120504	2485
Loligo pealei	AF120505	2221
Sepia elegans	AF120506-7*	
Gastropoda (14 sp.)		
Cocculina messingi	AF120508	1775
Entemotrochus adansonianus	AF120509	1991
Perotrochus midas	AF120510	1986
Haliotis tuberculata	AF120511	1809
Sinezona confusa	AF120512	1810
Diodora graeca	AF120513	1855
Clanculus cruciatus	AF120514	1733
Theodoxus fluviatilis	AF120515	1767
Viviparus georgianus	AF120516	1795
Truncatella guerinii	AF120517	1834
Truncatella sp. 2	AF120518	1827
Balcis eburnea	AF120519	1758
Rissoella caribea	AF120520	2239
Discodoris atromaculata	AF120521	1858
Scaphopoda (2 sp.)		
Dentalium pilsbryi	AF120522	1804
Rhabdus rectius	AF120523	1810
Bivalvia (40 sp.)		
Solemya velum	AF120524	1771
Nucula sulcata	AF120525	1765
Nucula proxima	AF120526	1766
Acila castrensis	AF120527	1765
Yoldia limatula	AF120528	1767
Nuculana minuta	AF120529	1770
Lithophaga lithophaga	AF120530	1767
Striarca lactea	AF120531	1765
Pteria hirundo	AF120532	1775
Lima lima	AF120533	1771
Limaria hians	AF120534	1767
Anomia ephippium	AF120535	1763
Psilunnio littoralis	AF120536	1766
Lampsilis cardium	AF120537	1765
Neotrigonia bednalli	AF120538	1765
Pandora sp.	AF120539	2143
Lyonsia hyalina	AF120540	1986
Cuspidaria cuspidata	AF120541-2*	
Cardiomya costellata	AF120543	1804
Myonera sp.	AF120544	1884
Chama gryphoides	AF120545*	
Codakia cfr. *orbiculata*	AF120546*	
Galeomma turtoni	AF120547	1775
Lasaea sp.	AF120548	1774
Cardita calyculata	AF120549	1727
Cardites antiquata	AF120550	1775
Astarte castanea	AF120551	1775
Dreissena polymorpha	AF120552	1782
Parvicardium exiguum	AF120553*	
Abra sp.	AF120554	1770
Ensis ensis	AF120555	1765
Calyptogena magnifica	AF120556	1777
Corbicula fluminea	AF120557	1777
Sphaerium striatinum	AF120558	1781
Mercenaria mercenaria	AF120559	1779
Mya arenaria	AF120560	1783
Varicorbula dissimilis	AF120561	1795
Gastrochaena dubia	AF120562	1777
Hiatella arctica	AF120563	1774
Bankia carinata	AF120564	1791
Phylum Brachiopoda (1 sp.)		
Argyrotheca cordata	AF119078	1762
Phylum Phoronida (2 sp.)		
Phoronis australis	AF119079	1767
Phoronopsis viridis	AF123308	1765
Phylum Bryozoa (3 sp.)		
Lichenopora sp.	AF119080	1785
Membranipora sp.	AF119081	1761
Caberea boryi	AF119082	1772
Nemertodermatida (1 sp.)		
Meara stichopi	AF119085	1768
Phylum Priapula (1 sp.)		
Tubiluchus corallicola	AF119086	1768
Phylum Onycophora (2 sp.)		
Peripatopsis capensis	AF119087	2174
Epiperipatus biolleyi	AFXXXXXX*	
Phylum Tardigrada (1 sp.)		
Macrobiotus hufelandi	X81442	1762

Downloaded From: https://complete.bioone.org/journals/American-Museum-Novitates on 12 Sep 2024
Terms of Use: https://complete.bioone.org/terms-of-use

TABLE 1—(*Continued*)

