Microbiology, Vol. 72, No. 4, 2003, pp. 389–402. Translated from Mikrobiologiya, Vol. 72, No. 4, 2003, pp. 437–452.

REVIEW PAPER

Copy Number of Ribosomal Operons in Prokaryotes and Its Effect on Phylogenetic Analyses

T. P. Tourova

Institute of Microbiology, Russian Academy of Sciences, pr. 60-letiya Oktyabrya 7, k. 2, Moscow 117312, Russia
e-mail: ttour@biengi.ac.ru

Received May 20, 2002

Abstract—Different aspects of the presence of multiple copies of ribosomal operons in prokaryotic genomes are reviewed. The structure of prokaryotic ribosomal operons is briefly described. The available data are summarized regarding the copy number of ribosomal genes in various prokaryotic genomes, the degree of polymorphism of their individual copies, and physiological and evolutional aspects of the presence of the multiple copies of ribosomal genes. The review also considers the influence of the presence of multiple copies of ribosomal genes on the results of identification of prokaryotic isolates and of the studies of prokaryotic diversity in environmental samples based on phylogenetic analysis of 16S rRNA gene sequences.

Key words: ribosomal operons, 16S rRNA genes, multiple copies, phylogenetic analysis.

Comparative analysis of ribosomal genes, or *ribosomal phylogenetics*, is currently widely used to infer evolutionary relationships between species, especially prokaryotic. The underlying principle of ribosomal phylogenetics postulates that ribosomal genes can serve as a molecular clock ideally suited for reconstructions of evolutionary history and for building a universal tree of living organisms [1, 2]. The choice of ribosomal genes as a molecular clock was strongly influenced by the fact that, their end products being RNAs, the selection pressure targets the primary sequences of these genes rather than those of the encoded proteins. Additional advantages of using ribosomal genes as phylogenetic markers are their universal distribution, high sequence conservation, functional stability, easiness of sequencing, and, finally, specific features of their primary and secondary structures due to the presence of stretches of different variability in the sequences.

Ribosomal phylogenetics provided microbiologists with a very convenient, universally applicable method of finding the place of any microorganism in a phylogenetic tree. It thus should come as no surprise that the comparison of sequences of the 16S rRNA genes is so widely applied to identifying and classifying prokaryotes that it has become a required part of studies in their systematics [3]. Tens of thousands of prokaryotic 16S rRNAs have been and continue to be sequenced and pooled in international public domain databases, easily accessible online. As a result, preliminary identification of an unknown microorganism is now more often done by automated sequencing of its 16S rRNA genes than by traditional microbiological, physiological, and biochemical diagnostic tools.

In addition, ribosomal phylogenetics has become a staple tool of *molecular ecology*, greatly expanding our knowledge of prokaryote diversity in a variety of natural ecosystems. Unlike traditional microbiological techniques, the molecular ecology approach does not require isolation of the species of an ecosystem as pure cultures, relying instead on the estimates of the diversity of 16S rRNA sequences in the total DNA pool of the ecosystem under study.

It was postulated in the early applications of ribosomal phylogenetics in prokaryote systematics that each pure strain (isolated in the laboratory) or each individual organism in an ecosystem (not isolated) is represented by a unique 16S rRNA sequence (a phylotype). Further extensive studies produced a wealth of experimental data suggesting that this principle is far from universal. The reason for this deviation is the presence of multiple copies of ribosomal genes in the genome of a single organism, making possible polymorphism of these genes within the genome. This constitutes not only a technical difficulty in carrying out phylogenetic analysis but also provided some authors with a ground to doubt the very possibility of the wide application of ribosomal phylogenetic data in prokaryote systematics [4]. Here we describe and discuss the major aspects of this problem.

STRUCTURE OF PROKARYOTIC RIBOSOMAL OPERONS

Every cell of a living organism possesses ribosomes, the protein-synthesizing structures whose numbers vary from 20000 to 50000 depending on the intensity of protein synthesis. Ribosomes are made of several

0026-2617/03/7204-0389$25.00

molecules of rRNA bound to a number of proteins. Ribosomes of prokaryotes (termed 70S particles according to their sedimentation properties) consist of small (30S) and large (50S) subunits. *Escherichia coli* has typical prokaryotic ribosomes, which, due to the extensive available information on their structure and functions, serve as a gold standard for analysis of ribosomal genes from other organisms. Small 30S subunits of bacterial (e.g., *E. coli*) ribosomes contain a single 16S rRNA molecule 1542 nucleotides (nt) long and 21 proteins of different molecular weight (S1–S21). Large 50S subunits contain two rRNAs: large 23S rRNA (2904 nt long) and small 5S rRNA (120 nt long); these two are bound to 34 protein molecules (L1–L34).

