FEBS Letters 429 (1998) 4–8 FEBS 20315

Secondary structure of some elements of 18S rRNA suggests that strongylid and a part of rhabditid nematodes are monophyletic

Vladimir V. Aleshin[a], Olga S. Kedrova[b], Irina A. Milyutina[b], Natalia S. Vladychenskaya[b], Nikolai B. Petrov[b,*]

[a] *Russian Institute of Physiology, Biochemistry and Feeding of Livestock, RAAS, Borovsk 240010, Kaluga Region, Russia*
[b] *Section of Evolutionary Biochemistry, A.N. Belozersky Institute of Physicochemical Biology, Moscow State University, Moscow, 119899, Russia*

Received 28 January 1998; revised version received 27 February 1998

Abstract Analysis of the secondary structure of 18S rRNA molecules in nematodes revealed some new traits in the secondary structure peculiar to their hairpin 17. Some of them are characteristic of all the nematodes, whereas others are characteristic exclusively of the order Rhabditida. The loss of a nucleotide pair in the highly conservative region of hairpin 17 distinguishes 18S rRNA of the Strongylida and some species of the Rhabditida from other nematodes and, moreover, from all other organisms. Hence, it is possible to regard the Strongylida and a part of the Rhabditida including *Caenorhabditis elegans* as a new monophyletic taxon.

Key words: 18S rRNA; Secondary structure; Hairpin 17; Molecular phylogeny; Nematode; Strongylid; Rhabditid

1. Introduction

18S ribosomal RNA genes are extensively used in investigations on molecular phylogeny and evolution [1–4]. A number of methods were developed for the phylogenetic reconstruction based on differences in 18S gene nucleotide sequences. However, sometimes these methods do not allow the reconstruction of the phylogenetic relationships among invertebrate groups [5–7]. In this connection, the analysis of some specific characters of 18S rRNA secondary structure may be considered a reasonable addition to the analysis of the 18S gene primary structures.

Although the secondary structure of 18S rRNA is rather uniform in all groups of eukaryotes, some elements of this structure sometimes may have unique features. Some of these traits are readily retained in the course of evolution of species. An example of such unique features is the secondary structure of hairpin 17 (according to Van de Peer et al.'s model [8]) observed in 18S rRNA of the nematode *Caenorhabditis elegans* and related nematodes. It may be hypothesized that comparative analysis of such unique features may be used to trace the possible connections between molecular structure and its manifestation in the cell and to elaborate a new approach in molecular phylogenetic studies.

In this work, the analysis of the primary and secondary structure of hairpin 17 of 18S rRNA was used to elucidate the phylogenetic relationships between the nematodes of the orders Strongylida and Rhabditida. In the contemporary taxonomy these two orders are considered two closely related taxa of the subclass Secernentea composed of free-living and parasitic forms. The outgroups for cladistic analysis were recruited from the non-secernentean nematodes for which differences in 18S rRNA genes were previously demonstrated [9].

2. Materials and methods

2.1. Biological material and DNA extraction

The animals investigated in the present study are *Pontonema vulgare* (Enoplida), *Paracanthonchus caecus* (Chromadorida), *Chromadoropsis vivipara* (Desmodorida), and *Daptonema procerus* (Monhysterida). They were collected by Prof. V.V. Malakhov in Kandalaksha Bay of the White Sea. Specimens of nematodes were fixed in 70% ethanol and DNA was extracted from several intact animals as described by Arrighi et al. [10], with some modifications.

2.2. Amplification and sequencing of the 18S rRNA genes

18S ribosomal RNA coding regions were amplified in polymerase chain reactions using two primers complementary to the 5′ and 3′ termini of eukaryotic 16S-like rRNA genes [11]. Full-length product of amplification was purified by agarose gel electrophoresis, cloned in pBluescript KS^+ plasmid and sequenced on both strands by the dideoxynucleotide method [12] using the Sequenase version 2.0 USB kit, the set of specific 18S rRNA primers and universal M13 sequencing primer.