	GenBank	bp
Phylum Arthropoda		
Chelicerata (49 sp.)		
Achelia echinata	AF005438	1795
Callipallene sp.	AF005439	1787
Endeis laevis	AF005441	1791
Colossendeis sp.	AF005440	1793
Limulus polyphemus	U91490	1759
Carcinoscorpius rotundicaudatus	U91491	1759
Belisarius xambeui	AF005442	1761
Pseudocellus pearsei	U91489	1763
Ricinoididae sp.	AF124930	1757
Gluvia dorsalis	AF007103	1761
Eusimonia wunderlichi	U29492	1762
Chanbria regalis	AF124931	1763
Stenochrus portoricensis	AF005444	1759
Trithyreus pentapeltis	AF124932	1761
Mastigoproctus giganteus	AF005446	1760
Paraphrynus sp.	AF005445	1760
Amblypygidae sp.	AF124933	1759
Liphistius bicoloripes	AF007104	1762
Nesticus celullanus	AF005447	1763
Roncus cfr. *pugnax*	AF005443	1761
Americhernes sp.	AF124934	1762
Opilioacarus texanus	AF124935	1763
Siro rubens	U36998	1763
Parasiro coiffaiti	U36999	1761
Stylocellus n.sp.	U91485	1763
Dalquestia formosa	AF124936	1761
Odiellus troguloides	X81441	1759
Opilio parietinus	AF124938	1761
Astrobunus grallator	AF124939	1761
Nelima sylvatica	U91486	1762
Leiobunum sp.	AF124940	1761
Hadrobunus cfr. *maculosus*	AF124941	1761
Caddo agilis	U91487	1766
Ischyropsalis luteipes	U37000	1758
Hesperonemastoma modestum	AF124942	1762
Sabacon cavicolens	AF124944	1762
Dicranolasma soerenseni	U37001	1756
Centetostoma dubium	U37002	1758
Nemastoma bimaculatum	AF124947	1758
Equitius doriae	U37003	1762
Triaenobunus sp.	AF124950	1763
Zuma acuta	AF124951	1762
Oncopus cfr. *alticeps*	U91488	1762
Scotolemon lespesi	U37005	1760
Maiorerus randoi	U37004	1763
Bishopella laciniosa	AF124952	1763
Gnidia holnbergii	U37006	1760
Pachyloides thorellii	U37007	1761
Hoplobunus sp.	AF124953	1762
Hexapoda (11 sp.)		
Podura aquatica	AF005452	1761
Acerentulus traeghardi	AF005453	1955
Campodea tillyardi	AF173234	1851
Campodeidae sp.	AF005455*	
Catajapyx sp.	AF005456*	
Dilta littoralis	AF005457	1792
Machiloides sp.	AFXXXXXX	1790
Lepisma sp.	AF005458	1785
Thermobius sp.	AFXXXXXX	1788
Tricholepidion gertschi	AFXXXXXX	1809
Myriapoda (38 sp.)		
Cylindroiulus punctatus	AF005448	1785
Polydesmus coriaceus	AF005449	1783
Scutigerella sp1	AF007106	1299
Scutigerella sp.2	AF005450*	
Hanseniella sp.	AF173237*	
Pauropodidae sp.	AF005451	2182
Scutigera coleoptrata	AF000772	1819
Thereuopoda clunifera	AF119088	1817
Allothereua maculata	AF173240*	
Lithobius variegatus	AF000773	1814
Australobius scabrior	AF173241	1815
Paralamyctes n.sp.	AF173242	1818
Lamyctes emarginatus	AF173244	2099
Henicops maculatus	AF173245	2231
Anopsobius n. sp.	AF173247	1944
Craterostigmus tasmanianus	AF000774	1814
Scolopendra cingulata	U29493	1841
Cormocephalus monteithi	AF173249	1842
Ethmostigmus rubripes	AF173250	1844
Alipes sp.	AF173251	1844
Rhysida nuda	AF173252	1846
Cryptops trisulcatus	AF000775	1819
Theatops erythrocephala	AF000776	1818
Scolopocryptops nigridus	AF173253	1817
Mecistocephalus sp.	AF173254	1820
Pseudohimantarium mediterraneum	AF000778	2157
Henia (Chaetechelyne) vesuviana	AF173255	2194
Pectiniunguis argentinensis	AF173256	2006
Schendylops pampeanus	AF173257	2053
Ballophilus australiae	AF173258	2108
Clinopodes cfr. *poseidonis*	AF000777	2224
Tasmanophilus sp.	AF173259	1930
Tuoba sydneyensis	AF173260	2083
Zelanion antipodus	AF173261	2194
Zelanion sp.	AF173262	2224
Ribautia n. sp.	AF173263	2218
Aphilodon weberi	AF173264	2015
Strigamia maritima	AF173265	2122
Phylum Enteropneusta (1 sp.)		
Glossobalanus minutus	AF119089	1776

Downloaded From: https://complete.bioone.org/journals/American-Museum-Novitates on 12 Sep 2024
Terms of Use: https://complete.bioone.org/terms-of-use

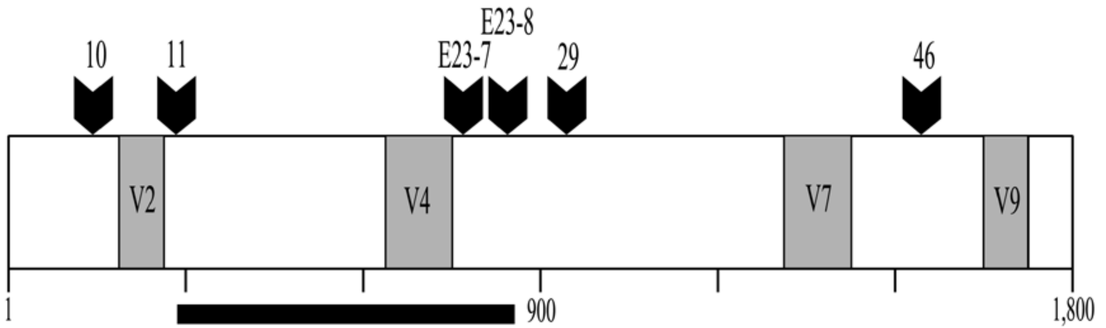

Fig. 1. Schematic representation of the 18S rRNA locus. The gray squares represent the variable regions V2, V4, V7, and V9 with insertions (V2: Onychophora, Geophilomorpha, Cephalopoda, Archaeogastropoda; V4: Hexapoda, Crustacea, Pauropoda, Holothuroidea, Chaetognatha, Platyhelminthes, Cephalopoda; V7: Onychophora, Hexapoda, Crustacea, Pauropoda, Chilopoda, Platyhelminthes, Hirudinea, Cephalopoda, Gastropoda; V9: Onychophora, Crustacea, Cephalopoda). The black arrowheads represent particular insertions (10: Pauropoda; 11: Onychophora; E23–7: Onychophora and Pauropoda; E23–8: Pauropoda; 29: Pauropoda; 46: Protura). The black bar represents the 500 bp deletion of the Symphyla.

da) and millipedes (class Diplopoda) are the two principal classes of myriapods. The two other classes of myriapods lack common names: Symphyla and Pauropoda. Prior to the analysis of Edgecombe et al. (1999), the only complete 18S rRNA sequence data for myriapods available at GenBank were eight centipedes and two millipedes (Giribet et al., 1996, 1999; Giribet and Ribera, 1998). Two centipede species of the order Geophilomorpha (*Clinopodes poseidonis* and *Pseudohimantarium mediterraneum*) present an insertion of about 300 bp at region V7, whereas all the other available sequences are fairly conserved in terms of primary sequence. However, a wider ongoing study on the 18S rRNA gene of myriapods suggests that they constitute one of the most interesting cases of 18S rRNA variation in any metazoan group.

Within the centipedes, the members of the order Geophilomorpha (15 species studied belonging to 9 families), excluding two species of the most basal family Mecistocephalidae (Edgecombe et al., 1999), exhibit insertions of about 300 bp in the region V7. This is an unusual example that shows exactly when the insertion occurred during the phylogenetic process, and illustrates the putative information of such insertions (fig. 2).

