Russian Journal of Genetics, Vol. 38, No. 8, 2002, pp. 877–894. Translated from Genetika, Vol. 38, No. 8, 2002, pp. 1043–1062.

Conditionally Neutral Phylogenetic Markers of Major Taxa: A New Aspect of the Evolution of Macromolecules

N. B. Petrov and V. V. Aleshin

Belozersky Institute of Physicochemical Biology, Moscow State University, Moscow, 119899 Russia; fax: (095) 939-31-81; e-mail: Petr@belozersky.msu.ru

Received December 4, 2001

Abstract—The current phase of molecular phylogenetics can be named the 18S rRNA gene era, which is now approaching the end. To date, almost all phyla of metazoans and many taxa of protists are represented in databases of 18S rRNA gene sequences. The elements of the phylogenetic tree of Metazoa inferred from 18S rRNA genes are characterized by unequal validity: some of them seem to be well grounded; others are not adequately supported, and probably will be revised later. The validity of phylogenetic reconstruction is influenced by two main factors: (1) erroneous grouping of long branches that occur because of abnormally high evolution rate; (2) deficit of phylogenetically informative characters. A method for overcoming these difficulties is suggested in addition to known tools: using phylogenetic markers that are stable within individual taxa and evolve by punctuated equilibrium. These markers are least influenced by the convergence caused by a high evolution rate of the entire gene. The nature of these markers of ancient taxa, paradoxical from the perspective of neutral evolution, is discussed, as well as their importance for establishing monophyly of both new large-scale taxonomic groups of invertebrates (Bilateria + Rhombozoa + Orthonectida + Myxozoa + Cnidaria + Placozoa and Echinodermata + Hemichordata) and some major taxa of Nematoda.

BRIEF HISTORY OF MOLECULAR PHYLOGENETICS

Molecular phylogenetics was born in the early 20th century, when immunological experiments showed a very close relationship between humans and apes [1]. Then, the history of molecular phylogenetics was discontinued until the emergence of methods for the determination of amino-acid sequences of proteins in the mid-1950s [2, 3], improved immunological methods [4], and methods for comparing DNA sequences based on DNA–DNA hybridization [5–7]. The accumulation of data that were obtained by these methods have stimulated the development of techniques for sequence alignment [8–10] and phylogenetic reconstruction [11–14], as well as extensive phylogenetic studies of different groups of organisms [15–18].

The development of both methods for the determination of nucleotide sequences and the cloning techniques for individual genes (especially using polymerase chain reaction) have initiated rapid accumulation of data on nucleotide sequences of various genes. Molecular phylogenetics entered the phase that can be named the era of 18S rRNA genes, because, by virtue of a number of reasons [19], exactly these genes promised to help in solving many phylogenetic problems. Certainly, other genes have been also actively used for this purpose [20, 21] but the scope of phylogenetic studies using the 5S [22, 23] and 18S rRNA genes is considerably larger than the scale of similar studies based on other genes. The available sample of 18S rRNA gene sequences currently covers all phyla of metazoans except Loricifera and Lobatocerebrida. The European Small Subunit Ribosomal RNA database (http://rrna.uia.ac.be) contains 6000 complete aligned sequences of the 18S rRNA genes of eukaryotes, of which more than 1500 organisms belong to metazoans [24].

The phase of phylogenetic studies using ribosomal RNA genes seems to approach its end. At present, we are on the verge of a new era, that of comparison of a multitude of various genes, and then complete genomes. In view of this, there is a need to review the past phase, estimate its advances, detect its weak points, outline the ways for overcoming these weak points, and determine the prospects of phylogenetic studies in the new, genomic era.

Even now, comparative studies of rRNA genes have significantly changed the views on the evolution of eukaryotes, permitted to resolve some old disputes, and introduced many new phylogenetic ideas [25–30]. Since the main advances of phylogenetic studies using rRNA genes of protists [31] and plants [32] have recently been reviewed, our review is devoted to analysis of similar data obtained mostly for metazoans.

ADVANCES OF THE PHASE OF TRADITIONAL PHYLOGENETICS

The metazoan phylogeny is characterized by a paradoxical situation: while all main animal phyla have been described as early as in the 19th century, the relationships between them are still debated. This situation can largely be explained by the fact that the famous

1022-7954/02/3808-0877$27.00 © 2002 MAIK “Nauka / Interperiodica”

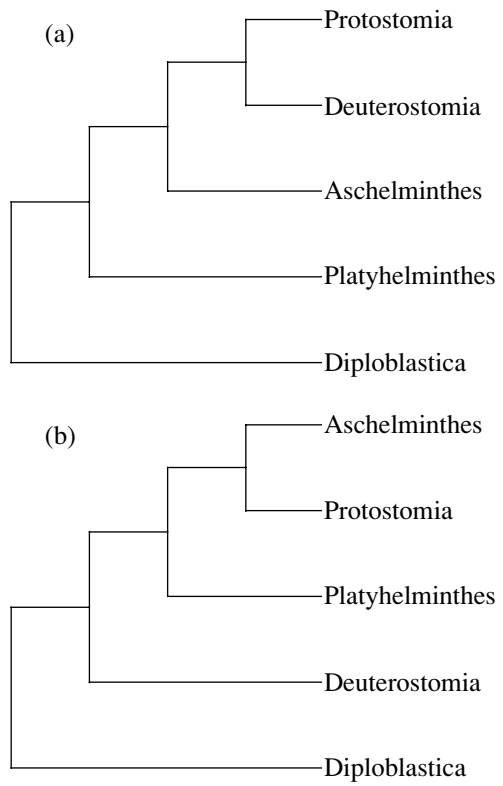


Fig. 1. Two main variants of the phylogeny of metazoans based on the traditional triad of methods: (a) acoelomate–coelomate hypothesis suggesting the transition from acoelomate ancestors through pseudocoelomates (Aschelminthes) to higher coelomate animals; (b) archicoelomate hypothesis suggesting the origin of Bilateria from a putative coelomate ancestor.

triad of basic methods of traditional phylogenetics (comparative anatomic, embryological, and paleontological methods) because of their specific limitations cannot provide comprehensive information for reliable phylogenetic reconstructions. Nevertheless, by the mid-20th century, the prevailing views on animal phylogeny were the concepts based on the idea of a progressive evolution from lower to higher forms with a gradual increase in the complexity of their organization. These views are included in the best known courses of invertebrate zoology [33] and reflected in later publications [34, 35].

According to these views, which were implemented in the so-called acoelomate–coelomate phylogenetic tree (Fig. 1a), the sponges (which have a low level of tissue organization, few types of differentiated cells, and lack nervous cells) were the first taxon detached from the main stem of animals. The next stage was the division of the remaining animals into Diploblastica and Triploblastica. In the embryogeny, the former produce only two blastophylla: the ectoderm, which gives rise to the epidermis and nervous cells, and the entoderm, which forms the alimentary cavity. The latter produce one more blastophyllum during their embryogenetic development, the mesoderm, which is situated between the ectoderm and the entoderm and produces muscles and internal organs. As diploblastic animals (hydroid and scyphoid medusas, hydropolyps, and corals) are usually characterized by radial symmetry, they are named Radiata. Triploblastic animals, by virtue of their bilaterally symmetric structure, are named Bilateria. The position of Ctenophora and Placozoa, as well as Mesozoa (wormlike acoelomate animals with a low number of cells), remained controversial. Typically, they were placed in the basal part of phylogenetic tree.

All Bilateria (triploblastic animals) were divided into acoelomates and coelomates according to the presence or absence of coelom, which was considered phylogenetically important. Acoelomates and coelomates were considered lower and higher animals, respectively. The latter were in turn divided into protostomes, deuterostomes, and lophophorates, whose position seemed to be intermediate between protostomes and deuterostomes. The phylogeny of protostomes was based on the sister relationship between Annelida and Arthropoda, which were considered close relatives within the group of articulate animals (Articulata). This system did not provide for a definite position of a number of animal taxa (Nematoda, Rotifera, Nematomorpha, Priapulida, Kinorhyncha, and Gastrotricha), which have an internal body cavity interpreted as residual blastocoel rather than a true coelom. Therefore, they were lumped together into a supraphylum of pseudocoelomates or aschelminths (Aschelminthes) and were placed between acoelomates and true coelomates.

This phylogenetic hypothesis is an implementation of the idea of a gradual progress in the level of organization from flatworms and roundworms forming the basal part of the Bilateria tree to two main branches of coelomates, arthropods and vertebrates. However, already studies of animal ultrastructure has shown that the structure of coelom is more plastic and diverse than it seemed at the level of light microscopy [36–39], so it cannot be considered a very conservative trait.

Another variant of the traditional tree (Fig. 1b) implies an early division of Bilateria into Protostomia and Deuterostomia. In this case, flatworms and roundworms are presented as secondarily simplified protostomes that have lost the coelom. Finally, one more variant implies an independent origin of the main trunks of Bilateria from radiate ancestors [40, 41].

In spite of a remarkable distinction between these hypothetical scenarios of animal evolution, it is difficult to make a choice between them, because they are based on the same data set, interpret it differently, and do not suggest any corollaries that could be tested using morphological and embryological methods.

CONSEQUENCES OF THE ERA OF 18S rRNA IN MOLECULAR PHYLOGENETICS

A comparison of the 18S rRNA genes provided phylogenetic studies with numerous new (molecular) characters that are completely independent from anatomic and embryological characters. Many of old hypotheses appeared to be incompatible with these new data. Thus, formerly popular hypotheses on the primary primitiveness of *Trichoplax* (Placozoa) [42] and Mesozoa [43–45] became a part of scientific history. The hypotheses implying that Ctenophora represent the highest stage in the evolution of Coelenterata and are the putative ancestors of Bilateria or Deuterostomia have little chance to survive, whereas the monophyly of Bilateria is currently beyond question [42]. These data became part of the 18S rRNA phylogenetics virtually at the same time with the discovery of wide homology of homeotic genes controlling the embryonic development. As a result, comparative anatomy faced the task of searching for remote homologies and revising the main structural patterns of Bilateria. It is now thought that cnidarians are the most probable sister group of Bilateria [46, 47], and the new models of evolution [48] do not repeat the old "ceriantarian" (archicoelomate) hypothesis. Even the idea on the taxonomic scope of Metazoa has drastically changed after it was found out that myxosporidia, which were formerly described in the courses of protistology, actually belong to triploblastic [49–51] or diploblastic [52, 53] animals.

The phylogeny of Bilateria is seen in a new light (Fig. 2). Instead of the former group Articulata (annelids and arthropods), the central position among Protostomia is occupied by Annelida and Mollusca, i.e., the trochozoan phyla having the ontogenetic stage of ciliary larva (trochophore). Many transitional taxa have received valid positions as relatives of trochozoans. This concerns the following taxa: Lophophorata, which seemed to be a connecting link between Protostomia and Deuterostomia [54, 55]; the phylum Nemertini, predatory sea worms having a circulatory system but lacking a coelom (which seemed to be the highest-level acoelomates, but at present are considered simplified descendants of coelomate ancestors [56]); and Pogonophora [57], the pride of Russian zoological science. The latter taxon has been interpreted by A.V. Ivanov as a new phylum of Deuterostomia [58] but was now returned to the class Polychaeta as a family of Siboglinidae or Lamellisabellidae Ushakov, 1933, in accordance with their first description [59, 60].

The greatest impact, due to exceptional uniqueness, was produced by the results of a comparison of 18S rRNA genes that revealed a close relationship between arthropods as higher (coelomate) animals and roundworms as lower animals and introduced a new group, molting animals (Ecdysozoa) [61]. Along with Arthropoda and Nematoda, Ecdysozoa include Onychophora, Tardigrada, and some phyla of pseudocoelomate worms. If the concept of Ecdysozoa is true, *Drosophila melanogaster* and *Caenorhabditis elegans* (the subjects of advanced genomic projects) represent closely related branches of the large phylogenetic tree of metazoans. In addition, the concept of Ecdysozoa is exemplary in respect of the fact that anatomic and embryological arguments against making annelids and arthropods closely related [62, 63] were not generally recognized until this concept was inferred from molecular characters (18S rRNA genes). After that, the concept of Ecdysozoa was repeatedly supported by cladistic analysis of morphological characters [28–30].