2.3. Alignment, tree construction and analysis of secondary structure elements

The complete or nearly complete 18S rRNA gene sequences of *Paracanthonchus caecus*, *Daptonema procerus*, *Pontonema vulgare*, and *Chromadoropsis vivipara* were submitted to GenBank under the accession numbers AF047888–AF047891. Other 18S rRNA gene sequences of the Nematoda and some metazoan phyla as well as plants and fungi species were taken from GenBank: Nematoda, *Enoplus brevis*, U88336; *Trichinella spiralis*, U60231; *Caenorhabditis elegans*, X03680; *Strongyloides stercoralis*, M84229; *Haemonchus contortus*, L04153; *Nematodirus battus*, U01290; *Cruznema tripartitum*, U73449; *Pellioditis typica*, U13933; *Pelodera strongyloides*, U13932; *Teratorhabditis palmarum*, U13937; *Mesorhabditis* sp., U73452; *Rhabditis myriophila*, U13936; *Rhabditiella axei*, U13934; *Aduncospiculum halicti*, U61759; *Zeldia punctata*, U61760; *Meloidogyne arenaria*, U42342; *Plectus* sp., U61761; *Ascaris* sp. (a nearly complete sequence was compiled from M58348, X06225, X05836, X06713, M74584, M74585); *Brugia malayi* (a nearly complete sequence was compiled from H30951, H31020, H35866, H39237, H39239, H39242, H39245, H48210, H48950, H48988, H52893, H77269, H91508, R86409, R95205); Nematomorpha, *Gordius aquaticus*, X87985; Rotatoria, *Brachionus plicatilis*, U29235; Gastrotricha, *Lepidodermella squamata*, U29198; Plathelminthes, *Stenostomum*, U95947; Cnidaria, *Anemonia sulcata*, X53498; Mollusca, *Mytilus edulis*, L24489; Chordata, *Homo sapiens*, M10098; Plantae, *Glycine max*, X02623; Fungi, *Saccharomyces cerevisiae*, J01353.

The sequences obtained were manually aligned and the subsets of this alignment were analyzed by both distance and maximum parsimony (MP) methods using bootstrap resampling [13] and appropriate programs within TREECON [14] and PHYLIP [15] packages as well as fastDNAml program [16]. Elements of the secondary structure of

*Corresponding author. Fax: (7) (095) 939-31-81.
E-mail: petr@bioevol.genebee.msu.su

PII S0014-5793(98)00550-X

some regions of the 18S rRNA molecule were constructed manually based on the model proposed by Van De Peer et al. [8].

3. Results

Fig. 1 depicts the primary structure of the region of the 18S rRNA gene corresponding to positions 1396–1424 of 18S rRNA of *Caenorhabditis elegans* (accession number X03680 in GenBank); it is the result of multiple alignment of this gene sequence of nematodes as well as of representatives of some phyla of metazoan animals, plants, and fungi. This region forms hairpin 17 in the secondary structure of the 18S rRNA molecule [8]. As is well known, evolutionarily conservative and variable regions alternate within the 18S rRNA molecule. The hairpin 17 region cannot be considered either evolutionarily conservative or evolutionarily variable. Though it has a high degree of conservativity, certain substitutions may take place there and become retained. It can be seen in Fig. 1 that some positions in the hairpin 17 region mark certain nematode taxa. Some of them are characteristic of all the nematodes, others mark single clades within this phylum. This is especially true with regard to the secondary structure of hairpin 17 of the 18S rRNA (Fig. 2). Fig. 2A shows the structure of hairpin 17 of the most different animals serving as outgroups relative to the Nematoda. It remains identical in such distant animals as vertebrates, mollusks, and worms. As for nematodes, some new traits appear in the secondary structure of their hairpin 17. Fig. 2B demonstrates the hairpin 17 structure in different nematodes excluding Rhabditida. Analysis of these structures shows that the unpaired G at position 17a of outgroups (Fig. 2A) has become a member of a UG pair (Fig. 2B, 17b) and the imperfect double helix ancestral structure could have been stabilized early in the Nematoda evolution. The distal part of this hairpin has become one nucleotide shorter and retained this state in all nematode lineages. The conservative pair UA of the outgroup has been replaced by the pair GC in most of the Nematoda (Fig. 2B, 17d), and the pair AU has been replaced by the pair GC (Fig. 2B, 17k).