Within the millipedes, members of the family Polyzonidae display sequences longer than 2700 bp. Data for one species of Pauropoda show that the 18S rRNA is ca. 2200 bp, with several small insertions (Giribet, 1997; Giribet and Ribera, 2000). But perhaps the most unusual case among metazoan 18S rRNA sequences is the Class Symphyla. Amplification of the 18S rRNA loci of three species belonging to two genera (two species of *Scutigerella* from northeastern Spain and the Canary Islands, respectively, and one species of *Hanseniella* from Australia) yielded a product band size of about 1350 bp. Sequencing this fragment suggests a deletion of about 500 bp in the central region of the molecule.

Although it might be conjectured that in this case a nonfunctional pseudogene has been sequenced, as occurred with the 18S rRNA locus of the platyhelminth *Dugesia mediterranea* (Carranza et al., 1996) and other dugesiids (Carranza et al., 1998a, 1998b), this seems improbable for several reasons. First, this sequence has been obtained from three different species and in two independent laboratories. Second, none of the highly conserved primers from the ‘‘deleted’’ region (forward primers 4F, 18Sa0.7, 18Sa0.79, 18Sa1.0; reverse primers 4R, 18Sb5.0, 18Sb3.9, 18Sb3.0 [Giribet et al., 1996; Whiting et al., 1997]) amplified any DNA fragment when combined with primers from other regions. If the 1350 bp fragment was a

Downloaded From: https://complete.bioone.org/journals/American-Museum-Novitates on 12 Sep 2024
Terms of Use: https://complete.bioone.org/terms-of-use

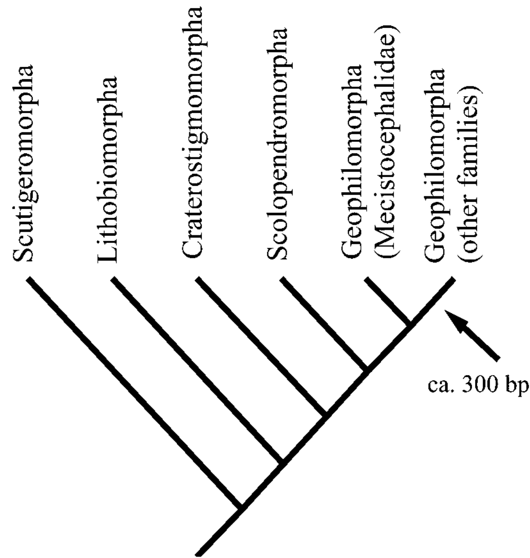

Fig. 2. Phylogenetic tree of the centipedes based on the combined analysis of Edgecombe et al. (1999). The arrow indicates where the insertion of ca. 300 bp at region V7 occurred during the evolution of centipedes.

pseudogene, we would expect to amplify fragments of the original gene when using the conserved primers located within the "deleted" region. Third, phylogenetic analyses including the symphylan sequences show that symphylans are arthropods related to other myriapods (Giribet, 1997; fig. 3). Fourth, amplification of DNA from an RNA source, as described by Carranza et al. (1998a), yielded a product band size of approximately 1350 bp, as expected. These facts demonstrate that a deletion of ca. 500 bp occurred in the common ancestor of these three symphylan species.

18S rRNA Variation in Metazoans

It seems that large insertions and deletions are not as constrained as was previously thought (e.g., Crease and Colbourne, 1998). These events occur in many metazoan taxa, and are commonly in regions V2, V4, V7, and V9. Other parts of region 23 (that includes the region V4) are also variable. In a phylogenetic study of about 150 arthropod 18S rRNA sequences (Giribet and Ribera, 2000), insertions at region E23–7 were observed in Onychophora and in Pauropoda while insertions at region E23–8 were observed in the pauropod species. Other insertions observed in particular taxa occur at sites 8 (in the millipede *Polyzonium*), 11 (in Onychophora), 29 (in Pauropoda), and 46 (in the proturan *Acerentulus traeghardi*). However, we only obtained sequences from one pauropod and one proturan, hence these results cannot be generalized to other members of such groups.

Certain taxa present insertions in variable regions whereas in the remaining regions the primary sequence may be conserved. For example, certain geophilomorph centipedes exhibit a small insertion at region V2 (between 10 and 80 bp compared to other centipedes) and a large insertion (about 300 bp) at region V7, whereas the remaining positions are conserved with respect to other centipedes. Other taxa not only present insertions in the variable regions, but also in the primary sequence. This is the case in the cephalopods, which have insertions in regions V2, V4, V7, and V9, and differ considerably from other molluscs in the primary sequence of the remaining regions.

Although several metazoans present insertions in the 18S rRNA, reduction of the 18S rRNA gene appears to be a rare event in evolution. To our knowledge, there are no other reported cases of 18S rDNA sequences with

Downloaded From: https://complete.bioone.org/journals/American-Museum-Novitates on 12 Sep 2024
Terms of Use: https://complete.bioone.org/terms-of-use