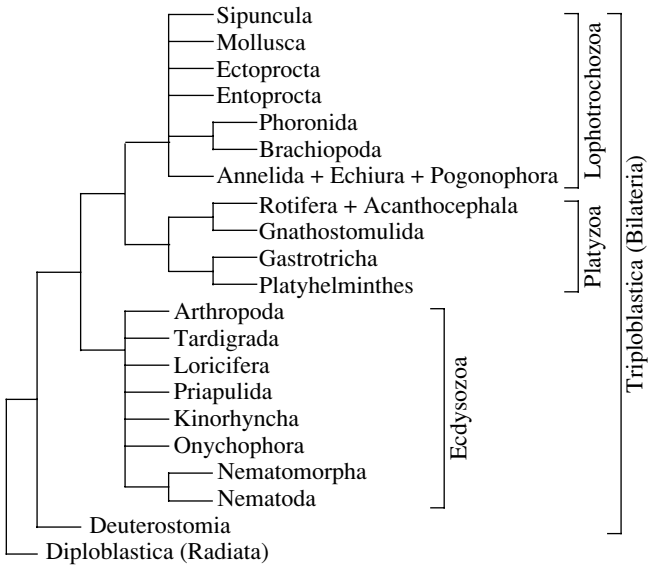


Fig. 2. A variant of the phylogeny of metazoans inferred from 18S rRNA genes and data obtained by cladistic analysis of morphological characters.

However, several key taxa of invertebrates, including Platyhelminthes, Acoela and Nemertodermatida (acoelous turbellaria), Chaetognatha, Bryozoa, Orthonectida, Rhombozoa, and Myxozoa, do not occupy a constant position on the phylogenetic tree inferred from 18S rRNA genes. For example, depending on the set of studied species and the algorithm of tree construction, flatworms either cluster together with coelomate trochozoans and lophophorates [61], or form a separate branch of Protostomia [28] (Fig. 2), or split into independent lines, of which some branch off the crown of the phylogenetic tree of Bilateria and others (Catenulida or Acoela) seem to be the first basal branch of Bilateria [64, 65]. A choice between these phylogenetic hypotheses entails radical changes in the concepts on the organization of the nearest common ancestor of bilaterians and on the pathways of their early evolution.

Ambiguous results on the position of the aforementioned taxa and an insufficient clarity of the relationships between animal phyla (a comb-shaped structure of the established branches in the phylogenetic tree shown in Fig. 2) reflect both obvious and hidden problems of phylogenetic analysis of molecular data.

PROBLEMS OF MOLECULAR PHYLOGENETICS

In spite of relative clarity of the methods of phylogenetic reconstruction based on molecular data, they are not devoid of serious and sometimes hidden problems.

One of these problems is related to the presence of repetitive sequences and to the necessity of discriminating between orthologous and paralogous homologs [66]. Because of concerted evolution, ribosomal genes are little different from one another [67–70], and in their case this problem is significant only in exceptional cases [71, 72]. Therefore, we shall not discuss it.

The second problem, the ambiguity of alignment, arises in consequence of a considerable sequence divergence, which is followed, along with multiple point substitutions, by deletions and (or) insertions of single nucleotides or relatively extended regions. In some cases, it is difficult to determine the boundary points of insertions and to match each nucleotide in one sequence with a homologous nucleotide in the other.

The third problem is a deficit of phylogenetically informative characters. It is manifested as low resolution of branching in many phyletic lineages and a low statistical support of groups, which raises doubts about the validity and reliability of the results obtained.

The fourth problem was termed long-branch attraction (LBA). Its effect is expressed as a distortion of tree topology by moving the groups that have 18S rRNA gene sequences characterized by an abnormally high evolution rate (secernentean nematodes [61], acoelous turbellaria [64], mesozoans, and myxozoans) into a basal position.

In what follows, we consider the latter two problems in more detail.

DEFICIT OF PHYLOGENETICALLY INFORMATIVE CHARACTERS

The most important reason for the difficulties concerning the use of molecular methods of phylogenetic reconstruction is related to the mere fact that, whatever gene is selected, some clades do not have molecular synapomorphies within it. A totality of 1800 nucleotide positions of a medium-size 18S rRNA gene cannot contain a million of outapomorphies for a million of taxa existing on the Earth. In some clades, synapomorphies could not occur because of a short time of the existence of the stem group before radiation; in other clades, prior to radiation, synapomorphies had only occurred in variable (rapidly changing) regions, and they have not been conserved for the time of further evolution, having undergone repeated substitutions. Apparently, in such cases, calculation tricks are useless to establish the monophyly of these groups, and the only hope of success is related to the opportunity to find molecular synapomorphies for these clades in other genes. Probably, using any individual gene, it is only possible to resolve branching in some parts of phylogenetic tree; the position of the remaining branches is either largely random or depending on the LBA effect (see below). This is one of the reasons for different results of phylogenetic reconstruction inferred from different genes. Another possible reason is related to evolution rate differences between genes. The effect of this factor on the results of phylogenetic reconstruction will be discussed below.

A discrimination between true and accidental groups requires information on phylogenetic markers fixed in different genes. And it is not that several genes represent a genome more adequately, but the possible absence of synapomorphies for many specific groups within a single gene. Two main approaches are possible for phylogenetic reconstruction inferred from many genes. The first is the simple combination of the sequences of different genes into a single set for alignment, which is then analyzed using conventional procedures. The second approach is the construction of a valid phylogenetic tree from tree fragments inferred from individual genes and containing both true and false elements of topology. This approach can be formalized, and it may be more fruitful than conventional procedures of the joint alignment of all available information, which is never complete.

Genomic projects show promise of a great advance in the accumulation of phylogenetic data, especially the information needed to establish the initial radiation in major taxa. However, even in this case, an ensemble of 10^8 nucleotides (in the smallest eukaryotic genomes) will be confronted with the ensemble of 10^6 taxa. Since even in the smallest genomes, an overwhelming majority of nucleotides is related to evolutionarily variable regions (which are useless in a comparison of remote taxa). Only a small part of each genome can be used for large-scale reconstruction. Thus, a switch to a comparison of many genes and even complete genomes will not definitively solve the problem of scarcity of phylogenetically informative molecular markers. According to the above estimate, a comparison of complete genomes can hardly guarantee success in finding obvious outapomorphies for each clade. Hence, even in this case, some taxa will be characterized based on homoplasies or unique combinations of nonunique characters.

THE LBA PROBLEM

Already the evidence on molecular DNA hybridization has shown that, in different phyletic lineages, differences in the rates of sequence evolution can reach several times [73]. Later, the accumulation of data on various genes (including the 18S rRNA genes) revealed many instances of abnormally high evolution rate in some phyletic lineages. The reasons for the accelerated rate of molecular clock in some phyletic lineages are nearly always unknown. One can only speculate which selective factors required a quick reorganization of the macromolecule or, conversely, made neutral and, con-

sequently, allowable substitutions in the domains that formerly were strictly controlled by stabilizing selection. In a phylogenetic reconstruction, the differences in the evolution rate pose a serious obstacle, which is very difficult to overcome [74].

In a simplistic form, the nature of this problem is as follows. Similar genes that have been little evolutionarily changed cluster together in a phylogenetic reconstruction, whereas a considerably changed gene will differ from homologous genes of other taxa and will occupy an isolated position within its family in the basal part of phylogenetic tree. In a phylogenetic reconstruction, two or more genes that are dissimilar to all other genes can group together as the remainder after the clustering of other sequences that are more similar to one another. The conditions for LBA and its effect on unrooted and rooted trees are schematically shown in Fig. 3.

Another reason for LBA is related to a limited number of DNA monomer types. Some newly fixed substitutions can coincide in individual positions of two phylogenetically distant but rapidly evolving genes by convergence, and this "noise" can contribute to erroneous grouping between long branches in the case of significantly reduced similarity with the nearest relatives. If a random, computer-generated sequence is included in a phylogenetic analysis, the most pronounced long branches can cluster together with this sequence in some bootstrap replications [75].

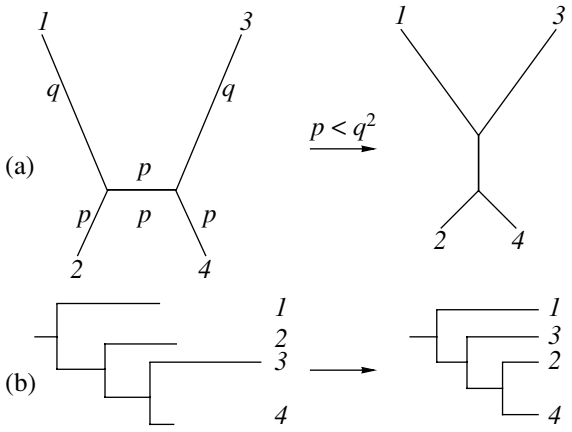


Fig. 3. LBA effect: (a) Scheme illustrating the appearance of wrong grouping in an unrooted tree [101]. If the probability of mutation is p and is lower than q^2, a stable group is formed of nonrelated long branches (*1*, *3*) by the maximum-parsimony method; (b) In the rooted tree, outgroup (*1*) attracts the long branch (*3*), masking its length and hampering its identification.

Identification of Long Branches and the Registration of the Effect of Their Attraction

Solving the LBA problem includes two aspects: first, it is necessary to identify long branches; second, to elucidate their nature, i.e., to determine if they reflect a long-term evolution (in this case, their basal position on the tree is a true reflection of phylogeny) or they represent the product of abnormally rapid evolution in particular phyletic lineages (in this case, their basal position on the tree of sequences represents an artifact caused by LBA).

The identification of long branches is not as simple as it may seem at first sight. The method for a comparison of evolution rates in different phyletic lineages, the so-called relative test, was suggested as early as at the dawn of molecular phylogenetics [76]. It consists in a pairwise determination of genetic distances between two studied taxa with respect to a certain outgroup. If several outgroups are used, it is possible to test the statistical significance of the between-distance differences, if any. However, if this procedure is used, the distance from the outgroup to a rapidly evolving taxon may be underestimated because of saturation of sequences with mutations [74]. The more distant the outgroup from the studied species, the more pronounced the underestimation of distance. A clear example is a comparison of mitochondrial HSP70 proteins. If α-proteobacteria are used as outgroup, the HSP70 protein of microsporidia seems to evolve at a very high rate (long branch); whereas, if an eukaryotic cytosol HSP70 protein serves as outgroup, the evolution rate of this gene in microsporidia appears to be normal, because not only the HSP70 proteins of microsporidia, but also the HSP70 proteins of all studied organisms, are positioned very distantly from the an outgroup [77, 78]. Thus, a wrong choice of outgroup can hide long branches, and they are best detected in an unrooted tree.

The molecular markers that most commonly used for reconstruction of phylogeny of phyla and classes are usually saturated with mutations, so the estimation of the variation in the rate of evolution requires other approaches. One of these is a comparison of evolutionary distances within the studied group, because the underestimation is less pronounced in the case of low distances [79].

Several indications were proposed to find out whether the early branches of sequence trees have a basal position because of high evolution rate rather than long-term evolution [80, 81]: (1) the branches leading to mutually attracting taxa are very long; (2) these branches can be grouped with random sequences; (3) some methods of reconstruction that are less sensitive to the LBA effect must discriminate mutually attracting branches; (4) the studied sequences have many unique substitutions; (5) the studied sequences significantly deviate from the consensus sequence for the studied sample; (6) compared to the typical length of 18S rRNA, the studied sequences are abnormally long or short; (7) a high relative apparent synapomorphy (RAS) taxon variance [82].