Fig. 2C,D describes the secondary structure of hairpin 17 within the order Rhabditida. Here it is possible to distinguish three evolutionary levels (Fig. 3). Level I corresponds to early branches of this order, Cephalobina (*Zeldia punctata*) and Diplogasterina (*Aduncospiculum halicti*). Their hairpins 17 do not differ in structure from that of non-rhabditidan nematodes. They retain the conservative pairs UA in position 17b, GC in position 17d and GC in position 17k (Fig. 2C). The preservation of the plesiomorphic state of these pairs makes it

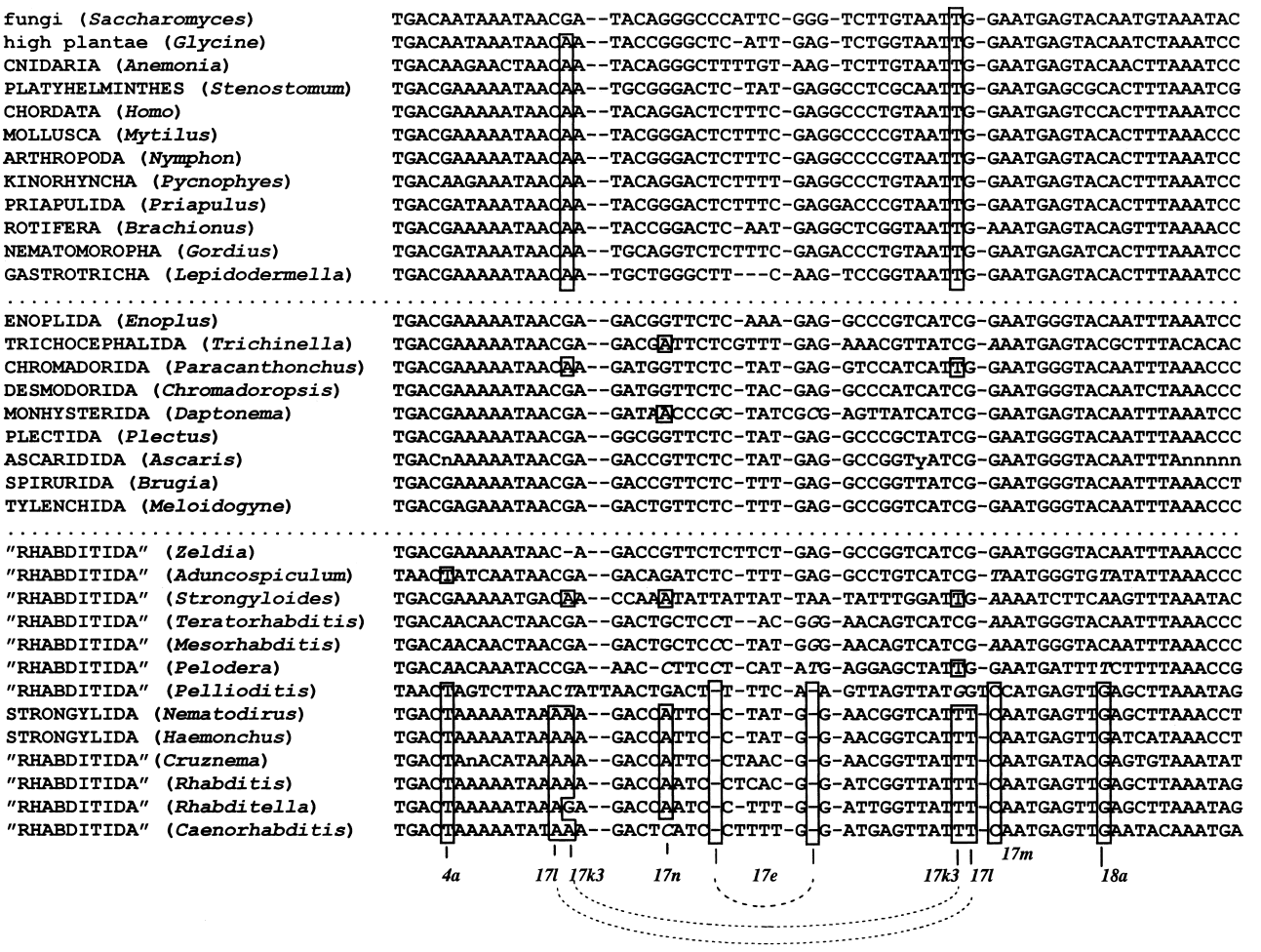

Fig. 1. Alignment of 18S rRNA gene sequence regions comprising hairpin 17 of the secondary structure of the 18S rRNA. The nucleotide positions which can be defined as synapomorphies for certain groups are boxed. Pairs of the complementary substitution are indicated by the same letters, for example 17l and 17k3. For other details, see text.