The genes for rRNAs of both small (16S rRNA) and large (23S and 5S rRNA) subunits are usually fused into a single operon within the genome of a prokaryotic microorganism and are cotranscribed in the direction 16S→23S→5S. A single primary transcript, which later generates the three types of rRNA, ensures that they are produced in equal amounts. In bacteria, the rare instances of noncanonical arrangement of ribosomal genes are usually encountered within organisms (often parasitic) with small genomes and low copy numbers of rRNA genes (1–2 per genome). For example, in the *Mycoplasma hyopneumoniae* genome, the 5S rRNA gene is separated from the fused 16S and 23S genes [5], whereas in *Mycoplasma gallisepticum* [6] and several *Rickettsia* species [7] the 16S rRNA gene is separated from the fused 23S and 5S genes.

In some prokaryotic species ribosomal genes were shown not only to be separated in the genome, but also to be transcribed independently. The rRNA genes for small (16S) and large (23S, 5S) ribosomal subunits of the hyperthermophilic bacterium *Thermus thermophilus* [8, 9] or of the parasitic bacterium *Wolbachia pipientis* [10] may serve as examples. The most unusual arrangement of ribosomal genes is found in the genome of the thermophilic archaeon *Termoplasma acidophilum*, where all rRNA genes—16S, 23S and 5S—are found in single separated copies and are independently transcribed [11, 12].

In addition, when ribosomal operons are found in several copies in the genome of an organism, some of the copies may be truncated. For example, the genomes of the actinomycetes *Thermobispora bispora* and *Thermomonospora chromogena* carry operons lacking one (5S) or even both (5S and 23S) rRNAs of the large ribosomal subunit [13, 14].

Aside from the three RNA-coding sequences, ribosomal operons contain intergenic spacers of varying lengths and one or more copies of tRNA genes. The spacers may lie before, between, or after the rRNA coding sequences, whereas the tRNA genes are usually found within the internal or 3'-terminal spacer. Transcription of a ribosomal operon produces an RNA molecule about 5000 nt long. Such transcript is to be processed to become fully functional. Before or during the processing, certain residues in the sequences of spacers, rRNA, and tRNA are specifically modified. Endonucleolytic cleavage of the primary transcript releases mature rRNA fragments, which then undergo additional modification, fold into a defined three-dimensional structure, and bind the set of ribosomal proteins to assemble into mature ribosomal subunits.

COPY NUMBERS OF RIBOSOMAL OPERONS IN PROKARYOTIC GENOMES

Unlike protein-coding genes, in which each transcript produces a number of molecules of the end product in multiple rounds of translation, the transcription of an rRNA gene produces only a single product molecule, the RNA transcript. This may be the reason why the high demand for rRNA synthesis must be met by multiple copies of the rRNA-coding genes; their copy numbers in eukaryotes may reach several thousand. Prokaryotic ribosomal operons may also be present in several copies, being an exception to the rule that most of the genes in prokaryotes are unique, i.e., are present as a single copy in the genome. The copy number of the ribosomal genes in prokaryotes is much less than in eukaryotes, averaging three or four per genome [15].

The copy numbers of ribosomal operons in a genome were experimentally determined for several groups of prokaryotes [15]. At present, a constantly expanding database containing data on the number and structure of ribosomal operons in prokaryotic genomes is available at the web site http://rrndb.cme.msu.edu/rrndb/servlet/controller. Nearly half of the prokaryotic species carry one or two copies of the rRNA genes in their genomes. Most of the organisms containing multiple ribosomal operons carry three to eight copies. The species with higher copy numbers are encountered much less frequently: *Clostridium perfringens* and *Streptomyces nodosus* have ten copies [16, 17], and *Bacillus cereus* has 9–12 copies [18]. The highest known copy number of 16S rRNA genes is 15 in *Clostridium paradoxum* [19].