A critical analysis of these criteria revealed that they are far from being of equal value and can be questioned

in certain cases [83]. For example, the method of relative apparent synapomorphy analysis (RASA) was specially developed to identify long branches [82]; however, it cannot discriminate between the long branches deriving from abnormally rapid evolution and the long branches that are the product of long-term evolution at a normal rate [74].

Thus, the LBA problem is difficult to solve. This fact is reflected in a long-term discussion between the advocates of the origin of Acoela from the evolutionary root of Bilateria [30, 65] and their opponents, who consider their basal position on the phylogenetic tree of 18S rRNA as an artifact of the high evolution rate of the 18S rRNA gene [28, 64, 84]. Another unsettled question that is directly related to the LBA problem is the affiliation of Myxozoa with Bilateria [49–51] or Cnidaria [52, 53, 83]. Grouping Myxozoa with cnidarians is observed only if a parasitic hydroid *Polypodium hydriforme* (which is the only coelenterate species comparable to Myxozoans by the abnormally high evolution rate for the 18S rRNA gene) is included in the studied species sample. Although this results in a very high probability of grouping Myxozoa and *P. hydriforme* together because of the LBA effect, this problem cannot be solved by the mere elimination of *P. hydriforme* from the analysis, because Myxozoa and *P. hydriforme* are characterized by unique cytological features that could be explained by their relationship [85]. Another example of discussion underlain by differences in the evolution rate of 18S rRNA genes is the putative sister relationship between two insect orders Diptera and Strepsiptera [83, 86], which is related to the LBA effect by the opponents of this hypothesis [81, 87]. Solving these problems by methods of molecular phylogenetics requires data on other genes or the development of more advanced tools for the analysis of the available 18S rRNA sequences.

Reasons for the Occurrence of Long Branches

The reason for the occurrence of long branches is an abnormally high evolution rate of all or some particular genes representing some phyletic lineages. The evolution rate of single genes or complete genomes is governed by two main factors: (1) mutation rate and (2) fixation probability of mutations [88]. The mutation rate of complete genomes is influenced by (1) generation time [89]; (2) efficiency of DNA repair systems [73]; (3) metabolism rate [90], which determines the concentration of free radicals and other mutagenic products. The mutation rate of single genes may be affected by the gene background [91] and some specific features of their structure, such as the presence of repetitive motifs, which are the targets of general genetic recombination and DNA polymerase slippage [92]. The fixation probability of mutations depends on initial mutation rate, selective value, effective population size, and generation time.

The primary structure of rRNA is under strong pressure of functional limitations, because it apparently preserves its functions invariable since the moment of its origin in primitive living organisms to the present time. Two nonalternative possibilities of fixation of an excessive number of mutations in a long branch can be proposed: (1) an increase only in the fixation rate of presumably neutral mutations within variable regions of molecule; (2) a qualitative change (diversification) in the types of permissible mutations. In the latter case, in certain species the changes are fixed in the regions that are prohibited from mutation in other species and are therefore considered functionally important.

The first possibility must be determined by general factors influencing the rate of neutral mutations, such as effective population size [93]. Apparently, this factor increases the rate of neutral molecular evolution in all genes rather than in a single gene. For example, an increased evolution rate of *Caenorhabditis elegans* genes not only represents a well-known source of LBA artifacts in phylogenetic trees inferred from 18S rRNA [61, 74], but also poses a serious problem in the case of using data on many genes [94]. In a study of early divergence in Metazoa [46], the same four species out of 23 studied appeared to be included in long branches in the phylogenetic trees inferred from 18S and 28S rRNA, though one cannot completely exclude some coordinated changes of large and small subunit ribosomal RNA in this case.

General factors are responsible for the increased evolution rate of genes encoded by mitochondrial genomes (Fig. 4), which is expressed in many phyletic lineages and is typical of all mitochondrial genes (the substitution rate in the genes of evolutionarily conserved proteins can be judged from the rate of synonymous substitutions in closely related species). A special three-times rule was formulated for the increased evolution rate of mitochondrial genes by contrast with nuclear genes [95]. The reason for the increase can hardly be a lower effective number of organelles than their host cells [95], because the egg cell usually has many mitochondria.

Another opportunity of a rapid accumulation of substitutions, which is related to an expansion of the range of permissible mutations into the areas of conserved regions, occurs in a gene-to-pseudogene transition or in the case of loss of functionality in a domain of multifunctional protein. For example, the presumed loss of the function of actin binding by the elongation factor 1α (EF-1α) genes in Ciliophora resulted in the fixation of mutations and the destruction of some conserved motifs [96]. These changes did not apparently affect the translation mechanism. At least the remaining genes that participate in translation, such as the 18S rRNA gene, evolve at a normal rate in most of Ciliophora. On the other hand, the increased variability of domains that control actin binding caused an increased variation of other EF-1α regions in Ciliophora [96].

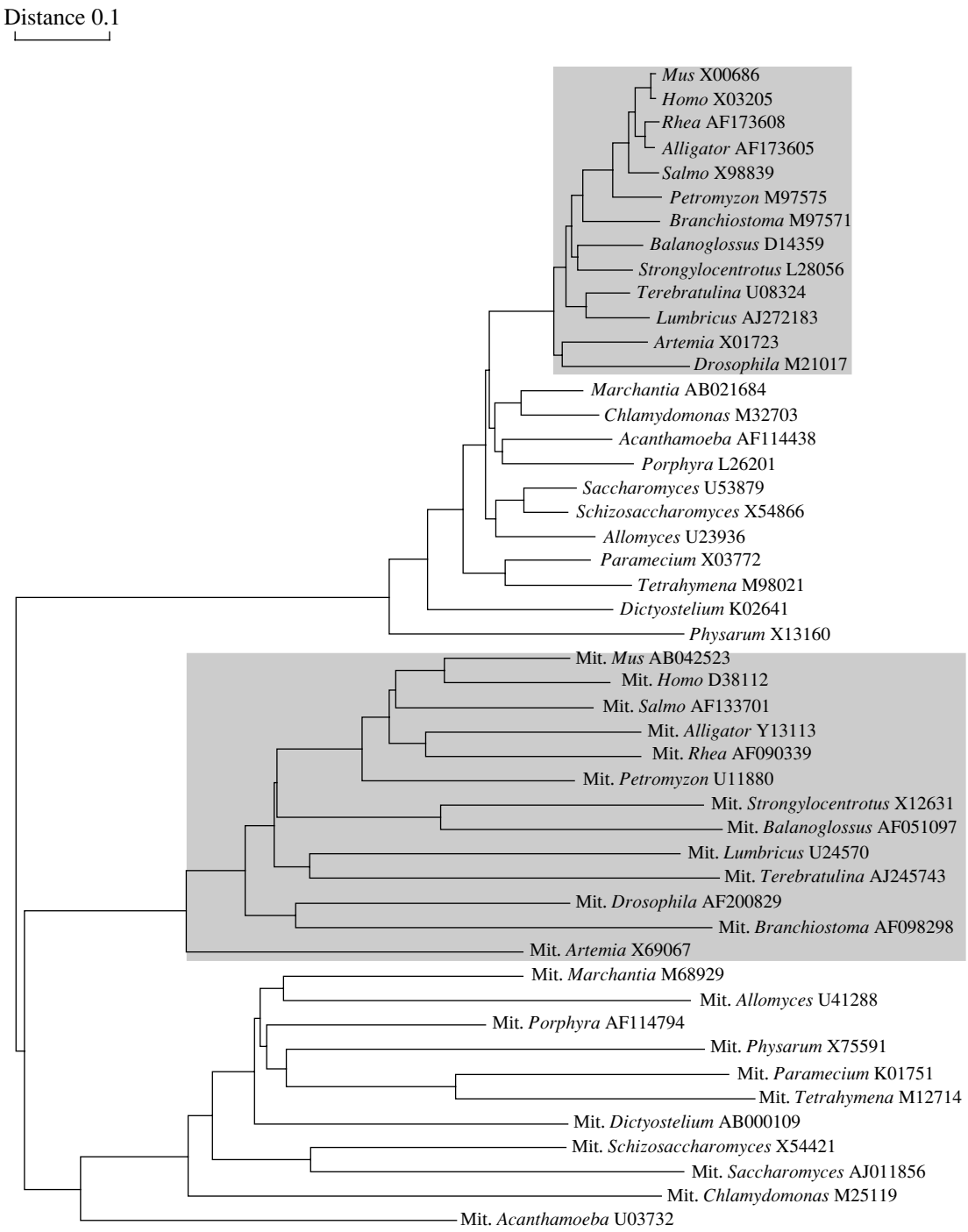


Fig. 4. Phylogenetic tree constructed based on genetic distances inferred from eukaryotic nuclear and mitochondrial small rRNA genes studied in the same sample of taxa. Taxa of metazoans are shaded. For mitochondrial genes, the length of branches is approximately three times longer than for nuclear 18S rRNA, which indicates a higher evolution rate of mitochondrial genes. GenBank accession numbers are given after species names.

It is evident that similar factors contribute to the cases of a reduced conservation of ribosomal genes. The ribosome represents a giant multidomain ribozyme. In spite of the advances of structural analysis [97], the functional importance of many domains remain unknown. Such domains include the V4 region that is distinguished for an exceptionally high level of variation. It is exemplified by extended insertions that are typical of the V4 region of Cicindellidae, Cladocera, and many other taxa; even closely related species significantly differ from one another in the length and nucleotide composition of these insertions [92, 98].

In these cases, the increased evolution rate of the V4 region is obviously related to the specific primary structure of this region, which is saturated with short repeats, and is caused by DNA polymerase slippage. However, if the V4 region is excluded from the alignment, the branches that lead to Cicindellidae on the phylogenetic tree of 18S rRNA remain still longer than in the case of species belonging to other families [92]. In other words, the conserved 18S rRNA gene regions of Cicindellidae are also allowed to have an increased variability that does not directly depends on DNA polymerase slippage and is probably correlated with substitutions in the V4 region.

The initial changes in the V4 region of the 18S rRNA gene of Cicindellidae and Cladocera, which cause a local increase in the mutability owing to DNA polymerase slippage, sometimes are developed to a hypertrophied stage. Simple (high-entropy) sequences deriving from DNA polymerase slippage can hardly be functionally important and controlled by directional selection. Apparently, the evolution of the V4 region is caused by nothing more than molecular mechanisms underlying a slippage-induced increase in the rate of certain mutations [92]. An analogy can be suggested with hypertrophied organs in paleontological series, which were explained by vital force or by orthoselection. The only difference is that hypertrophy characterizing a phylogenetic series of macromolecules is open to experimental study.

Approaches to Solving the LBA Problem

Grouping sequences by the "negative" trait of differences from all other sequences in the studied sample poses serious methodological problems in the area of molecular phylogenetics, because the modern paradigm of phylogenetics is based on identification of clades by synapomorphies. Some authors try to solve this troublesome problem by rejecting any distance-based methods of tree construction in favor of cladistic approaches, for example, the maximum-parsimony method [99]. However, this method can also produce the LBA-related clustering artifacts [100, 101]. Therefore, various artificial tricks were suggested to reduce the contribution of the LBA effect: (1) markedly differing sequences should not be included in the analysis [61]; (2) sequences with abnormal evolution rate should be "diluted" by sequences with normal evolution rate [102]; (3) hypervariable gene regions should not be taken into consideration (either excluded from alignment or assigned a lower weight in the analysis); and (4) the tree construction algorithms and sequence evolution models should be less dependent on evolution rate, which more accurately match the real process [103].

The first strategy was implemented in the studies devoted to the establishment of the Ecdysozoa group [27, 61], in the study that substantiated independence of Acoela [65], and in other works. Another variant of this strategy is the choice of a gene within the studied group, which evolves with minimum deviations from the molecular clock hypothesis [76]. For example, the phylogeny of some protist taxa, which are inferred from mitochondrial proteins HSP70 [77, 78] and cpn60 [104], are apparently not affected by LBA, if α-proteobacteria are used as outgroup.