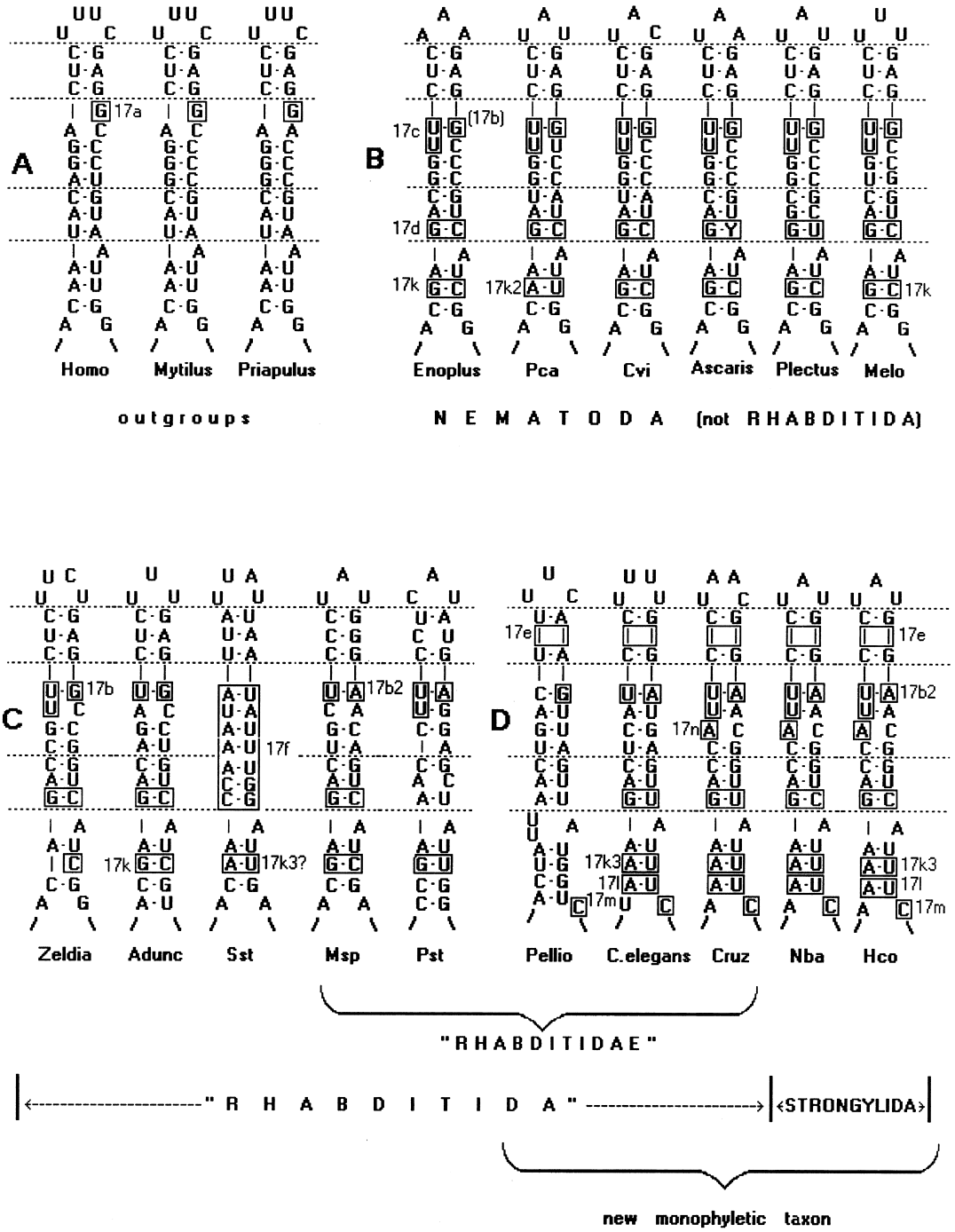

Fig. 2. The secondary structure of the hairpin 17 regions of the 18S rRNA of the nematodes. The pairs of nucleotides peculiar to certain groups are boxed. A: The representatives of the outgroups. B: The representatives of the non-rhabditid nematodes (Pca, *Paracanthonchus caecus*; Cvi, *Chromadoropsis vivipara*; Melo, *Meloidogyne arenaria*), C and D: The representatives of the traditional orders Rhabditida and Strongylida (Adunc, *Aduncospiculum halicti*; Sst, *Strongyloides stercoralis*; Msp, *Mesorhabditis* sp.; Pst, *Pelodera strongyloides*; Pellio, *Pellioditis typica*; Cruz, *Cruznema tripartitum*; Nba, *Nematodirus battus*; Hco, *Haemonchus contortus*). Within the braces, there are representatives of the traditional family Rhabditidae and the new monophyletic taxon including the species of the order Strongylida and part of the species of the traditional family Rhabditidae.

possible to place these branches as ancestral to the Rhabditidae family. In the family Rhabditidae two other levels may be distinguished. Level II corresponds to the genera *Mesorhabditis*, *Teratorhabditis* and *Pelodera*. At this level, the first character of hairpin 17 peculiar to the Rhabditida originated: the replacement of G by A in the conservative pair UG and the formation of the pair UA (17b2). At the same time, the structure of hairpin 17 of *Mesorhabditis*, *Teratorhabditis* and *Pelodera* has