In some cases, e.g., in *B. cereus*, variations in the number of ribosomal operons (9 or 12) in different strains have been directly related to their genome size [18]. In general, however, the genome size and the ribosomal genes copy number correlate only slightly [15, 20]. Both these parameters, however, were shown to be stable for large phylogenetic groups [15]. For example, genome sizes of most of the *Archaea* are similar, and they all possess one or two copies of the ribosomal genes, although exceptions are known (*Methanococcus vannieli* carries four copies of the rRNA genes in its genome [21]).

Comparison of closely related organisms reveals that they usually have similar numbers of ribosomal genes [15], making it possible to extrapolate the data from one species to a related species for which the number of copies of the rRNA genes is not known. Such an extrapolation is often helpful, for example,

when estimating the number of organisms in environmental samples using competitive PCR. This rule is not without exception either: the genome of *Aquifex pyrophilus* contains six copies of 16S rRNA [22], whereas all closely related species have no more than two. A similar situation is observed with different species of extremely halophilic archaea, in which the copy number of the ribosomal genes varies between one and four [23].

Finally, in the genomes of some species, the number of multiple copies of the ribosomal operons is not invariable. Intrachromosomal recombination in *Bacillus subtilis* may lead to spontaneous deletions and the disappearance of one of its ten copies of the rRNA operon [24]. A deletion of the same nature was shown to affect one of the several ribosomal operons in *E. coli* [25, 26]. Such deletions may account for strain variability in the number of ribosomal operons in the genome, which was detected in several species, including *B. subtilis* [27], *B. cereus* [18], *Paenibacillus polymyxa* [28], and *Streptococcus dysgalactiae* [29].

PHYSIOLOGICAL ROLE OF MULTIPLE COPY NUMBER OF RIBOSOMAL OPERONS

It is widely believed that the existence of multiple copies or ribosomal genes in prokaryotes is needed to maintain the high rate of rRNA synthesis and thus the large number of ribosomes required for fast cell growth and division [30]. In particular, direct estimates of the transcription rate of the rRNA operon show that one operon is insufficient to produce the number of ribosomes required for the highest growth rate observed in *E. coli* [31]. This view is supported by the existence of a general correlation between the growth rate of prokaryotic species and the number of ribosomal operons in their genomes [15]. For example, slowly dividing bacteria, both free-living (such as a strain of marine oligotrophic bacterium *Sphingomonas* [32]) and parasitic (such as the obligate intracellular parasite *Rickettsia prowazekii* [33]) possess only one ribosomal operon in their genomes. Most of the slowly growing species of the genus *Mycobacterium* also have only one ribosomal operon [15], whereas *M. fortuitum* and related species grow faster and have two operons [34]. It has been suggested that slowly growing species branched off the main line of evolution of mycobacteria with a deletion of one of their two ribosomal operons [35]. The tendency to lose one of the two ribosomal operons may be connected with the mechanism of drug resistance found in some mycobacterial species and based on a mutation in the single 16S rRNA gene [36].

Nevertheless, there exist organisms, such as many archaea, in which a single ribosomal operon provides for a high growth rate [37]. In addition, a spontaneous deletion or inactivation of one of several ribosomal operons hardly affects the growth rate of many microorganisms [24–26]. In a series of *E. coli* strains specifically designed to have ribosomal operons inactivated, even the strain with only one of the seven operons remaining functional produced nearly half the amount of rRNA as compared with the wild type strain [38]. Thus, the functional role of the multiple copies of ribosomal operons is likely not limited to a mere support of high cell growth rates but may be an adaptive mechanism. Support for this idea comes from the observation that the genomes of symbiotic or parasitic species existing in very stable environments contain a small number of ribosomal operons [39–41]. The predominant occurrence of single-copy ribosomal operons in archaea can also be explained by their highly specialized nature allowing them to exist in very narrow ecological niches. Additional evidence for the adaptive role of multiple rRNA operons was provided by the experiments with *E. coli* that showed that only five of its seven ribosomal operons are sufficient to support the optimal growth rate in rich media, but all seven are required for quick adaptation to new nutrient sources and temperature conditions [42]. At the same time, individual copies of ribosomal genes in the common laboratory strains of *Paenibacillus polymyxa* may be expressed differently depending on the growth conditions and the age of the culture [28]. A recent study of soil bacteria representing different phylogenetic groups has shown that the number of ribosomal operons in the genomes of species in a bacterial community correlates with the rate of their adaptation to a change in the nutrient source [20]. After a new nutritional substrate had been introduced into soil, the community composition changed, as was evidenced by an increase in the average number of copies of ribosomal operons in the isolated strains from one to four to five to seven. It has been suggested that such a mechanism is common for the ecological strategies of environmental bacterial populations, responding to a change in the nutrient supply by changes in the composition of the population.