It is evident that the trick of selective use of data is very vulnerable to criticism, because the mere indication of the fact that some sequences are related to long branches (sources of artifacts) is definitely not sufficient to substantiate the exclusion of some scientific data (nucleotide sequences) from analysis. Moreover, this trick must not be used, if the task is exactly the determination of position on the tree for the studied sequence, to which a long branch leads. The aforementioned example of *P. hydriforme* 18S rRNA gene sequence is very appropriate. In this case, another trick may be helpful, which is also related to the choice of taxa. Its strategy consists in a "dilution" of long branches by adding new taxa to the studied sample. This strategy was used in a study of the taxonomic position of Arthropoda [27] and supported the hypothesis of monophyly of molting animals, Ecdysozoa [61]. This trick is efficient even if the added sequences also evolve at a high rate [105]. For example, according to the first fragmentary data on the ribosomal genes of amoebas, no common characters were found in their sequences, and all amoebas seemed to be a polyphyletic set of nonrelated protists originated from different unknown ancestors [106]. However, adding increasingly more sequences outlined a trend of classifying all amoeboid protists into two or three monophyletic groups [107].

The third trick, which consists in excluding hypervariable nucleotide positions saturated with substitutions from alignment, makes possible to reduce the noise. In so doing, artificial data weighing also takes place, as well as the underestimation of distances between species. The necessity of considering the heterogeneity of evolution rate among individual positions along the strand of DNA molecule was pointed out relatively long ago [108]. However, the first study, in which the heterogeneity of positions by evolution rate was actually taken into consideration, appeared almost a decade later [109]. Depending on whether the heterogeneity by positions is considered or not, different protist taxa appear as the earliest branch on the trees inferred from eukaryotic 18S rRNA [110]. Thus, microsporidia were moved from a basal position, which they were given in an earlier reconstruction [111, 112], to the crown of eukaryotic tree and were placed in the immediate vicinity of fungi, if using the correction for heterogeneity of individual rRNA sites by evolution rate [107]. This phylogenetic result is in complete agreement with the data on other genes [113–116].

It is well known that the rate of substitutions is unequal for different positions along a functional molecule sequence. For example, five classes of positions were distinguished in the 18S rRNA gene [117], in

which variability estimated in relation to the mean variability for the entire molecule is 0.0918, 0.324, 0.977, 2.38, and 5.74, and approximately one-third of this gene positions are invariant. These variability levels indicate that rRNA regions differ in the degree of conservation, and the rate of differences reaches many times. It is thought that the variability of positions by evolution rate along the entire rRNA sequence is best described by a γ distribution [118]. Different variants of this model for the estimation of variability of gene regions are realized in a number of computer procedures, e.g., TREECON [119], TREE-PUZZLE 5.0 [120], etc. Based on the results of calculations, each position can be assigned a weight (category) which is in inverse proportion with the evolution rate. This weight is used at the next stage of tree construction. Conventional phylogenetic software packages (fastDNAml [121], PHYLIP [122], and PAUP [123]) provide for the possibility of reading initial data (alignment) together with the categories for each position. A good choice of categories improves the resultant tree, but the procedure of trait weighing, albeit formalized, allows a great degree of subjectivity, for example, in the choice of the number of categories. And, most importantly, the categories imply a constant probability of fixation for substitutions in a certain site, whereas its change in time is actually characterized by discrete steps in accordance with the model of punctuated equilibrium [124].

Finally, the fourth trick consists in using both methods of phylogenetic reconstruction that are less sensitive to the LBA effect and more realistic models of evolution of nucleotide sequences in the estimation of genetic distances between species. For example, a maximum likelihood (ML) approach was specially developed to avoid this effect [125]; however, this expectation comes true in the only case, if the applied model of evolution perfectly matches the input data, which is not fulfilled in most cases, because the actual parameters of the fixation of mutations in the course of evolution are mostly unknown [118].

ANALYSIS OF SECONDARY-STRUCTURE ELEMENTS OF 18S rRNA: ANOTHER METHOD FOR OVERCOMING THE DIFFICULTIES

A less obvious approach for overcoming the aforementioned difficulties is a cladistic analysis of specific evolutionary changes of both primary and the predicted secondary structure of macromolecule. The presence of specific changes (synapomorphies or outapomorphies in terms of cladistics) in the analyzed sequence indicates that it belongs to a certain clade, even if multiple substitutions in variable regions do not allow its definite position on the tree because of LBA-caused artifacts. These macromolecular characters can be used as input data for computer procedures of molecular morphometry [126]. Although formalization of their choice is difficult, using these particularly indicative characters in a phylogenetic study, in essence, represents an extreme case of conventional weighing of macromolecule regions.

The procedure of verification of tree topology by single markers may be criticized, but it is justified for two reasons. First, higher-rank taxa can preserve a very small number of synapomorphies. For example, in the order Enoplida (marine nematodes), only two specific nucleotides were preserved in the 18S rRNA gene (one in each of the loops of hairpins 35 and 48), if random coincidences in hypervariable regions are not taken into consideration [127]. Second, such markers are very stable within groups having different taxonomic rank. Thus, analyzing the secondary structure of eukaryotic rRNA, we found discrete states for individual secondary-structure elements of rRNA that are markers of different eukaryotic taxa, whose rank ranges from supertaxa uniting phyla to orders and species groups within an individual phylum. A characteristic evolutionary feature of these markers is their stability, which is noted even in phyletic lineages with an abnormally high evolution rate of the 18S rRNA gene. For example, a nematode-specific symmetric helix structure in hairpin 17 is preserved in a rhabditid nematode *Pelodera strongyloides*, which is characterized by a record high evolution rate of the 18S rRNA gene [128]. Thus, individual elements of secondary structure can be good phylogenetic markers, which are particularly reliable in the determination of the position of rapidly evolving phyletic lineages. The examples of these markers will be presented in a special section.

THE MARKERS OF MAJOR TAXA DO NOT COMPLY WITH SIMPLE LAWS OF NEUTRAL EVOLUTION

The basis of modern paradigm of molecular evolution is formed by two closely related concepts: the hypothesis of significant neutral component in the evolution of macromolecules [93, 129] and the concept of molecular clock [130]. Actually, even in coding sequences, many nucleotide substitutions do not result in changes in the encoded protein because of degenerate amino-acid code, and, in the protein, many of substitutions of amino acid residues with chemically similar amino acid residues do not significantly affect its properties. The evolution of variable regions of macromolecules conforms with predictions of the neutral evolution and molecular clock theories. Permanent mutation and random propagation of individuals carrying different allele variants provide a permanent and rapid change of genes not related to directional selection. In species and populations with any degree of relationship, the results of this process are found in the regions of mitochondrial genomes, in spacers and variable regions of ribosomal operons, and in fractions of repetitive sequences. These genomic regions are used to discriminate taxonomic units of lower rank, i.e., populations and species [131–135].

Which would be specific features of the markers of major ancient taxa, whose subtaxa exist, for example, since the Cambrian period? Apparently, their common features must not be destroyed by neutral mutations, as in spacer and intron sequences. Can such evolutionary conservation be accompanied by neutrality of these markers, at least at the time of their origin? Does not this indicate that the patterns of evolution are different in the markers of major (phyla and classes) and minor taxa and that using the methods of molecular biology would help to draw the notorious distinction between micro- and macroevolution? Are the markers of major taxa nevertheless adaptive? However, in this case, it is difficult to imagine what can be an adaptive change in macromolecules fulfilling permanent functions in the cell. If the changes are adaptive, a new question arises that concerns the probability of convergence by these marker characters.

VARIABLE rRNA REGIONS AS A DEPOSITARY OF MARKER CHARACTERS

A paradoxical nature of the markers of major taxa is expressed in the situation that at least some of them are localized in evolutionarily variable regions of macromolecule. It is not easy to find out these markers, because the procedure of sequence alignment of hypervariable regions is difficult. The computer-aided search of them is particularly complicated because of their high variability. First, there is the possibility of random gap placement in the computer procedures of sequence alignment, which causes these markers to be excluded from the analysis; second, even if the alignment is accurate, these markers will according to their evolution rate be considered variable, and, consequently, of low value for phylogenetic analysis. In addition, the contribution of individual amino-acid residues to the support of the corresponding clades is low against the background of the great noise of extended regions with multiple substitutions in particular positions, especially in the case of using the distance-based methods of tree construction.

The examples of the markers of taxa of different rank, which were found using the analysis of secondary-structure elements in the 18S rRNA molecule, are presented below. The order of their presentation corresponds to the localization of these markers in the variable regions of the 18S rRNA molecule, which have a standard numeration from 1 to 9 [24].

Hypervariable Region V1

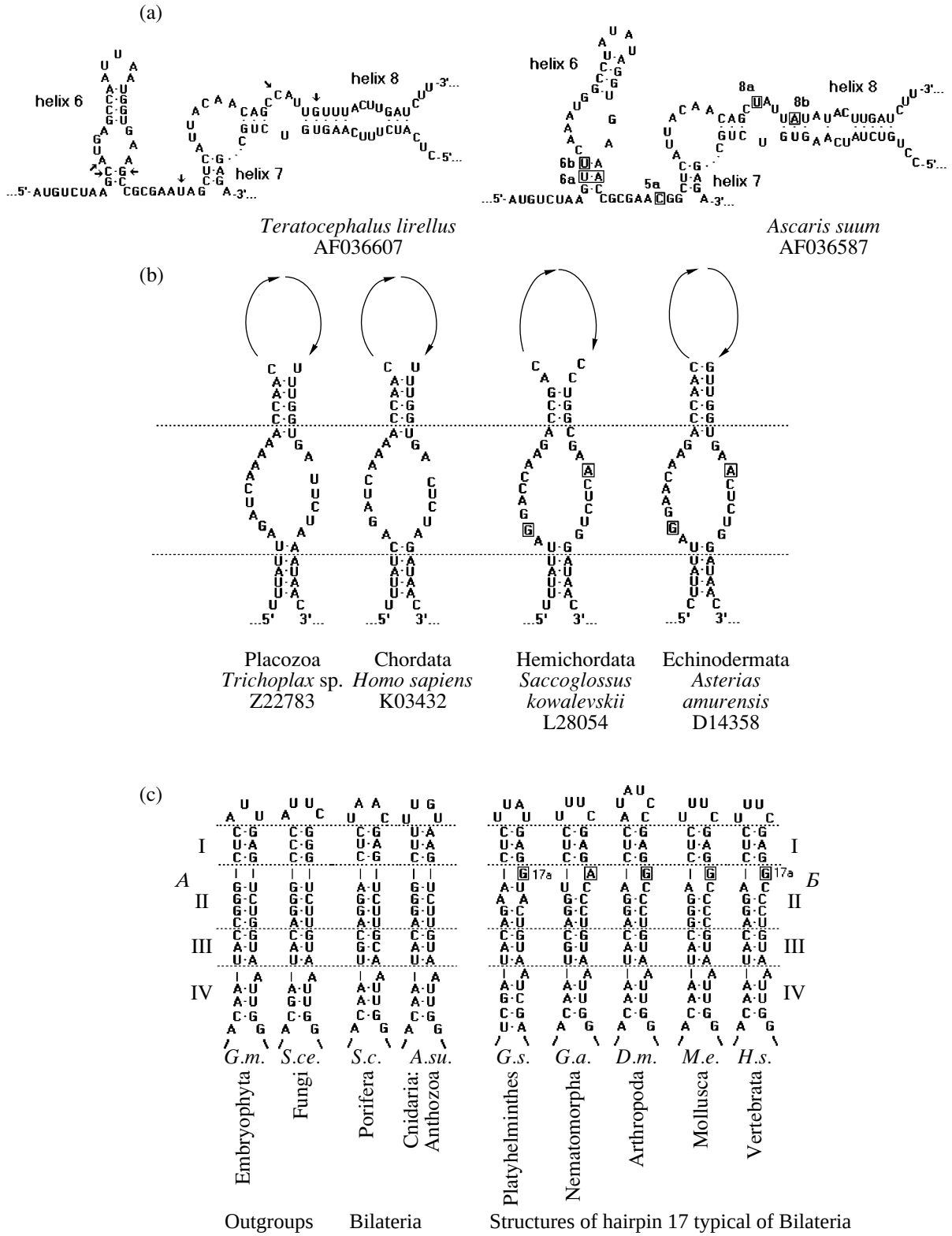
In this region, a set of six nucleotide substitutions was found in hairpin 6 and the adjacent hairpins 7 and 8 (Fig. 5a). This set is synapomorphic for secernentean nematodes that constitute the Rhabditia group of nematode orders, which is the nematode subtaxon characterized by the largest number of species included and the greatest importance for humans. The origin of the Rhabditia group has long been controversial, because its nearest ancestors that are assumed from the totality of morphological characters, the orders Teratocephalida and Plectida, fall within this group on the phylogenetic trees inferred from 18S rRNA genes and join to the spirurid–ascaridid orders, whose 18S rRNA genes appeared to be the least changed among Rhabditia. By contrast, morphologically primitive free-living Rhabditida occupy a basal position in relation to Teratocephalida and Plectida because of a marked divergence of the 18S rRNA genes and the consequent LBA effect.