retained a number of ancestral features (Fig. 2C). Level III is marked by a number of a new characters. In addition to the above-mentioned character 17b2, they include 17e, 17k3, 17l, 17m, and 17n. Remarkably, these characters are observed not only in certain rhabditid genera such as *Rhabditis*, *Rhabditella*, *Cruznema*, *Pellioditis*, *Caenorhabditis*,

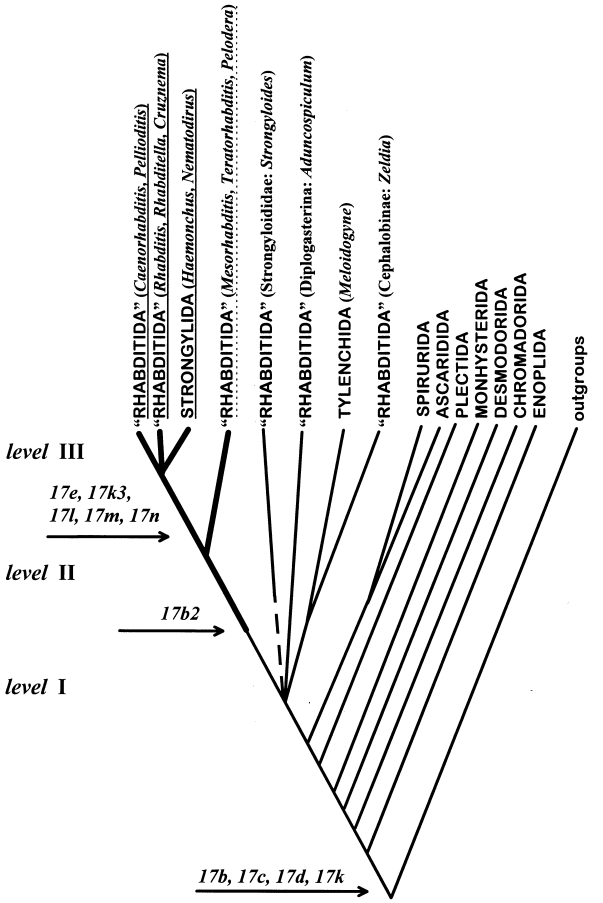

Fig. 3. Phylogenetic tree of nearly complete 18S rRNA gene sequences of nematodes. This topology is a consensus of trees generated by NJ, MP, and ML methods. Arrows indicate the branches where the nucleotide substitutions have occurred, which can be defined as synapomorphies for the groups above these branches. Branches of the Strongylida and part of the Rhabditida cluster are shown in bold.

but in representatives of strongylids (*Haemonchus* and *Nematodirus*) as well (Fig. 2D). The most important is character 17e (a loss of a nucleotide pair in a highly conservative region, Fig. 2D). Based on this character, it is possible to distinguish the 18S rRNA of strongylids and related rhabditids from that of other nematodes and, moreover, from that of all other organisms. Though not absolutely conservative, all other substitutions taken together with the most obvious character 17e also demonstrate the relatedness of rhabditids of the genera *Rhabditis*, *Rhabditella*, *Cruznema*, *Pellioditis*, *Caenorhabditis* and strongylids.