SEQUENCE CONSERVATION BETWEEN MULTIPLE COPIES OF RIBOSOMAL GENES

The discovery of the wide occurrence of multiple-copy ribosomal operons in the genomes of prokaryotic organisms inevitably raised the question of how highly conserved the different copies are. The putative differences between them may be found at different scales of organization: (i) the sequences of the 16S, 23S, and 5S rRNA genes and the spacer regions; (ii) the structure of the operons; (iii) the location of the operons in the genome map; and (iv) the functioning of the operons. The phylogenetic analysis is primarily concerned with the first level of organization, since the reliability of this technique may be influenced by the heterogeneity of the sequences of the most widely used molecular marker, the 16S rRNA gene, and, to a lesser extent, of the 5S and 23S rRNA genes and the spacers.

The comparison of sequences of the 16S rRNA genes in the GenBank database revealed a significant amount of intraspecies and intrastrain divergence that is not only attributable to experimental errors [43]. As one

of the possible reasons for this divergence, the authors proposed the heterogeneity of several copies of the 16S rRNA gene sequences within the genome of an individual organism. The reported cases of a reliably established difference between individual copies of these genes may be rather arbitrarily pooled into two classes: (i) *microheterogeneity*, when the differences are minimal (less than 2%), and (ii) *macroheterogeneity*, when the differences are quite significant (more than 2%).

According to this classification, most prokaryotic species from different phylogenetic groups fall into the former class. The 16S rRNA gene copies were found to be identical in several streptococcal strains [44], a strain of *Renibacterium salmoninarum* [45], a strain of *Bifidobacterium bifidum* [46], several strains of *Mycoplasma hominis* [47], and a strain of *Rhodobacter sphaeroides* [48]. Little difference was found between several copies of the 16S rRNA gene in a strain of *Prevotella bryantii* [49], several cultured strains and environmental isolates of the genus *Vibrio* [50], a strain of *E. coli* [51], several strains of the genus *Mycobacterium* [52, 53], a strain of *Phytoplasma* [54], several strains of *Mycoplasma* [55, 56], a strain of *Streptomyces coelicolor* [57], a strain of *Bacillus subtilis* [58], a strain of *Streptomyces nodosus* [17], and a strain of *Clostridium perfringens* [59].

Even the microheterogeneity of the 16S rRNA gene copies, however, may make direct sequencing of PCR products impossible. Such a situation occurred in a strain of *Mycobacterium celatum* during the sequencing of this gene, two copies of which only differed in having an insertion of a single nucleotide in one position and a few more nucleotide substitutions [53]. In another slowly growing strain of *Mycobacterium*, two copies of the gene diverged by substitutions in 18 positions [52]. The microheterogeneity of multiple copies of the 16S rRNA gene in genomes of several streptomycete strains was directly evident from the sequencing gels, with more than one nucleotide found at certain positions [60].

The microheterogeneity of copies of the 16S rRNA gene may also be detected by gradient gel electrophoresis. This technique allows separation of nucleotide fragments with different sequences based on their G+C content. For example, temperature gradient gel electrophoresis (TGGE) detected several characteristic bands during separation of PCR products of a 347-base-pair-long variable region of the 16S rRNA gene from several strains of *P. polymyxa*. The pattern of the bands was strain-specific, suggesting the presence of several heterogeneous copies of this gene in each of these strains [28]. Subcloning of the individual copies of the 16S rRNA gene from the type strain of this species revealed differences in one to eight nucleotides in a total of ten positions.

In most cases, the sequence differences in microheterogeneous copies of the 16S rRNA gene do not occur at random but cluster within the variable regions of the gene, although rare occurrences of differences in the conservative regions are also known [52, 55, 56, 61]. The nucleotide substitutions may be both single and compensatory, but they do not change the secondary structure of the respective rRNA regions.