This artifact is eliminated in view of the discovery of the aforementioned synapomorphies of Rhabditia in the hypervariable region V1 (Fig. 5a), which definitely indicate the monophyly of Rhabditia in relation to Teratocephalida and Plectida.

Hypervariable Region V2

The sequence of this region produces an eukaryote-specific hairpin E10_1. In the stem of this hairpin, an additional nucleotide pair (Fig. 5b) represents a unique synapomorphy of Hemichordata and Echinodermata that confirms the most important recent phylogenetic discovery on the relationship between the phyla of Deuterostomia, which was made based on a comparison of 18S rRNA genes. The scheme prevailing in textbooks, which consider Hemichordata (*Balanoglossus*, Pterobranchia) the nearest relatives of Chordata, appeared to be wrong in the light of molecular data. In phylogenetic trees inferred from 18S rRNA, Hemichordata and Echinodermata represent a single clade [136, 137], which confirms the monophyly of the unusual taxon Ambulacralia that was long ago proposed by Mechnikov on the basis of embryological observations [138]. The synapomorphy that represents an additional nucleotide pair in the helix of hairpin E10_1 in the hypervariable region V2 is a good marker of clade Ambulacralia. Chordata preserve a plesiomorphic (ancestral) state or this region.

Fig. 5. Secondary structure of some parts of hypervariable regions V1, V2, and V3. (a) Hairpin 6 and the adjacent hairpin 7 and 8 in *Teratocephalus lirellus* (Teratocephalida) and *Ascaris suum* (Rhabditia). The ancestral (plesiomorphic) state of these positions (Teratocephalida) and the derivative (apomorphic) state (Rhabditia) are designated by arrows and rectangles, respectively. (b) Hairpin E10_1 in Ambulacralia (Hemichordata + Echinodermata) and the remaining animal species. Additional nucleotides in the internal loop of the hairpin of Ambulacralia are boxed. (c) Hairpin 17 of variable region V3. (A) Ancestral state (symmetric upper part of hairpin) in all eukaryotes except Bilateria. (B) Derivative state (asymmetric upper part of hairpin) caused by insertion of additional nucleotide (G or A; designated as 17a). *G.m., Glycine max* X02623; *S.ce., Saccharomyces cerevisiae* J01353; *S.c. Scypha ciliata* L10827; *A.su., Anemonia sulcata* X53498; *G.s., Gyrodactylus salaris* Z26942; *G.a., Gordius albopunctatus* U88337; *D.m., Drosophila melanogaster* M21017; *M.e., Mytilus edulis* L24489; *H.s., Homo sapiens* K03432.



Hypervariable Region V3

Two main states of hairpin 17 (Fig. 5c) are observed in the secondary structure of the V3 region of 18S rRNA, of which one (symmetric helix 17) is ancestral (plesiomorphic), because it is found in most eukaryotes except Bilateria. The other state (asymmetric helix 17 due to the occurrence of an unpaired purine nucleotide) is derivative (apomorphic), because it is typical of most Bilateria [124]. In turn, several minor states are found within Bilateria, which can be specific of either individ-

ual phyla or subtaxa within phyla. For example, a reversion to the state resembling the ancestral symmetric helix 17 is typical of nematodes, and further modification of this state indicates the monophyly of some Rhabditida and Strongylida [128].

Hypervariable Region V4

Region V4 is the most variable region of the 18S rRNA molecule. This region attracts particular attention, because it contains supplementary hairpins that are eukaryote-specific secondary-structure elements. Some of these hairpins may be lost in some taxa whereas new supplementary hairpins may occur in other taxa. In a recent revision of the secondary structure of this region based on data on 3000 eukaryotic species, 11 major taxa were revealed, in which deviations from canonical helix shapes were found in all or some species [139]. For example, a supplementary helix was found in crustaceans belonging to the order Cladocera [98]. The same helix was found in Cyclestheria representing another group of crustaceans (Conchostraca), which indicates the monophyly of Cladocera and Cyclestheria. Unique supplementary hairpins that are found in the terminal part of this region are also typical of taxa having a higher rank, for example, Kinetoplastida and Euglenida [139].

Hypervariable Region V7

A specific two-nucleotide deletion and the corresponding changes in the predicted secondary structure of 18S rRNA in the region of hairpins 42 and 44 were found in bilaterians, coelenterates, and *Trichoplax* (Placozoa), whereas an ancestral state of these characters was found in unicellular eukaryotes, sponges, and Ctenophora [47]. This supports the closest relationship of Bilateria with Cnidaria rather than Ctenophora. A characteristic feature is that such animal taxa as Myxozoa and Mesozoa (whose phylogenetic position represents a complicated problem) also have a derivative state of these characters, which supports the monophyly of the Placozoa + Cnidaria + Myxozoa + Mesozoa + Bilateria group.

A plesiomorphic state of another secondary-structure element of 18S rRNA, hairpin 43, is represented by short helices (4 to 6 nucleotides) alternating with short unpaired regions that form bulges. The changes of hairpin 43 are related to more or less extended insertions into these bulges. These changes are mainly observed in taxonomic groups having a greater size of the 18S rRNA gene. Thus, the occurrence of extended insertions is followed by the formation of new specific helices, which results in an increase of the total hairpin size and in the corresponding change of its configuration. Such changes of hairpin 43 that are markers of taxa having an order rank were registered in Crustacea [98] and Strepsiptera [140].

Hypervariable Region V9

The hypervariable region V9 is packed into helix 49, which adjoins the decoding center of ribosomes. This region is characterized by a sharp gradient of evolutionary conservation from the positions of the basal part of helix that are virtually similar in all eukaryotes to the positions of the apical part of helix that are different even in closely related species. Nematode orders Enoplida and Dorylaimida preserved all evolutionarily conserved positions in this region, whereas 35 point substitutions, which form a specific secernentean stem of hairpin 49, were fixed in a large group of nematode orders (Rhabditida, Strongylida, Tylenchida, Ascaridida, Spirurida, Oxyurida, Teratocephalida, Plectida, Araeolaimida, and Monhysterida). Although the changes this region do not cause any predicted changes in the configuration and size of hairpin 49, they affect the most conserved sites, which discriminates these nematodes not only from other animals, but also from other eukaryotes. Thus, the secernentean stem supports the monophyly of Chromadoria and helps to determine the order of separation of major nematode taxa from the main trunk on their phylogenetic tree [141].

Nature and the Phylogenetic Importance of Conditionally Neutral Markers

The aforementioned examples show the existence of discrete states for the individual elements of the primary and secondary structure of eukaryotic rRNA, which are nonrandomly distributed throughout the phylogenetic tree. The alternative states of these elements are markers of different taxa, whose rank ranges from supertaxa that unite phyla to orders and species groups within an individual phylum. A characteristic feature of their evolution consists in the fact that, after a short-term transition from one state to another, they enter a phase of long-term stability, which is noted even in phyletic lineages with an abnormally high rate of evolution of the 18S rRNA gene. In essence, this group of molecular characters, which is chosen from their totality by their value as phylogenetic markers for ancient phyletic lineages, evolves by punctuated equilibrium [124]. Earlier, the punctuated equilibrium was described for morphological characters in paleontological series [142].

How we can explain the occurrence of punctuated equilibrium in the evolution of macromolecules? Physiological function of the stable but not absolutely invariant elements is mostly unknown. By analogy with the amino-acid residues that interact in a native ribosome and consequently undergo coordinated changes (for example, in the complementary strands of helices) [143], we can suppose that the amino-acid residues that are characterized by variable evolution rate are involved in the interaction with a ligand (RNA or protein). A mutation change of such a residue requires a coordinated compensation in the structure of ligand. If this compensation is realized, further reversion is ham-

pered, because a concurrent elimination of structural compensation of ligand is also required to preserve normal functioning. For this reason, such changes occur very rarely, and the changes that occur in the stem group are capable of clearly marking monophyletic lineages. As we see in the cases of the formation of secernentean stem [141] or the accumulation of additional changes outside the V4 region of Cicindellidae rRNA [92], the fixation of key mutations in a macromolecule promotes the fixation of a whole bunch of mutations, which therefore became allowed in evolutionarily conserved sites as well. They are realized in the subsequent evolution of this taxonomic group and are not found beyond its limits [141].

Thus, these markers must not be considered neutral. Their occurrence cannot be predicted judging from the time of existence of clade by calibration of molecular clock. However, they must not also be considered adaptive, because it is beyond reason to interpret the changes as adaptation to any new, particular functions. We suggest to name these markers conditionally neutral.

CONCLUSION

The current state of phylogenetics gives grounds to believe that this field, which is traditionally related to zoology and botany, will be fully supplied with tools that have been ascribed to molecular biology, while traditional methods based on morphological characters will progressively assume lesser and lesser importance, at least for recent forms. For example, in the field of microbiology, data on the 16S rRNA structure (which sometimes represent the only available information on noncultivated microorganisms found in natural habitats) are now required for characterization of any prokaryotic species [144]. In the near future, the complete genome sequences will be determined in all major taxa of bacteria. More than 50 complete genomes have already been included in public databases [145], and several times more genomes are classified as secret by research firms. Modern molecular methods provide the possibility of isolating and sequencing of the required gene in numerous species, including the cases of gene isolation from tissue fragments of rare museum specimens, subfossil remnants, or even individual protist cells extracted from the substrate using a micromanipulator.

A great number of efficient and sometimes very elaborate algorithms have been developed for molecular data analysis, and these algorithms are permanently being improved. Based on these algorithms, freely distributed user-friendly computer programs of tree construction have been designed, which are either particularized or included in program packages of general purpose. This gives the impression that the phylogenetic reconstruction actually becomes a purely technical task, which can be readily accomplished by methods that are immune from subjectivism. This view gains strength from the fact that the hypotheses on phylogenetic relationship between remote taxa by morphological or embryological characters appear to be vague and controversial, especially if they are limited to traditional arguments and do not involve genetic data on morphogenes.

However, at the current stage of development (which can be named the era of the 18S rRNA gene), molecular phylogenetics faces a number of problems. These problems are difficult to solve and are likely to pass to the next developmental stage of molecular phylogenetics. We are currently on the verge of this new stage, which can be named the genomic era of molecular phylogenetics. Actually, both LBA effect (which is related to the irregular evolution rate of single genes or complete genomes in different phyletic lineages) and the deficit of phylogenetically informative characters can hardly be overcome by the mere transition to a comparison of complete genomes. It can be supposed that, for example, the use of many genes or even complete genomes can enhance the LBA effect in some cases, in which not only individual genes, but also complete genomes, evolve at a high rate in some phyletic lineages. Due to LBA, multigenic comparisons of genes differing in evolution rate can be an additional source of the lack of coincidence of tree topology and can therefore reduce the potential of phylogenetically informative characters of complete genomes. The irregularity of distribution of phylogenetic signals by taxa can also be of negative value, for example, in view of actual differences between the trunk groups by the time of existence.