There are other hairpins in 18S rRNA where mutations common to rhabditids and strongylids are observed (helix 4, the hairpin 18 loop, Fig. 1), but the mutations of hairpin 17 are the most demonstrative as they result in non-compensated changes in nucleotide pairs number (17e) in the conservative region of the hairpin. Nucleotide substitutions common to both strongylids and certain rhabditids may be considered synapomorphies acquired by these nematodes. Thus, it is possible to regard them as a monophyletic group; the formation of this group was marked by quite definite evolutionary events in the region of hairpin 17 of the 18S rRNA gene. By contrast, the traditional order 'Rhabditida' should be abandoned as a paraphyletic one, whereas the 'order' Strongylida should be included in the family Rhabditidae.

The monophyly of Strongylida and part of Rhabditidae is confirmed by the topology of the tree constructed using complete 18S rRNA gene sequences (Fig. 3). Comparison of 18S rRNA gene sequences by distance and parsimony methods revealed that Strongylida are not an independent branch of the Secernentea, but are to be included in the Rhabditida as one of a number of their branches [9]. Similar results on the relationships of these taxa were obtained recently by Blaxter et al. [17].

Some other taxonomic conclusions follow from the differences in hairpin 17. For example, by these criteria the genera *Caenorhabditis* and *Pellioiditis* are close to the genera *Rhabditis* and *Rhabditella* and should be removed from the subfamily Peloderina Andrássy to the subfamily Rhabditinae Örley. The genus *Cruznema* should be moved to the same subfamily from the subfamily Mesorhabditinae Andrássy.

4. Discussion

It is well known that in phylograms constructed using traditional algorithms based on the quantitative dissimilarity between nucleotide sequences, branches with an extremely high rate of evolution may be artificially clustered [18,19]. Extremely high evolutionary rate is characteristic of nematodes of the order Rhabditida. It is highly typical of *Strongyloides sercoralis* and *Pellioditis typica* and especially of *Pelodera strongyloides*, a record holder among all the metazoans. That is why there is a certain difficulty in placing these nematodes on a phylogenetic tree using not only morphological but molecular characters of 18S rRNA as well. Even with the same sequence and alignment sets these species are located in different manners in the trees generated by neighbor joining (NJ), MP, and maximum likelihood (ML) methods [9]. The reason for artifacts occurring in tree topology is that repeat and reverse mutations may arise in variable sites resulting in occasional sequence similarity. In such situations, when considering similarity based on certain characters, objective preference should be given to some characters rather than to other ones in reconstruction of the real phylogenetic linkages. The example of such phylogenetically significant characters was described in nematodes at the hairpin 49 region [9]. The characters regarded in this work, although being nucleotide residues representing only a small part of the total 18S rRNA molecule, beyond any doubt have great importance for phylogenetic reconstruction. The revealed synapomorphies are detected in hairpin 17 even in such aberrant sequences as those of *Pelodera strongyloides* and *Pellioditis typica*. Thus, it may be concluded that reverse mutations of these nucleotides occur very rarely due to their important role in secondary structure maintenance and thereby ribosome functioning. In *Strongyloides stercoralis*, many nucleotides appear to be substituted even at the usually conservative sites, destroying substantially the similarity of the 18S rRNA sequence to that of any known nematode species and disturbing the position of this species in the phylogenetic tree. But even so, the 'nematode' pattern differing from that of the majority of animals is maintained through retaining the number of nucleotides comprising hairpin 17. This number in *S. stercoralis* is initial for all the nematodes, and, therefore, the loss of nucleotide pair 17e typical for Rhabditinae and Strongylida species is not

observed in this case. It may be hypothesized that *S. stercoralis* branched before differentiation of the Rhabditina (Fig. 3).