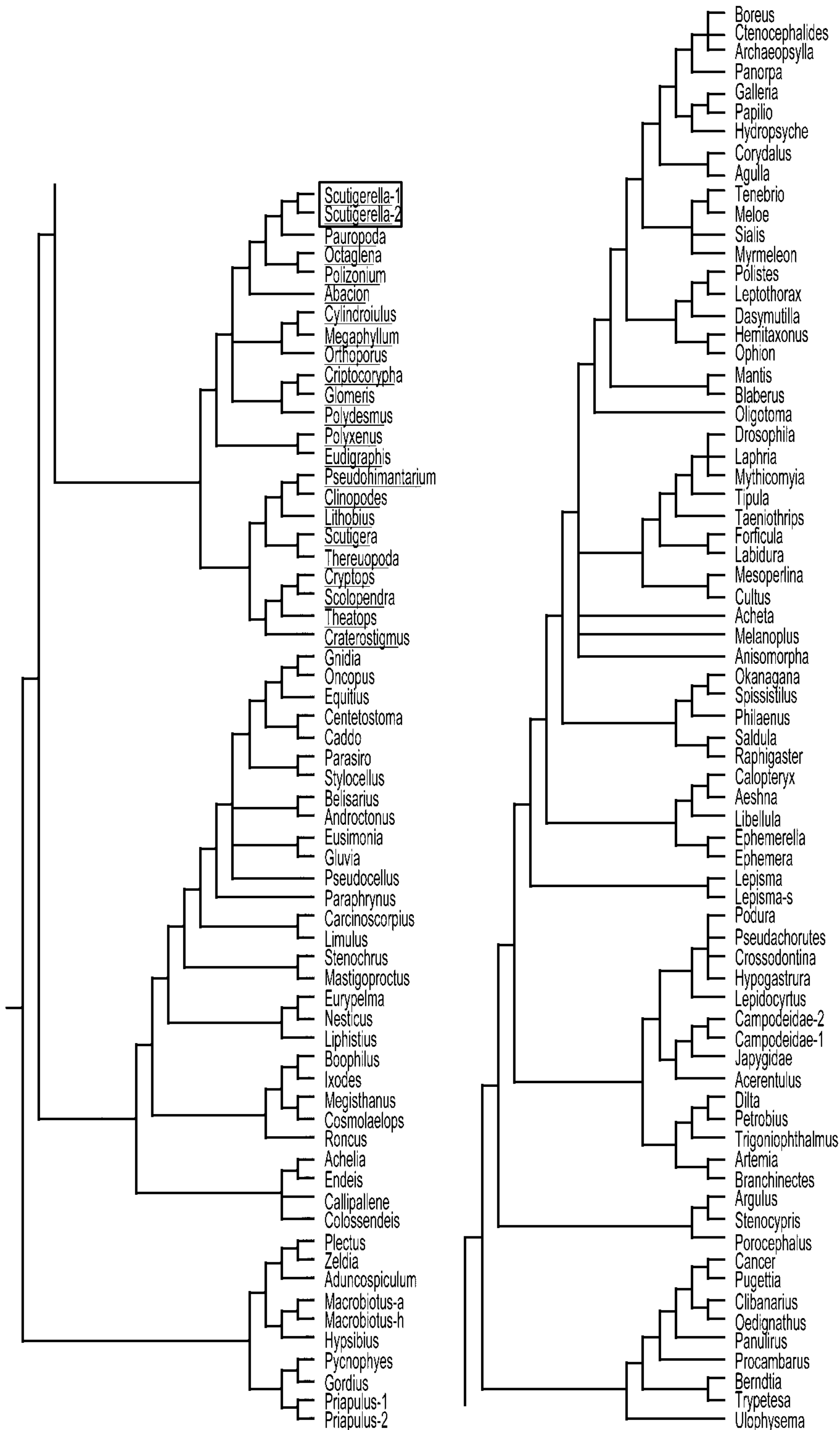

Fig. 3. Phylogenetic tree based on 18S rRNA sequence data indicating the position of two symphylans (box) with respect to other myriapods (underlined taxa) in a phylogenetic analysis of arthropods (from Giribet, 1997). The two symphylans appear related to other myriapods.

Downloaded From: https://complete.bioone.org/journals/American-Museum-Novitates on 12 Sep 2024
Terms of Use: https://complete.bioone.org/terms-of-use

a deletion of the magnitude of that observed in the symphylans (ca. 500 bp). The deletion corresponds to the central region of the molecule (approximately from region 14 to E23–9). The remaining primary sequence is fairly conserved, which we assume may be functional. However reconstruction of a global secondary structure cannot be conducted using the comparative method. Another case of sequence length reduction in the 18S rRNA locus occurs in the Dicyemid mesozoans (three species of the genus *Dicyema*), with a total gene sequence of about 1670 bp that presents two major deletions at the variable sites V2 and V9.

The geophilomorphan centipedes that present the insertion of about 300 bp at region V7 (13 species) also display a large insertion (about 300 bp) at the D3 expansion fragment of the large subunit rRNA locus (28S rRNA). Neither of the insertions at region V7 of the 18S rRNA or at the D3 expansion fragment of the 28S rRNA locus have been found in the putative most basal geophilomorph family Mecistocephalidae (Giribet et al., 1999; Edgecombe et al., 1999). This apparent correlation between the insertions at region V7 of the 18S rRNA locus and at the D3 expansion fragment of the 28S rRNA locus (also observed in the cephalopod *Loligo*) could suggest a possible interaction between these subunits in the ribosome.

18S rRNA Variable Sites and Phylogenetic Analyses

The presence of certain variable sites in the 18S rRNA molecule can hinder phylogenetic analyses at the alignment step, a fact that is promoted by several researchers who avoid the use of ribosomal genes for drawing inferences (e.g., Ayala et al., 1998). In general, researchers using ribosomal genes exclude the variable regions from their phylogenetic inference step because of uncertainties in alignments (e.g., Giribet et al., 1996, 1999). But since data removal from phylogenetic analyses is also problematic (see Gatesy et al., 1993), and automatic alignments are explicit, other researchers prefer the use of automatic alignments exploring different cost matrices (e.g., Wheeler, 1995). Regardless of which of these options is the best (philosophically, computationally, or practically), the extreme length variation of some of the SSU rRNA helices may constitute a serious problem in the phylogenetic analysis of certain data sets.

A clear example is illustrated in figure 4, where we show the variable region V7 of 17 centipede taxa (see table 2 for the taxonomy). The fragment as a whole can be considered homologous because it is located between two conserved regions that constitute a stem. The most common sequence length for the V7 region in centipedes ranges between 65 and 70 bp (character state present in the five recognized orders of centipedes). However three taxa (*Scolopendra, Ethmostigmus,* and *Alipes*; family Scolopendridae) display sequences between 93 and 94 bp, and four taxa (*Pseudohimantarium, Henia, Clinopodes,* and *Zelanion*; belonging to four families of Geophilomorpha) exhibit sequences between 354 and 384 bp. These two clades are well defined morphologically and could also be characterized in terms of sequence length or perhaps secondary structure topology. However, a standard phylogenetic analysis of this fragment cannot be conducted successfully when base-to-base homology is required, as is the case with multiple sequence alignments. This situation is frustrating since the sequence data clearly display historical information that cannot be used phylogenetically.