The modern theory of molecular evolution is based on the concepts of neutrality and molecular clock. They perfectly describe a wide range of genetic events that occur in reproductively isolated populations, but they are of little use for the description of a large class of adaptive or conditionally neutral characters with a variable evolution rate, which are under pressure of functional limitations. However, as discussed above, these characters are of supreme importance for solving phylogenetic problems concerned ancient phyletic lineages (phyla and classes). Thus, further advances of the new stage of molecular phylogenetics will be related to the development of novel ideas and new approaches to known phylogenetic problems and to further generalization of our knowledge about molecular evolution rather than to the progress of sequencing of complete genomes.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant nos. 99-04-48840, 00-15-97905, and 01-04-48832).

REFERENCES

1. Nuttal, G.H.F., *Blood Immunity and Blood Relationship*, Cambridge: Cambridge Univ. Press, 1904.

2. Brown, H., Sanger, F., and Kitai, R., The Structure of Pig and Sheep Insulins, *Biochem. J.*, 1955, vol. 60, no. 5, pp. 556–565.

3. Anfinsen, C.B., *The Molecular Basis of Evolution*, New York: Wiley, 1959.

4. Goodman, M., Poulik, E., and Poulik, M.D., Variations in the Serum Specificities of Higher Primates Detected by Two-Dimensional Starch Gel Electrophoresis, *Nature*, 1960, vol. 188, pp. 78–79.

5. Schildkraut, C.L., Marmur, J., and Doty, P., The Formation of Hybrid DNA Molecules and Their Use in Studies of DNA Homologies, *J. Mol. Biol.*, 1961, vol. 3, no. 5, pp. 596–617.

6. McCarthy, B.J. and Bolton, E.T., An Approach to the Measurement of Genetic Relatedness among Organisms, *Proc. Natl Acad. Sci. USA*, 1963, vol. 50, no. 1, pp. 156–164.

7. Hoyer, B.H., McCarthy, B.J., and Bolton, E.T., A Molecular Approach in the Systematics of Higher Organisms, *Science*, 1964, vol. 144, no. 3621, pp. 959–967.

8. Needleman, S.B. and Wunsch, C.D., A General Method Applicable to the Search for Similarities in the Amino Acid Sequence of Two Proteins, *J. Mol. Biol.*, 1970, vol. 48, no. 3, pp. 443–453.

9. Dayhoff, M.O., *Atlas of Protein Sequence and Structure*, vol. 5, suppl. 3, Washington, DC: Natl. Biomed. Res. Found., 1978.

10. Waterman, M.S., Smith, T.F., and Beyer, W.A., Some Biological Sequence Metrics, *Adv. Math.*, 1976, vol. 20, no. 4, pp. 367–387.

11. Cavalli-Sforza, L.L. and Edwards, A.W.F., Phylogenetic Analysis: Models and Estimation Procedures, *Am. J. Hum. Genet.*, 1967, vol. 19, no. 3, pp. 233–257.

12. Fitch, W.M. and Margoliash, E., Construction of Phylogenetic Trees, *Science*, 1967, vol. 155, no. 760, pp. 279–284.

13. Felsenstein, J., Maximum-Likelihood and Minimum-Steps Methods for Estimating Evolutionary Trees from Data on Discrete Characters, *Syst. Zool.*, 1973, vol. 22, no. 2, pp. 240–249.

14. Nei, M., Interspecific Gene Differences and Evolutionary Time Estimated from Electrophoretic Data on Protein Identity, *Am. Nat.*, 1971, vol. 105, no. 4, pp. 385–398.

15. Mednikov, B.M., Popov, L.S., Antonov, A.S., and Belozerskii, A.N., Features of the DNA Primary Structure as a Criterion for Constructing a Natural Fish Taxonomic System, *Zh. Obshch. Biol.*, 1973, vol. 34, no. 4, pp. 516–529.

16. Maxson, L.R. and Wilson, A.C., Albumin Evolution and Organismal Evolution in Tree Frogs (Hylidae), *Syst. Zool.*, 1975, vol. 24, no. 1, pp. 1–15.

17. Poltaraus, A.B., Petrov, N.B., and Antonov, A.S., Estimation of the Extent of Ribosomal Gene Conservation in Several Phyla of Invertebrates, *Biokhimiya*, 1980, vol. 45, no. 1, pp. 165–172.

18. Sibley, C.G. and Ahlquist, J.E., *Phylogeny and Classification of the Birds of the World: A Study in Molecular Evolution*, Yale Univ. Press, 1990.

19. Hillis, D.M. and Dixon, M.T., Ribosomal DNA: Molecular Evolution and Phylogenetic Inference, *Quart. Rev. Biol.*, 1991, vol. 66, no. 4, pp. 411–453.

20. Friedlander, T.P., Regier, J.C., and Mitter, C., Nuclear Gene Sequences for Higher Level Phylogenetic Analysis: 14 Promising Candidates, *Syst. Biol.*, 1992, vol. 41, no. 5, pp. 483–490.

21. Sidow, A.R. and Thomas, W.K., A Molecular Evolutionary Framework for Eukaryotic Model Organisms, *Curr. Biol.*, 1994, vol. 4, no. 7, pp. 596–603.

22. Hori, H. and Osawa, S., Origin and Evolution of Organisms as Deduced from 5S Ribosomal RNA Sequences, *Mol. Biol. Evol.*, 1987, vol. 4, no. 5, pp. 445–472.

23. Joffe, B.I., Valiejo Roman, K.M., Birstein, V.Ya., and Troitsky, A.V., 5S rRNA Sequences of 12 Species of Flatworms: Implications for the Phylogeny of the Platyhelminthes, *Hydrobiology*, 1995, vol. 305, no. 1, pp. 37–43.

24. Van de Peer, Y., De Rijk, P., Wuyts, J., *et al.*, The European Small-Subunit Ribosomal RNA Database, *Nucleic Acids Res.*, 2000, vol. 28, no. 1, pp. 175–176.

25. Garey, J.R. and Schmidt-Rhaesa, A., The Essential Role of "Minor" Phyla in Molecular Studies of Animal Evolution, *Am. Zool.*, 1998, vol. 38, no. 6, pp. 907–917.

26. Adoutte, A., Balavoine, G., Lartillot, N., *et al.*, The New Animal Phylogeny: Reliability and Implications, *Proc. Natl Acad. Sci. USA*, 2000, vol. 97, no. 9, pp. 4453–4456.

27. Giribet, G. and Ribera, C., The Position of Arthropods in the Animal Kingdom: A Search for a Reliable Outgroup for Internal Arthropod Phylogeny, *Mol. Phyl. Evol.*, 1998, vol. 9, no. 3, pp. 481–488.

28. Giribet, G., Distel, D., Polz, M., *et al.*, Triploblasic Relationships with Emphasis on the Acoelomates and the Position of Gnathostomulida, Cycliophora, Plathelminthes, and Chaetognatha: A Combined Approach of 18S rDNA Sequences and Morphology, *Syst. Biol.*, 2000, vol. 49, no. 3, pp. 539–562.

29. Peterson, K.J. and Eernisse, D.J., Animal Phylogeny and the Ancestry of Bilaterians: Interferences from Morphology and 18S rDNA Gene Sequences, *Evol. Dev.*, 2001, vol. 3, no. 3, pp. 170–205.

30. Zrzavý, J., Mihulka, S., Kepka, P., *et al.*, Phylogeny of the Metazoa Based on Morphological and 18S Ribosomal DNA Evidence, *Cladistics*, 1998, vol. 14, no. 3, pp. 249–285.

31. Mikryukov, K.A., Ultrastructural Analysis and Comparison of the Ribosomal RNA Genes as Methods to Construct Protist Taxonomy, *Zool. Zh.*, 1999, vol. 78, no. 8, pp. 901–915.

32. Antonov, A.S., Are There Molecular Prerequisites to Revising the Phylogeny and Taxonomy of Higher Plants?, *Zh. Obshch. Biol.*, 1999, vol. 60, no. 3, pp. 245–276.