Confirmed cases of macroheterogeneity of the 16S rRNA genes in prokaryotes are much less frequent. The most thoroughly studied and proven macroheterogeneous system is the two copies of the 16S rRNA in the genome of the halophilic archaeon *Haloarcula marismortui*, which possesses two spatially separated ribosomal operons, both of which are transcribed [62]. These operons are significantly different in several respects: they have different sequences of the 16S, 23S, and 5S genes; different distribution of tRNA genes within the operon; and their transcripts are processed differently. However, the rRNA products of both operons are nearly equally represented in the cellular ribosomal pool [63, 64]. The sequence of the 16S rRNA gene of this species was investigated in most detail [63]. Both copies of this gene were of the same length and did not carry deletions or insertions but differed by nucleotide substitutions in 74 positions, or about 5% of the total number of nucleotides. These differences were not evenly distributed along the gene but were clustered in its three variable regions. To exclude the possibility that this macroheterogeneity reflected contamination of the culture, this study was followed with experiments on a culture grown from a single cell; the results of the earlier study were fully confirmed [65]. In the same study, the expression of both copies in the same cell was confirmed by in situ fluorescent hybridization. The study of other species of the genus *Haloarcula* revealed that *H. argentinensis* contains only one copy of the 16S rRNA gene in its genome, whereas *H. vallismortis*, *H. quadrata*, and "*H. sinaiiensis*" have two heterogeneous copies [66, 67].

Two other cases of macroheterogeneity of the 16S rRNA genes are provided by two species of actinomycetes. Four ribosomal operons, two of which were complete and the other two were not, were found in the genome of *Thermobispora bispora*. The 23S rRNA genes in all these operons were nearly identical, whereas the 16S rRNA genes belonged to two types, each of which contained nearly identical sequences, but between the types the sequences differed in 98 positions (6.4%) and in six small insertions/deletions mostly clustered in the variable regions of the gene. Both types of 16S rRNA genes were normally expressed [13]. Six ribosomal operons, including two incomplete ones, were discovered in the genome of another actinomycete, *Thermomonospora chromogena*. Five of these operons belonged to one type, with the 16S and 23S genes within the type nearly identical in sequence, whereas another type consisted of one complete operon. The 16S and 23S rRNA genes of these two types differed by 6 and 10%, respectively, and were normally expressed in the cell [14].