A new method to analyze DNA sequence data that does not require base-to-base correspondences was recently developed by Wheeler (1999) and is discussed in greater detail there. Briefly, the method, named ‘‘fixed character states’’, optimizes DNA sequence data without employing multiple sequence alignments by treating entire homologous stretches of sequence data as characters. The set of specific sequences exhibited by the terminal taxa constitutes the character states. Thus the number of states is equal to the number of unique sequences (or homologous fragments) exhibited by the data. In the example illustrated here, there is one character (region V7) with 17 states (as many as different taxa). Other situations could arise where the number of states would be smaller than the number of taxa if two or more were to share identical sequences. The salient fea-

Downloaded From: https://complete.bioone.org/journals/American-Museum-Novitates on 12 Sep 2024
Terms of Use: https://complete.bioone.org/terms-of-use

```
Scutigera
       1 ACGATCGATT TAGGCGAGCT GTTTCCTTCC CCCGNGGNTG GAGCGGCACT GCCTCCGTCG GTCGACAA

Thereuopoda
       1 ACGATCGATT TGGGCGAGCT GTTTCCTGCC TTCACGGTAG GAGCGGCACT GCCTCCGTTG GTCGATAA

Lithobius
       1 ACGACTGATC CCGGGGTGCC GGGCCCTCTT CGNGGGGGAA CGGTGTTGCC TCCGTCAGTT GTTCG

Australobius
       1 ACGACCGATC CCGGGGTGCC GGTGCCCTCT TCGGGGGGAA CGGTGTTGCC TCTGTCGGTT GATCG

Craterostigmus
       1 ACGACCGATC CCGGGGTGCC GTCTCCTTCC TCGTGATGGA GCGGCGTTAC CTCCGTCGGC CGATCG

Scolopendra
       1 ACGTCCGATC CTGGGGTGCC GGTGCCTCCT AAACCTCCGC TCTTTCGAAA AGAGTGGGAG GCGGGGGAAC GGCTTTGCCT CTGTCGGACG ATTG

Ethmostigmus
       1 ACGTCCGATC TCGGGGTGCC GGTGCCTCCT AAACCCCCCC TTCTTCGATG GAGCGGGGGG CGGGGGAACG GCTTTGCCTC TGTCGGATGA TCG

Alipes
       1 ACATTCGATC TCGGGGTGCT GGCACCTCCT ACACCCCCCC TTCTTTCATG GAGTGGGGGG CGGGGGAACA GCTTTGCCTC TGTCGGATGA TCG

Cryptops
       1 ACGTCCGATC TCGAGGTGCC GTTACCTTtC TCCTCGTGAG GGGTtCGGCT TTGCCTCTGT CCGACGATTA

Theatops
       1 ACGTCCGATC TCGGGGTGCC GTCTCCTTCT CCTCGMGAGA GGTGCGGCTT TGCCTCTGTC GGACGATTG

Scolopocryptops
       1 ACGTCCGATC TCGGGGTGCC GTTGCCATCT CCTCGTGAGG GGCTCGGCCT GACCTCTGTT GGACGATTG

Mecistocephalus
       1 ACGACTGGTC TTGGGGTGCT GGTTCTATTC CTTCATGGGT AGCCAGCTTT TGCCTCCGTC GGTCGATTT

Nodocephalus
       1 ACGACTGATC TCGGGGTGCT GGATCTATTC CTTCGTGGAT AGCCGGTAGT TGCCTCTGTT CGTCGTCCGA

Pseudohimantarium
       1 ACGACTGATC CCGGGGTGCC GGTGTCCCCC CTTCTGTCGC TTTAATTTTT TGTCTGCGGC ATGTTGCCGT TTGCTTTCTT GGGTGTATCT TGCTGATCCC
     101 YTACTCATGT TTTCTATCAC CTCCCCACTC ATCGAGTGTT TTGCGGCTGG TTTCTGCCTC TGGTCGGTAT TGCATTTACG CCRTCGCGGG TATCGTGCGT
     201 ATCGGTCATG TGGTTGCCTC GTTGCTTTCG CTCTGTGTTG TGGTGTGTGT GTGTGTGGGG CGTGTTGAGG GCAAAGGCAT TATGATTCTC GAGAGGAGTA
     301 GCGGTCTTGC GGGACATTAA TGGGCGGTCA GGCTAGGTGG GGTCGCACAC GGCGTTGCYT CTGTCAGTCG ACAGG

Henia
       1 ACGACCGATC CCGGGGTGCC GGTGCTCCCA TCTTGCTTCT GTTTGTCCGT TTTTCGCTCA GGCGACTCTC GCGTCTCGCC TCTCTCTGCG GACGTTCGGT
     101 TCAACTCGGC GCGGCAGAGG CCACGCTCTC GACTTCCCCT CGTGTTTCGG CGGTTGCGTC GAGGTTGTCT CCGTGTCTCC CTTTCTCGGA AGGGCTGCGC
     201 GGCGGTCGGC CTTCGTCGTC ATCGTTCTGG CCGGGGGTGC ATGTTCGCGG GCGGGGATGA AACCGCTTCG AGCGCCGAGC GTCGGGAGTC CGGGCGTTTC
     301 TGCAGATAGC CGAAAGCGGG TTCGGGCGGC AGTTGTCGTT AGCATGGGAA AGCACCCGGC AGTTGCCTCC GTCGGTCGAA TACG

Clinopodes
       1 ACGACCGATC CCGGGGTGCC GGCGCTCCCA TCTWGMTTCT GTTTGTTTCT GTCCGGCGYT GAGACGCTCT TTGTGGCTGC TGCGATTTTC TTGCTCCTCC
     101 GGGTTTTTTC TWTCNCCYGC CGGGCGGGAR AAAGAARGGT CTGATGCGTG GGGTGTGTGT GTGAAGCTGT CTTTYTCTCT CTTTCGAGGG GGGAGAGGCT
     201 GTCTTTTTCT CGCTCGCTCT AAGCTCTTTC CCYTTTCTTT CTTTCCGTTT ACGGAGGTGA TKGATATTTT CTCGGCTCTC GAGCGGGTCT TTCAAGGCGC
     301 ACGAATCGTC SACTGCGTTG GTCGAGACGG CAGTTGTCGT TAGCATGGGT CAGCGTGTGG CGTTGCCTCT GTCGGTCGAW TACG

Zelanion
       1 ACGACCGATC CCGGGGTGCC GGCGCCTTCT CGTCTTGCGG CCGTTTCGCT CCCCTCTCGG GTTTCGCGTC GTTCTCTTTT CCCTTCCTCA CATCCGTCGC
     101 GGGGAGGCGG TGCGTGCGTT TCACTCGCGG TAGATGTCGC GGAGTTGGCG GAATAAGGCT GCCGCTCTCC CTTTCGGGGG TTGAGTGGGG TCCCGTCGCT
     201 CCGTTTCACT CTCGGGCGTA CGCGCGCATT CGTTTYTCCC TTCTTCGGGG ACTGGGGCTC GAGGGATCGA GGCGCGCGGG ACTTTGCTGA GGGACCGGCG
     301 ACGGTTGTCA TTTGCGCGAG TCGGCTGTCC GGCGTTGCCT CTGTCGGTCG ATCG
```