33. Dogel', V.A., *Zoologiya bespozvonochnykh* (Zoology of Invertebrates), Moscow: Vysshaya Shkola, 1981.

34. Salvini-Plawen, L.V., Early Evolution and the Primitive Groups, *Mollusca*, 1985, vol. 10, no. 2, pp. 59–150.

35. Meglitsch, P.A. and Schram, F.R., *Invertebrate Zoology*, Oxford Univ. Press, 1991, 3rd ed.

36. Ruppert, E.E., Introduction to the Aschelminth Phyla: A Consideration of Mesoderm, Body Cavities, and Cuticle, *Microscopic Anatomy of Invertebrates*, vol. 4: *Aschelminthes*, Harrison, F.W. and Ruppert, E.E., Eds., New York: Wiley–Liss, 1991, pp. 1–17.
37. Kristensen, R.M., Are Aschelminthes Pseudocoelomate or Acoelomate?, *Body Cavities: Function and Phylogeny: Selected Symposia and Monographs U.Z.I., 8*, Lanzavvecchia, G., Valvassori, R., and Candia Carnevalli, M.D., Eds., Modena: Mucchi, 1995, pp. 41–43.
38. Wallace, R.L., Ricci, C., and Melone, G., A Cladistic Analysis of Pseudocoelomate (Aschelminth) Morphology, *Invert. Biol.*, 1996, vol. 115, no. 2, pp. 104–112.
39. Zavarzina, E.G. and Tsetlin, A.B., Structure of the Body Cavity in Several Annelidae: Preliminary Results, *Zool. Zh.*, 1990, vol. 69, no. 7, pp. 31–41.
40. Beklemishev, K.V., On the Possible and Actual Directions of Invertebrate Evolution, *Zh. Obshch. Biol.*, 1974, vol. 35, no. 2, pp. 209–222.
41. Nielsen, C., Scharf, N., and Eibye-Jacobsen, D., Cladistic Analyses of the Animal Kingdom, *Biol. J. Linn. Soc.*, 1996, vol. 57, pp. 385–410.
42. Wainright, P.O., Hinkle, G., Sogin, M.L., and Stickel, S.K., Monophyletic Origins of the Metazoa: An Evolutionary Link with Fungi, *Science*, 1993, vol. 260, no. 5106, pp. 340–342.
43. Katayama, T., Wada, H., Furuya, H., *et al.*, Phylogenetic Position of the Dicyemid Metazoa Inferred from 18S rDNA Sequences, *Biol. Bull.*, 1995, vol. 189, no. 2, pp. 81–90.
44. Pawlowski, J., Montoya-Burgos, J.I., Fahrni, J.F., *et al.*, Origin of the Mesozoa Inferred from 18S rRNA Gene Sequences, *Mol. Biol. Evol.*, 1996, vol. 13, no. 8, pp. 1128–1132.
45. Kobayashi, M., Furuya, H., and Holland, P.W.H., Dicyemids Are Higher Animals, *Nature*, 1999, vol. 401, no. 6755, p. 762.
46. Medina, M., Collins, A.G., Silberman, J., and Sogin, M., Evaluating Hypotheses of Basal Animal Phylogeny Using Complete Sequences of Large and Small Subunit rRNA, *Proc. Natl. Acad. Sci. USA*, 2001, vol. 98, no. 17, pp. 9707–9712.
47. Aleshin, V.V. and Petrov, N.B., Molecular Evidence for Regress in Evolution of Multicellular Animals, *Zh. Obshch. Biol.*, 2002, vol. 63, no. 3, pp. 195–208.
48. Dewel, R.A., Colonial Origin for Metazoa: Major Morphological Transition and the Origin of Bilaterian Complexity, *J. Morphol.*, 2000, vol. 243, no. 1, pp. 35–74.
49. Smothers, J.F., von Dohlen, C.D., Smith, L.H., Jr., and Spall, R.D., Molecular Evidence That the Myxozoan Protists Are Metazoans, *Science*, 1994, vol. 265, no. 5179, pp. 1719–1721.
50. Cavalier-Smith, T., Allosopp, M.T.E.P., Chao, E.E., *et al.*, Sponge Phylogeny, Animal Monophyly, and the Origin of the Nervous System: 18S rRNA Evidence, *Can. J. Zool.*, 1996, vol. 74, no. 11, pp. 2031–2045.
51. Anderson, C.R., Canning, E.U., and Okamura, B., A Triploblast Origin for Myxozoa?, *Nature*, 1998, vol. 392, no. 6674, pp. 346–347.
52. Siddal, M.E., Martin, D.S., Bridge, D., *et al.*, The Demise of a Phylum of Protists: Phylogeny of Myxozoa and Other Parasitic Cnidarians, *J. Parasitol.*, 1995, vol. 81, no. 6, pp. 961–967.
53. Kent, M.L., Andree, K.P., Bartholomew, J.P., *et al.*, Recent Advances in Our Knowledge of the Myxozoa, *J. Eukaryot. Microbiol.*, 2001, vol. 48, no. 4, pp. 395–413.
54. Halanych, K.M., Bacheller, J.D., Aguinaldo, A.M., *et al.*, Evidence from 18S Ribosomal DNA That the Lophophorates Are Protostome Animals, *Science*, 1995, vol. 267, no. 5204, pp. 1641–1643.
55. Conway Morris, S., Cohen, B.L., Gawthrop, A.P., *et al.*, Lophophorate Phylogeny, *Science*, 1996, vol. 272, no. 5259, pp. 282–283.
56. Turbeville, J.M., Field, K.G., and Raff, R.A., Phylogenetic Position of Phylum Nemertini, Inferred from 18S rRNA Sequences: Molecular Data as a Test of Morphological Character Homology, *Mol. Biol. Evol.*, 1992, vol. 9, no. 2, pp. 235–249.
57. Winnepenninckx, B., Backeljau, T., and De Wachter, R., Phylogeny of Protostome Worms Derived from 18S rRNA Sequences, *Mol. Biol. Evol.*, 1995, vol. 12, no. 4, pp. 641–649.
58. Ivanov, A.V., Pogonophora, *Fauna SSSR* (Fauna of the Soviet Union), Leningrad: Nauka, vol. 75.
59. Ushakov, P., Eine neue Form aus der Familie Sabellidae (Polychaeta), *Zool. Anz.*, 1933, vol. 104, pp. 205–208.
60. Rouse, G.W. and Fauchald, K., Cladistics and Polychaetes, *Zool. Scripta*, 1997, vol. 26, no. 2, pp. 139–204.
61. Aguinaldo, A.M., Turbeville, J.M., Linford, L.S., *et al.*, Evidence for a Clade Nematodes, Arthropods and Other Molting Animals, *Nature*, 1997, vol. 387, no. 6632, pp. 489–493.
62. Eernisse, D.L., Albert, J.S., and Anderson, F.E., Annelida and Arthropoda Are Not Sister Taxa: A Phylogenetic Analysis of Spiralian Metazoan Morphology, *Syst. Biol.*, 1992, vol. 41, no. 3, pp. 305–330.
63. Dzik, J. and Krumbiegel, G., The Oldest "Onychophoran" Xenusion: A Link Connecting Phyla?, *Lethaia*, 1989, vol. 22, no. 2, pp. 169–181.
64. Carranza, S., Baguñà, J., and Riutort, M., Are the Platyhelminthes a Monophyletic Primitive Group? An Assessment Using 18S rDNA Sequences, *Mol. Biol. Evol.*, 1997, vol. 14, no. 5, pp. 485–497.
65. Ruiz-Trillo, I., Riutort, M., Littlewood, D.T., *et al.*, Acoel Flatworms: Earliest Extant Bilaterian Metazoans, Not Members of Platyhelminthes, *Science*, 1999, vol. 283, no. 5409, pp. 1919–1923.
66. Fitch, W.M., Aspects of Molecular Evolution, *Annu. Rev. Genet.*, 1973, vol. 7, pp. 343–380.
67. Arnheim, N., Krystal, M., Schmickel, R., *et al.*, Molecular Evidence for Genetic Exchange among Ribosomal Genes on Nonhomologous Chromosomes in Man and Apes, *Proc. Natl. Acad. Sci. USA*, 1980, vol. 77, no. 12, pp. 7323–7327.
68. Dover, G.A., A Molecular Drive through Evolution, *BioScience*, 1982, vol. 32, no. 6, pp. 526–533.
69. Dover, G., Braun, S., Koen, E., *et al.*, Dynamics of Genome Evolution and Species Differentiation, *Genome Evolution*, Dover, G.A. and Flavell, R.B., Eds., London: Academic, 1982, pp. 329–356.

70. Ratner, V.A., Zharkikh, A.A., Kolchanov, N.A., *Problems of the Theory of Molecular Evolution,* Novosibirsk: Nauka, 1985.
71. Carranza, S., Giribet, G., Ribera, C., *et al.*, Evidence That Two Types of 18S rDNA Coexist in the Genome of *Dugesia* (*Schmidtea*) *mediterranea* (Platyhelminthes, Turbellaria, Tricladida), *Mol. Biol. Evol.*, 1996, vol. 13, no. 6, pp. 824–832.
72. Carranza, S., Baguna, J., and Riutort, M., Origin and Evolution of Paralogous rRNA Gene Clusters within the Flatworm Family Digesiidae (Platyhelminthes, Tricladida), *J. Mol. Evol.*, 1999, vol. 49, no. 2, pp. 250–259.
73. Britten, R.J., Rates of DNA Sequence Evolution Differ between Taxonomic Groups, *Science*, 1986, vol. 231, no. 4744, pp. 1393–1398.
74. Philippe, H., Opinion: Long Branch Attraction and Protist Phylogeny, *Protist*, 2000, vol. 151, no. 4, pp. 307–316.
75. Wheeler, W.C., Nucleic Acid Sequence Phylogeny and Random Outgroups, *Cladistics*, 1990, vol. 6, no. 4, pp. 363–367.
76. Sarich, V.M. and Wilson, A.C., Generation Time and Genomic Evolution in Primates, *Science*, 1973, vol. 179, no. 78, pp. 1144–1147.
77. Germot, A., Philippe, H., and Le Guyader, H., Evidence for Loss of Mitochondria in Microsporidia from a Mitochondrial-Type HSP70 in *Nosema locustae*, *Mol. Biochem. Parasitol.*, 1997, vol. 87, no. 2, pp. 159–168.
78. Germot, A. and Philippe, H., Critical Analysis of Eukaryotic Phylogeny: A Case Study Based on the HSP70 Family, *J. Eukaryot. Microbiol.*, 1999, vol. 46, no. 2, pp. 116–124.
79. Bhattacharya, D., Stickel, S.K., and Sogin, M.L., Molecular Phylogenetic Analysis of Actin Genic Regions from *Achlya bisexualis* (Oomycota) and *Costaria costata* (Chromophyta), *J. Mol. Evol.*, 1991, vol. 33, no. 6, pp. 525–536.
80. Stiller, J. and Hall, B., Long-Branch Attraction and the rDNA Model of Early Eukaryotic Evolution, *Mol. Biol. Evol.*, 1999, vol. 16, no. 9, pp. 1270–1279.
81. Huelsenbeck, J.P., Is a Felsenstein Zone a Fly Trap?, *Syst. Biol.*, 1997, vol. 46, no. 1, pp. 69–74.
82. Lyons-Weiler, J. and Holzer, G.A., Escaping from the Felsenstein Zone by Detecting Long Branches in Phylogenetic Data, *Mol. Phylogenet. Evol.*, 1997, vol. 8, no. 3, pp. 375–384.
83. Siddal, M.E. and Whiting, M.F., Long-Branch Abstraction, *Cladistics*, 1999, vol. 15, no. 1, pp. 9–24.
84. Berney, C., Pawlowski, J., and Zaninetti, L., Elongation Factor 1α Sequences Do Not Support an Early Divergence of the Acoela, *Mol. Biol. Evol.*, 2000, vol. 17, no. 7, pp. 1032–1039.
85. Uspenskaya, A.V. and Raikova, E.V., Cytological Aspects of the Similarity and Difference between Myxozoa and Cnidaria, *Tsitologiya*, 2001, vol. 43, no. 3, pp. 284–309.
86. Whiting, M.F., Carpenter, J.C., Wheeler, Q.D., and Wheeler, W.C., The Strepsiptera Problem: Phylogeny of the Holometabolous Insect Orders Inferred from 18S and 28S Sequences and Morphology, *Syst. Biol.*, 1997, vol. 46, no. 1, pp. 1–68.
87. Carmean, D. and Crespi, B.J., Do Long-Branches Attract Flies?, *Nature*, 1995, vol. 373, no. 6516, p. 666.
88. Li, W.-H., *Molecular Evolution*, Sunderland, Mass.: Sinauer, 1997.
89. Kohne, D.E., Evolution of Higher-Organism DNA, *Q. Rev. Biophys.*, 1970, vol. 3, no. 3, pp. 1–48.
90. Martin, A.P. and Palumbi, S.R., Body Size, Metabolic Rate, Generation Time, and the Molecular Clock, *Proc. Natl. Acad. Sci. USA*, 1993, vol. 90, no. 9, pp. 4087–4091.
91. Wolfe, K.H., Sharp, R.M., and Li, W.-H., Mutation Rates Differ among Regions of the Mammalian Genome, *Nature*, 1989, vol. 337, no. 6204, pp. 283–285.
92. Hancock, J.M. and Vogler, A.P., How Slippage-Derived Sequences Are Incorporated into rRNA Variable-Region Secondary Structure: Implications for Phylogeny Reconstruction, *Mol. Phylogenet. Evol.*, 2000, vol. 14, no. 3, pp. 366–374.
93. Kimura, M., *The Neutral Theory of Molecular Evolution*, Cambridge: Cambridge Univ., 1983.
94. Mushegian, A.R., Garey, J.R., Martin, J., and Liu, L.X., Large-Scale Taxonomic Profiling of Eukaryotic Model Organisms: A Comparison of Orthologous Proteins Encoded by the Human, Fly, Nematode, and Yeast Genomes, *Genome Res.*, 1998, vol. 8, no. 6, pp. 590–598.
95. Palumbi, S.R., Cipriano, F., and Hare, M.P., Predicting Nuclear Gene Coalescence from Mitochondrial Data: The Three-Times Rule, *Evol. Int. J. Org. Evol.*, 2001, vol. 55, no. 5, pp. 859–868.
96. Moreira, D., Le Guyadier, H., and Philippe, H., Unusually High Evolutionary Rate of the Elongation Factor 1α Genes from the Ciliophora and Its Impact on the Phylogeny of Eukaryotes, *Mol. Biol. Evol.*, 1999, vol. 16, no. 2, pp. 234–245.
97. Wimberly, B.T., Brodersen, D.E., Clemons, W.M., Jr., *et al.*, Structure of the 30S Ribosomal Subunit, *Nature*, 2000, vol. 407, no. 6802, pp. 327–339.
98. Crease, T.J. and Taylor, D.J., The Origin and Evolution of Variable-Region Helices in V4 and V7 of the Small-Subunit Ribosomal RNA of Branchiopod Crustaceans, *Mol. Biol. Evol.*, 1998, vol. 15, no. 11, pp. 1430–1446.
99. Fitch, W.M., On the Problem of Discovering the Most Parsimonious Tree, *Am. Nat.*, 1977, vol. 111, no. 2, pp. 223–257.
100. Cavender, J.A., Taxonomy with Confidence, *Math. Biosci.*, 1978, vol. 103, no. 3, pp. 271–280.
101. Felsenstein, J., Cases in Which Parsimony or Compatibility Methods Will Be Positively Misleading, *Syst. Zool.*, 1978, vol. 27, no. 4, pp. 401–410.
102. Hillis, D.M., Inferring Complex Phylogenies, *Nature*, 1996, vol. 383, no. 6596, pp. 130–131.
103. Kim, J., Kim, W., and Cunningham, C.W., A New Perspective on Lower Metazoan Relationships from 18S rDNA Sequences, *Mol. Biol. Evol.*, 1999, vol. 16, no. 3, pp. 423–427.
104. Roger, A.J., Svard, S.G., Tovar, J., *et al.*, A Mitochondrial-like Chaperonine 60 Gene in *Giardia lamblia*: Evidence That Diplomonads Once Harbored an Endosymbiont Related to the Progenitor of Mitochondria,

Proc. Natl. Acad. Sci. USA, 1998, vol. 95, no. 1, pp. 229–234.