The above results demonstrate that the details of 18S rRNA secondary structure may serve as markers of the evolutionary process in certain lineages. The secondary structure of the 18S rRNA molecule is now known perfectly well; the real existence of double helices in its molecule is confirmed by compensatory changes in evolution maintaining the double helix conformation [20]. Much less is known about the three-dimensional structure of 18S rRNA. The conservativity of the number of nucleotides in some helices or loops in most evolutionary lineages serves as an indicator of their connection with some aspects of ribosome function through influence on the rRNA folding. Among the characters analyzed in this work at least two (17b and 17e) are connected with changes in the number of nucleotides (17b) or nucleotide pairs (17e) in the hairpin. They are probably responsible for positioning of hairpin 17. Since they are not compensated at the rDNA level, it may be supposed that certain changes occur in some other ribosome component which compensate the mutations 17b and 17e at a higher structure level.

Acknowledgements: The authors thanks Drs. A.S. Antonov, V.V. Malakhov, E.D. Krasnova, S.E. Spiridonov, and D.A. Voronov for collection of the nematode species and helpful discussion and Dr. M. Blaxter for providing a copy of the manuscript submitted for publication. This work was supported by RFFI Grant 96-15-97970.

References

[1] Field, K.G., Olsen, G.J., Lane, D.L., Giovannoni, S.J., Ghiselin, M.T., Raff, E.C., Pace, N.R. and Raff, R.A. (1988) Science 239, 748–753.
[2] Wainright, P.O., Hinkle, G., Sogin, M.L. and Stickel, S.K. (1993) Science 260, 340–342.
[3] Aleshin, V.V., Vladychenskaya, N.S., Kedrova, O.S., Milyutina, I.A. and Pertov, N.B. (1995) Mol. Biol. 29, 1408–1426 (in Russian).
[4] Vladychenskaya, N.S., Kedrova, O.S., Milyutina, I.A., Okstein, I.L., Aleshin, V.V. and Petrov, N.B. (1995) Dokl. Akad. Nauk. Russia 344, 119–121 (in Russian).
[5] Winnepenninckx, B., Backeljau, T. and De Wachter, R. (1995) Mol. Biol. Evol. 12, 641–649.
[6] Winnepenninckx, B., Backeljau, T., Mackey, L.Y., Brooks, J.M., De Wachter, R., Kumar, S. and Garey, J.R. (1995) Mol. Biol. Evol. 12, 1132–1137.
[7] Aguinaldo, A.M., Turbeville, J.M., Linford, L.S., Rivera, M.C., Garey, J.R., Raff, R. and Lake, J.A. (1997) Nature 387, 489–493.
[8] Van De Peer, Y., Nicolai, S., De Rijk, P. and De Wachter, R. (1996) Nucleic Acids Res. 24, 86–91.
[9] Aleshin, V.V., Kedrova, O.S., Milyutina, I.A., Vladychenskaya, N.S. and Petrov, N.B. (1998) Russ. J. Nematol. 6 (in press).
[10] Arrighi, F.E., Bergendahl, J. and Mandel, M. (1968) Exp. Cell Res. 50, 47–53.
[11] Medlin, L., Elwood, H.J., Stickel, S. and Sogin, M.L. (1988) Gene 71, 491–499.
[12] Sanger, F., Nicklen, S. and Coulson, R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[13] Felsenstein, J. (1985) Evolution 39, 738–791.
[14] Van De Peer, Y. and De Wachter, R. (1994) Comput. Appl. Biosci. 10, 569–570.
[15] Felsenstein, J. (1993) PHYLIP – Phylogeny Inference Package, version 3.5, University of Washington, Seattle, WA.
[16] Olsen, G.J., Matsuda, H., Hagstrom, R. and Overbeek, R. (1994) Comput. Appl. Biosci. 10, 41–48.
[17] Blaxter, M.L., De Ley, P., Garey, J.R., Liu, L.X., Scheldeman, P., Vierstaete, A., Vanfleteren, J.R., Mackey, L.Y., Dorris, M., Frisse, L.M., Vida, J.T. and Thomas, W.K. (1998) Nature (in press).
[18] Penny, D., Hendy, M.D. and Steel, M.A. (1991) in: Phylogenetic Analysis of DNA Sequences (Miyamoto, M.M. and Cracraft, J., Eds.), pp. 155–183, Oxford University Press, New York.
[19] Swofford, D.L. and Olsen, G.J. (1990) in: Molecular Systematics (Hillis, D.M. and Moritz, C., Eds.), pp. 411–499, Sinauer, Sunderland, MA.
[20] Wheeler, W.C. and Honeycutt, R.L. (1988) Mol. Biol. Evol. 5, 90–96.