Fig. 4. Variable region (V7) of the 18S rRNA locus of 17 species of centipedes.

ture of this method is the treatment of length variation. Since the sequence variation is expressed through a series of transformations between states, indel or "gap" variation only occurs as transformations between sequences, not globally among all the sequence data. This has the effect of moderating the difficulties presented by extreme nucleotide length variation at the expense of treating strings of bases as character states instead of individual nucleotides.

A matrix of transformation costs is created to relate the states to one another. The cells of this matrix are defined as the minimum transformation cost required between each pair of states based on insertion-deletion and base substitution costs (as in the calculation of an alignment score). The next operation

Downloaded From: https://complete.bioone.org/journals/American-Museum-Novitates on 12 Sep 2024
Terms of Use: https://complete.bioone.org/terms-of-use

TABLE 2
Taxonomy of the Centipede Species (Chilopoda) Represented in Figures 4 and 5

Order Scutigeromorpha	
F. Scutigeridae	*Scutigera coleoptrata*
	Thereuopoda clunifera
Order Lithobiomorpha	
F. Lithobiidae	*Lithobius variegatus*
	Australobius scabrior
Order Craterostigmomorpha	
F. Craterostigmidae	*Craterostigmus tasmanianus*
Order Scolopendromorpha	
F. Scolopendridae	*Scolopendra cingulata*
	Ethmostigmus rubripes
	Alipes crotalus
F. Cryptopidae	*Cryptops trisulcatus*
	Theatops erythrocephala
	Scolopocryptops nigridus
Order Geophilomorpha	
F. Mecistocephalidae	*Mecistocephalus* sp.
	Nodocephalus doii
F. Himantariidae	*Pseudohimantarium mediterraneum*
F. Dignathodontidae	*Henia (Chaetechelyne) vesuviana*
F. Geophilidae	*Clinopodes* cfr. *poseidonis*
F. Chilenophilidae	*Zelanion antipodus*

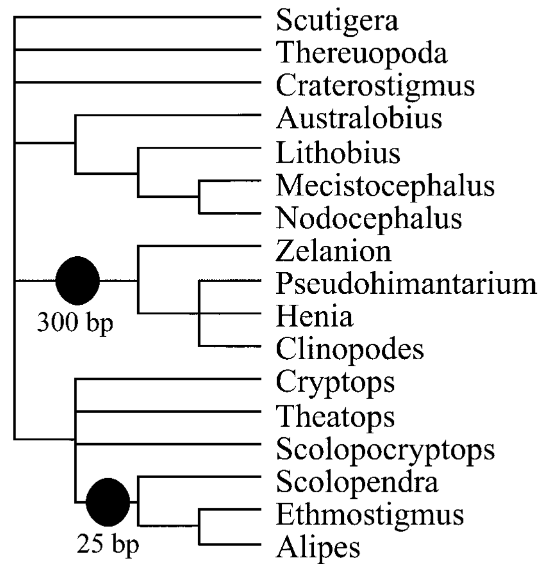

Fig. 5. Phylogenetic analysis of the data from fig. 4 using the "fixed character states" method of Wheeler (1999) implemented in the computer program POY (Gladstein and Wheeler, 1997). Commands: poy -fixedstates -noleading -norandomizeoutgroup -gap 1 -maxtrees 20 -multibuild 10 -seed−1 -slop 2 -checkslop 5. The two circles illustrate the insertions of the Geophilomorpha (ca. 300 bp), and the Scolopendridae (ca. 25 bp).

uses this transformation matrix to diagnose a specific phylogenetic topology by means of existing dynamic programming techniques (Sankoff and Rousseau, 1975) with the number of states greatly expanded. This method has been implemented in the computer program POY (Gladstein and Wheeler, 1997) specifying the option **-fixedstates** (available via anonymous ftp at the site ftp.amnh.org / pub/molecular/poy/).

To illustrate how the method works empirically, we have analyzed the 17 sequences presented in figure 4 (and table 2) using the fixed character states method. The tree obtained (fig. 5) shows some lack of resolution, but also shows certain clades highly consistent with the current morphology of the group, as well as with the molecular analyses of Giribet et al. (1999) and Edgecombe et al. (1999). Scolopendromorpha is recognized as a clade that in turn includes a monophyletic clade, the family Scolopendridae, presenting an insertion of about 25 bp. Another clade is defined by the insertion of about 300 bp that groups all geophilomorph species except the mecistocephalids (*Mecistocephalus* and *Nodocephalus,* the most basal group that lacks the insertion). This is encouraging considering that this topology corresponds to the analysis of just a few bases from the variable region V7. Thus, this method facilitates use of all the information (variable and conserved) from ribosomal genes.