105. Hendy, M. and Penny, D., A Framework for the Quantitative Study of Evolutionary Trees, *Syst. Zool.*, 1989, vol. 38, no. 3, pp. 297–309.
106. Sims, G.P., Rogerson, A., and Aitken, R., Primary and Secondary Structure of the Small-Subunit Ribosomal RNA of the Naked, Marine Amoeba *Vannella anglica*: Phylogenetic Implications, *J. Mol. Evol.*, 1999, vol. 48, no. 6, pp. 740–749.
107. Milyutina, I.A., Aleshin, V.V., Mikrjukov, K.A., *et al.*, The Unusually Long Small-Subunit Ribosomal RNA Gene Found in Amitochondriate Amoeboflagellate *Pelomyxa palustris*: Its rRNA Predicted Secondary Structure and Phylogenetic Implication, *Gene*, 2001, vol. 272, nos. 1–2, pp. 131–139.
108. Olsen, G., Earliest Phylogenetic Branching: Comparing rRNA-Based Evolutionary Trees Inferred with Various Techniques, *Cold Spring Harb. Symp.*, 1987, vol. 52, pp. 825–837.
109. Van de Peer, Y., Rensing, S.A., Maier, U.G., and De Wachter, R., Substitution Rate Calibration of Small-Subunit Ribosomal RNA Identifies Chlorarachniophyte Endosymbionts as Remnants of Green Algae, *Proc. Natl Acad. Sci. USA*, 1996, vol. 93, no. 15, pp. 7732–7736.
110. Philippe, H., Lopez, P., Brinkmann, H., *et al.*, Early Branching or Fast Evolving Eukaryotes? An Answer Based on Slowly Evolving Positions, *Philos. Trans. R. Soc. London, B*, 2000, vol. 267, no. 1449, pp. 1213–1221.
111. Vossbrinck, C.R., Maddox, J.V., Friedman S., *et al.*, Ribosomal RNA Sequence Suggests Microsporidia Are Extremely Ancient Eukaryotes, *Nature*, 1987, vol. 326, no. 6111, pp. 411–414.
112. Kamaishi, T., Hashimoto, T., Nakamura, Y. *et al.*, Protein Phylogeny of Translation Elongation Factor EF-1α Suggests Microsporidians Are Extremely Ancient Eukaryotes, *J. Mol. Evol.*, 1996, vol. 42, no. 2, pp. 257–263.
113. Edlind, T.D., Li, J., Visvesvara, G.S., *et al.*, Phylogenetic Analysis of β-Tubulin Sequences from Amitochondrial Protozoa, *Mol. Phylogenet. Evol.*, 1996, vol. 5, no. 2, pp. 359–367.
114. Keeling, P. and Doolittle, W.F., α-Tubulin from Early-Diverging Eukaryotic Lineages and the Evolution of the Tubulin Family, *Mol. Biol. Evol.*, 1996, vol. 13, no. 10, pp. 1297–1305.
115. Keeling, P., Luker, M.A., and Palmer, J.D., Evidence from β-Tubulin Phylogeny That Microsporidia Evolved from within the Fungi, *Mol. Biol. Evol.*, 2000, vol. 17, no. 1, pp. 23–31.
116. Van de Peer, Y., Ben Ali, A., and Meyer, A., Microsporidia: Accumulating Molecular Evidence That a Group of Amitochondriate and Suspectedly Primitive Eukaryotes Are Just Curious Fungi, *Gene*, 2000, vol. 246, nos. 1–2, pp. 1–8.
117. Van de Peer, Y., Neefs, J.-M., De Rijk, P., and De Wachter, R., Reconstructing Evolution from Eukaryotic Small Ribosomal Subunit RNA Sequences: Calibration of the Molecular Clock, *J. Mol. Evol.*, 1993, vol. 37, no. 2, pp. 221–232.
118. Yang, Z., Among-Site Rate Variation and Its Impact on Phylogenetic Analyses, *Trends Ecol. Evol.*, 1996, vol. 11, no. 4, pp. 367–370.
119. Van de Peer, Y. and De Wachter, R., TREECON for Windows: A Software Package for the Construction and Drawing of Evolutionary Trees for the Microsoft Windows Environment, *Comput. Appl. Biosci.*, 1994, vol. 10, no. 5, pp. 569–570.
120. Strimmer, K. and von Haeseler, A., Quartet Puzzling: A Quartet Maximum Likelihood Method for Reconstructing Tree Topologies, *Mol. Biol. Evol.*, 1996, vol. 13, no. 2, pp. 964–969.
121. Olsen, G.J., Matsuda, H., Hagstrom, R., and Overbeek, R., Fast DNAml: A Tool for Construction of Phylogenetic Trees of DNA Sequences Using Maximum Likelihood, *Comput. Appl. Biosci.*, 1994, vol. 10, no. 1, pp. 41–48.
122. Felsenstein, J., *PHYLIP* (Phylogeny Inference Package) Version 3.5c., Seattle: Univ. Washington, 1993.
123. Swofford, D.L., *PAUP*: Version 4.0*, Sunderland, Mass.: Sinauer, 1994.
124. Aleshin, V.V., Vladychenskaya, N.S., Kedrova, O.S., *et al.*, The Secondary Structure of Hairpin 17 of Primitive Multicellular Animal *Rhopalura ophiocomae* (Mesozoa: Orthonectida) as an Example of "Punctuated Equilibrium" in Evolution of the 18S Ribosomal RNA, *Mol. Biol.* (Moscow), 1999, vol. 33, no. 2, pp. 319–329.
125. Felsenstein, J., Evolutionary Trees from DNA Sequences: A Maximum Likelihood Approach, *J. Mol. Evol.*, 1981, vol. 17, no. 6, pp. 368–376.
126. Billoud, B., Guerruchi, M.-A., Masselot, M., and Deutsch, J.S., Cirripede Phylogeny Using a Novel Approach: Molecular Morphometrics, *Mol. Biol. Evol.*, 2000, vol. 17, no. 10, pp. 1435–1445.
127. Rusin, L.Yu., Aleshin, V.V., Vladychenskaya, N.S., *et al.*, The Primary Structure of Loops 35 and 48 of the Small rRNA Suggests that Trefusiidae Are a Subtaxon of Enoplida (Nematoda), *Mol. Biol.* (Moscow), 2001, vol. 35, no. 5, pp. 912–919.
128. Aleshin, V.V., Kedrova, O.S., Milyutina, I.A., *et al.*, Secondary Structure of Some Elements of 18S rRNA Suggests That Strongylid and a Part of Rhabditid Nematodes Are Monophyletic, *FEBS Lett.*, 1998, vol. 429, no. 1, pp. 4–8.
129. King, J.L. and Jukes, T.H., Non-Darwinian Evolution, *Science*, 1969, vol. 164, no. 881, pp. 788–798.
130. Zuckerkandl, E. and Pauling, L., Molecules as Documents of Evolutionary History, *J. Theor. Biol.*, 1965, vol. 8, no. 2, pp. 357–366.
131. Grechko, V.V., Fedorova, L.V., Fedorov, A.N., *et al.*, Restriction Mapping of Highly Repetitive DNA Sequences in Studying the Genetic Relationships among Primitive Animal Taxa, *Mol. Biol.* (Moscow), 1997, vol. 31, no. 2, pp. 244–252.
132. Brykov, V.A., Kirillova, O.N., Kukhlevskii, A.D., *et al.*, Analysis of the Mitochondrial DNA Variation in Calico Salmon *Oncorhynchus keta* (Walbaum) from Populations of Primorye and Sakhalin Rivers, *Genetika* (Moscow), 2000, vol. 36, no. 10, pp. 1388–1393.
133. Filipenko, M.L., Timofeeva, O.A., Gusachenko, A.M., *et al.*, Phylogeny Reconstruction in Grasshoppers of the Family Acrididae (Orhoptera) on the Basis of Nucleotide Sequences of the Mitochondrial 16S Ribosomal

RNA Gene, *Genetika* (Moscow), 2000, vol. 36, no. 10, pp. 1355–1361.

134. Sherbakov, D.Y., Kamaltynov, R.M., Ogarkov, O.B., and Verheyen, E., Patterns of Evolutionary Change in Baikalian Gammarids Inferred from DNA Sequences (Crustacea, Amphipoda), *Mol. Phylogenet. Evol.*, 1998, vol. 10, no. 2, pp. 160–167.
135. Martin, P., Kaygorodova, I., Sherbakov, D.Y., and Verheyen, E., Rapidly Evolving Lineages Impede the Resolution of Phylogenetic Relationships among Clitellata (Annelida), *Mol. Phylogenet. Evol.*, 2000, vol. 15, no. 3, pp. 355–368.
136. Halanych, K.M., The Phylogenetic Position of the Pterobranch Hemichordates Based on 18S rDNA Sequence Data, *Mol. Phylogenet. Evol.*, 1995, vol. 4, no. 1, pp. 72–76.
137. Cameron, C.B., Garey, J.R., and Swalla, B.J., Evolution of the Chordate Body Plan: New Insights from Phylogenetic Analyses of Deuterostome Phyla, *Proc. Natl. Acad. Sci. USA*, 2000, vol. 97, no. 9, pp. 4469–4474.
138. Metschnikoff, E., Uber die systematische Stellung der Balanoglossus, *Zool. Anz.*, 1881, vol. 15, no. 78, pp. 139–143.
139. Wuyts, J., De Rijk, P., Van de Peer, Y., *et al.*, Comparative Analysis of More Than 3000 Sequences Reveals the Existence of Two Pseudoknots in Area V4 of Eukaryotic Small-Subunit Ribosomal RNA, *Nucleic Acids Res.*, 2000, vol. 28, no. 23, pp. 4698–4708.
140. Choe, C.P., Hwang, U.W., and Kim, W., Putative Secondary Structures of Unusually Long Strepsipteran SSU rRNAs and Its Phylogenetic Implications, *Mol. Cells*, 1999, vol. 9, no. 2, pp. 191–199.
141. Aleshin, V.V., Kedrova, O.S., Milyutina, I.A., *et al.*, Relationships among Nematodes Based on the Analysis of 18S rRNA Gene Sequences: Molecular Evidence for Monophyly of Chromadorian and Secernentean Nematodes, *Russ. J. Nematol.*, 1998, vol. 6, no. 2, pp. 175–184.
142. Eldredge, N. and Gould, S.J., On Punctuated Equilibria, *Science*, 1997, vol. 276, no. 5311, pp. 338–341.
143. Woese, C.R., Gutell, R., Gupta, R., and Noller, H.F., Detailed Analysis of the Higher-Order Structure of 16S-like Ribosomal Ribonucleic Acids, *Microbiol. Rev.*, 1983, vol. 47, no. 4, pp. 621–669.
144. Turova, T.P., The Use of the Data of DNA–DNA Hybridization and Analysis of the 16S rRNA Genes to Solve Taxonomic Problems, as Demonstrated with the Example of the Order Haloanaerobiales, *Mikrobiologiya*, 2000, vol. 69, no. 6, pp. 741–752.
145. Borinskaya, S.A. and Yankovskii, N.K., Structure of the Prokaryotic Genomes, *Mol. Biol.* (Moscow), 1999, vol. 33, no. 6, pp. 941–957.