CONCLUSIONS

Long SSU rRNA appears to be more common than claimed by some authors (e.g., Crease and Colbourne, 1998) based on its occurrence in at least seven metazoan phyla: Platyhelminthes, Mollusca, Onychophora, Arthropoda, Chaetognatha, Echinodermata, and Mesozoa. However, large deletions appear to be rare, having so far only been found in one group of arthropods (Symphyla) and in mesozoans. Probably many more taxa dis-

Downloaded From: https://complete.bioone.org/journals/American-Museum-Novitates on 12 Sep 2024
Terms of Use: https://complete.bioone.org/terms-of-use

play extraordinary 18S rRNA genes, but this will not be discovered until sampling within each phylum increases. Nonetheless, the existence of variable regions should not discourage the use of ribosomal genes in phylogenetic analyses, especially when secondary structure predictions are combined with novel methods of DNA sequence data analysis. In this sense, the characterization of secondary structural features by means of the comparative method, and the use of these features (homologous regions) as characters with multiple states provides a powerful approach for the analysis of such data using the fixed character states method. Maybe it is at such levels that secondary structure information can best contribute to phylogenetic analyses.

ACKNOWLEDGMENTS

We thank Jerry Harasewych for discussion and Lorenzo Prendini and Michael Whiting for many helpful suggestions. G. G. was supported by a Lerner Gray research fellowship from the AMNH.

REFERENCES

Adoutte, A., and H. Philippe
1993. The major lines of metazoan evolution: summary of traditional evidence and lessons from ribosomal RNA sequence analysis. *In* Y. Pichon (ed.), Comparative molecular neurobiology: 1–30. Basel: Birkhäuser Verlag.

Aguinaldo, A. M. A., J. M. Turbeville, L. S. Lindford, M. C. Rivera, J. R. Garey, R. A. Raff, and J. A. Lake
1997. Evidence for a clade of nematodes, arthropods and other moulting animals. Nature 387: 489–493.

Ayala, F. J., A. Rzhetsky, and F. J. Ayala
1998. Origin of the metazoan phyla: Molecular clocks confirm paleontological estimates. Proc. Natl. Acad. Sci. USA 95: 606–611.

Bhattacharya, D. B., B. Surek, M. Rüsing, S. Damberger, and M. Melkonian
1994. Group I introns are inherited through common ancestry in the nuclear-encoded rRNA of Zygnematales (Charophyceae). Proc. Natl. Acad. Sci. USA 91: 9916–9920.

Carranza, S., G. Giribet, C. Ribera, J. Baguñà, and M. Riutort
1996. Evidence that two types of 18S rDNA coexist in the genome of *Dugesia* (*Schmidtea*) *mediterranea* (Platyhelminthes, Turbellaria, Tricladida). Mol. Biol. Evol. 13: 824–832.

Carranza, S., D. T. J. Littlewood, K. A. Clough, I. Ruiz-Trillo, J. Baguñà, and M. Riutort
1998a. A robust molecular phylogeny of the Tricladida (Platyhelminthes, Seriata) with a discussion on morphological synapomorphies. Proc. R. Soc. London B 265: 631–640.

Carranza, S., I. Ruiz-Trillo, D. T. J. Littlewood, M. Riutort, and J. Baguñà
1998b. A reappraisal of the phylogenetic and taxonomic position of land planarians (Platyhelminthes, Turbellaria, Tricladida) inferred from 18S rDNA sequences. Pedobiologia 42: 433–440.

Chalwatzis, N., A. Baur, E. Stetzer, R. Kinzelbach, and F. K. Zimmermann
1995. Strongly expanded 18S rRNA genes correlated with a peculiar morphology in the insect order of Strepsiptera. Zoology 98: 115–126.

Chalwatzis, N., J. Hauf, Y. Van de Peer, R. Kinzelbach, and F. K. Zimmermann
1996. 18S ribosomal RNA genes of insects: Primary structure of the genes and molecular phylogeny of the Holometabola. Ann. Entomol. Soc. Am. 89: 788–803.

Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter
1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry 18: 5294–5299.

Crease, T. J., and J. K. Colbourne
1998. The unusually long small-subunit ribosomal RNA of the crustacean, *Daphnia pulex*: sequence and predicted secondary structure. J. Mol. Evol. 46: 307–313.

De Wachter, R., J. M. Neefs, A. Goris, and Y. Van de Peer
1992. The gene coding for small ribosomal subunit RNA in the basidiomycete *Ustilago maydis* contains a group I intron. Nucleic Acids Res. 20: 1251–1257.

Dover, G. A.
1982. Molecular drive: a cohesive model of species evolution. Nature 299: 111–117.

Edgecombe, G. D., G. Giribet, and W. C. Wheeler
1999. Phylogeny of Chilopoda: Combining 18S and 28S rRNA sequences and morphology. Boletin S.E.A. 26: 293–331.

Downloaded From: https://complete.bioone.org/journals/American-Museum-Novitates on 12 Sep 2024
Terms of Use: https://complete.bioone.org/terms-of-use

Gatesy, J., R. DeSalle, and W. C. Wheeler
1993. Alignment-ambiguous nucleotide sites and the exclusion of systematic data. Mol. Phylogenet. Evol. 2: 152–157.
Giribet, G.
1997. Filogenia molecular de Artrópodos basada en la secuencia de genes ribosomales. Ph.D. thesis, Universitat de Barcelona, Barcelona, 221 pp.
Giribet, G., and C. Ribera
1998. The position of arthropods in the animal kingdom: a search of a reliable outgroup for internal arthropod phylogeny. Mol. Phylogenet. Evol. 9: 481–488.
2000. A review of arthropod phylogeny: new data based on ribosomal DNA sequences and direct character optimization. Cladistics 16: 204–231.
Giribet, G., S. Carranza, J. Baguñà, M. Riutort, and C. Ribera
1996. First molecular evidence for the existence of a Tardigrada + Arthropoda clade. Mol. Biol. Evol. 13: 76–84.
Giribet, G., S. Carranza, M. Riutort, J. Baguñà, and C. Ribera
1999. Internal phylogeny of the Chilopoda (Arthropoda, 'Myriapoda'): a combined approach using complete 18S rDNA and partial 28S rDNA sequences. Phil. Trans. R. Soc. London B 354: 215–222.
Gladstein, D. S., and W. C. Wheeler
1997. POY: the Optimization of Alignment Characters. New York: American Museum of Natural History.
Gunderson, J. H., M. L. Sogin, G. Wollett, M. Hollingdale, V. F. de la Cruz, A. P. Waters, and T. F. McCutchan
1987. Structurally distinct, stage-specific ribosomes occur in *Plasmodium.* Science 238: 933–937.
Gutell, R. R., B. Weiser, C. R. Woese, and H. F. Noller
1985. Comparative anatomy of 16S-like ribosomal rRNA. Progr. Nucl. Acids. Res. Mol. Biol. 32: 155–216.
Hendriks, L., R. De Baere, C. van Broeckhoven, and R. De Wachter
1988a. Primary and secondary structure of the 18S ribosomal RNA of the insect species *Tenebrio molitor.* FEBS Lett. 232: 115–120.
Hendriks, L., C. van Broeckhoven, A. Vandenberghe, Y. Van de Peer, and R. De Wachter
1988b. Primary and secondary structure of the 18S ribosomal RNA of the bird spider *Eurypelma californica* and evolutionary relationships among eukaryotic phyla. Eur. J. Biochem. 177: 15–20.
Hillis, D. M., and M. T. Dixon
1991. Ribosomal DNA: molecular evolution and phylogenetic inference. Q. Rev. Biol. 66: 411–453.
Hinkle, G., D. D. Leipe, T. A. Nerad, and M. L. Sogin
1994. The unusually long small subunit ribosomal RNA of *Phreatamoeba balamuthi.* Nucleic Acids Res. 22: 465–469.
Kjer, K. M.
1995. Use of rRNA secondary structure in phylogenetic studies to identify homologous positions: An example of alignment and data presentation from the frogs. Mol. Phylogenet. Evol. 4: 314–330.
Kwon, O. Y., K. Ogino, and H. Ishikawa
1991. The longest 18S ribosomal RNA ever known. Nucleotide sequence and presumed secondary structure of the 18S rRNA of the pea aphid, *Acyrthosiphon pisum.* Eur. J. Biochem. 202: 827–833.
Neefs, J. M., and R. De Wachter
1990. A proposal for the secondary structure of a variable area of eukaryotic small ribosomal subunit RNA involving the existence of a pseudoknot. Nucleic Acids Res. 18: 5695–5704.
Picon, M., R. R. Gutell, R. Ehrlich, and A. Zaha
1996. Characterization of a flatworm ribosomal RNA-encoding gene: promoter sequence and small subunit rRNA secondary structure. Gene 171: 215–220.
Qari, S. H., I. F. Goldman, N. J. Pieniazek, W. E. Collins, and A. A. Lal
1994. Blood and sporozoite stage-specific small subunit ribosomal RNA-encoding genes of the human malaria parasite *Plasmodium vivax.* Gene 150: 43–49.
Sanger, F, S. Nicklen, and A. R. Coulsen
1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA 74: 5463–5468.
Sankoff, D., and P. Rousseau
1975. Locating the vertices of a Steiner tree in arbitrary space. Math. Prog. 9: 240–246.
Sogin, M. L.
1991. Early evolution and the origin of eukaryotes. Curr. Opin. Genet. Dev. 1: 457–463.
Van de Peer, Y., and R. De Wachter
1997. Evolutionary relationships among the eukaryotic crown taxa taking into ac-

Downloaded From: https://complete.bioone.org/journals/American-Museum-Novitates on 12 Sep 2024
Terms of Use: https://complete.bioone.org/terms-of-use

count site-to-site rate variation in 18S rRNA. J. Mol. Evol. 45: 619–630.

Van de Peer, Y., A. Caers, P. De Rijk, and R. De Wachter
1998. Database on the structure of small ribosomal subunit RNA. Nucleic Acids Res. 26: 181–185.

Waters, A. P., C. Syin, and T. F. McCutchan
1989. Developmental regulation of stage-specific ribosome populations in *Plasmodium.* Nature 342: 438–440.

Wheeler, W. C.
1995. Sequence alignment, parameter sensitivity, and the phylogenetic analysis of molecular data. Syst. Biol. 44: 321–331.
1999. Fixed character states and the optimization of molecular sequence data. Cladistics 15: 379–385

Whiting, M. F., J. M. Carpenter, Q. D. Wheeler, and W. C. Wheeler
1997. The Strepsiptera problem: phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology. Syst. Biol. 46: 1–68.

Wilcox, L. W., L. A. Lewis, P. A. Fuerst, and G. L. Floyd
1992. Group I introns within the nuclear-encoded small-subunit rRNA gene of three green algae. Mol. Biol. Evol. 9: 1103–1118.

Winnepenninckx, B., and T. Backeljau
1996. 18S rRNA alignments derived from different secondary structure models can produce alternative phylogenies. J. Zool. Syst. Evol. Res. 34: 135–143.

Downloaded From: https://complete.bioone.org/journals/American-Museum-Novitates on 12 Sep 2024
Terms of Use: https://complete.bioone.org/terms-of-use

Recent issues of the *Novitates* may be purchased from the Museum. Lists of back issues of the *Novitates* and *Bulletin* published during the last five years are available at World Wide Web site http://nimidi.amnh.org. Or address mail orders to: American Museum of Natural History Library, Central Park West at 79th St., New York, NY 10024. TEL: (212) 769-5545. FAX: (212) 769-5009. E-MAIL: scipubs@amnh.org

♾ **This paper meets the requirements of ANSI/NISO Z39.48-1992 (Permanence of Paper).**

Downloaded From: https://complete.bioone.org/journals/American-Museum-Novitates on 12 Sep 2024
Terms of Use: https://complete.bioone.org/terms-of-use