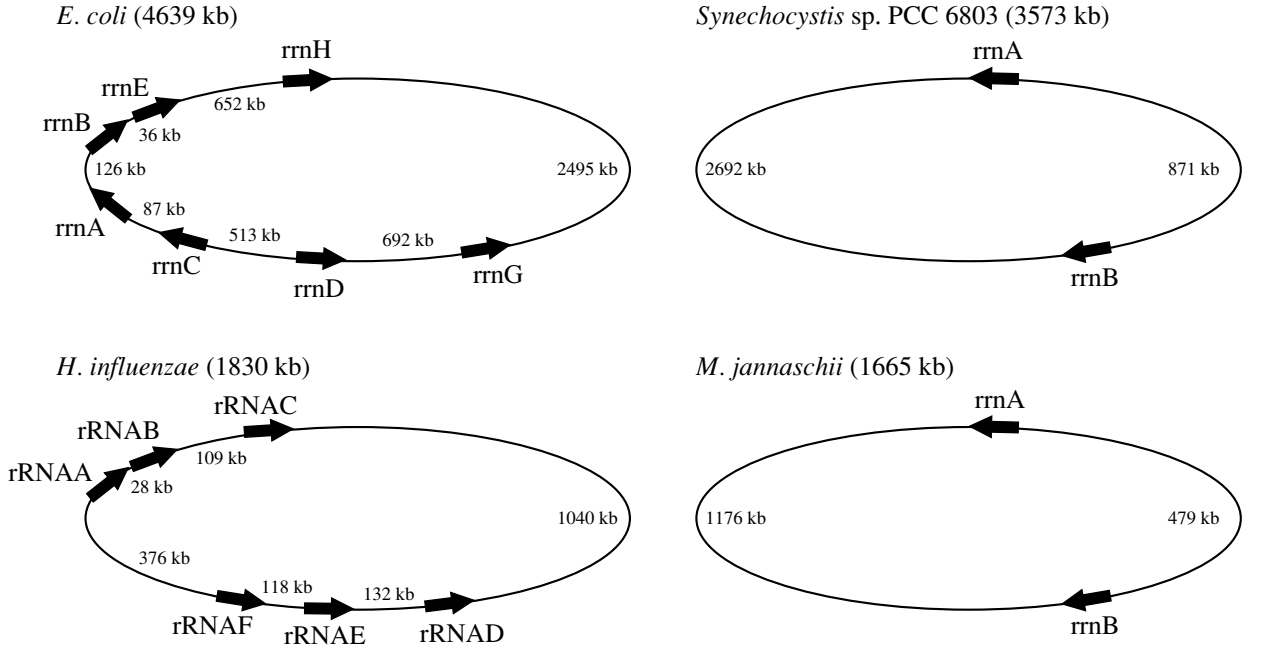

Fig. 1. Positions of ribosomal genes in genome maps of several prokaryotic species [75]. The full genome sizes (kb) are given in parentheses. Arrows indicate locations of multiple copies of ribosomal genes; the distances between the copies (kb) are shown.

Intervening sequences (IVS) are a special case of macroheterogeneity of 16S rRNA genes. Various kinds of these short inserts are found in the 16S rRNA gene sequences of both archaeal and bacterial genomes [68]. Ribosomal genes of several bacteria contain a special type of IVS excised during processing, which yields a fragmented rRNA product. The presence of IVS of this type in the 16S rRNA genes was established for *Caedibacter caryophila* [69], *Campylobacter helveticus* [70], and *Helicobacter canis* [71], possibly contributing to the heterogeneity of different copies of the gene. One to four IVS of different length were found in the 16S rRNA genes of various thermophilic species of the genera *Clostridium*, *Thermoanaerobacter*, and *Thermoanaerobacterium* [72]. The genome of one such species, *Clostridium paradoxum*, contains 15 copies of the 16S rRNA gene, which are macroheterogeneous (up to 3.2%) due mostly to the presence of IVS of different length and structure in the hypervariable 5'-terminal region of the gene and single nucleotide substitutions in the rest of the gene [19]. Northern blot analysis revealed that the 16S rRNA product lacks the IVS-coded sequences, but their absence is due to the expression of the two copies of the gene lacking the IVS rather than to IVS excision during processing. It is believed that the IVS-containing copies of the gene are silent or serve as a backup, but, unlike in *E. coli* [26], a shift towards unfavorable growth conditions (temperature, pH, salt concentration, age of the culture) did not turn on the expression of the additional gene copies in this microorganism.

Another type of IVS was initially discovered in the thermophilic bacterium *Desulfotomaculum australicum* [73] and later in other species of this genus. Unlike the case discussed above, this IVS, an insert in a hypervariable 5'-terminal region of the 16S rRNA gene, is expressed. In the final rRNA product, the IVS forms a helix which is about 40 nucleotide pairs longer than the regular helix present in the 16S rRNA of other prokaryotes. We observed significant macroheterogeneity of two copies of the 16S rRNA gene, caused by the presence of similar IVS, in another species of the same genus, *D. kuznetsovii* [74]. Comparison of the full sequences of the two copies of the 16S rRNA gene in this species revealed a very high degree of their heterogeneity (8.3%). The differences between the copies occurred mostly in long (105–118 nt) inserts in the variable 5'- and 3'-terminal regions of the 16S rRNA genes; several minor nucleotide substitutions were also found in other variable parts of the gene. Although the expression of 16S rRNA genes in *D. kuznetsovii* was not specifically addressed, it is reasonable to believe that the rRNA contains the IVS sequences, because the IVS formed a secondary structure similar to that found in *D. australicum*.

MULTIPLE COPIES OF RIBOSOMAL OPERONS IN EVOLUTION

Amplification of a single-copy ancestral DNA sequence is a well-known way for families of repetitive sequences in a genome of a single organism to appear. The amplified repetitive sequences may be either located in tandem or dispersed over the genome.

A comparison of full genomes for six archaeal and 19 bacterial species revealed that their ribosomal operons are usually dispersed (Fig. 1) [75]. At the same time, clusters of tandem ribosomal operons are characteristic for some bacilli [76], whereas two ribosomal operons of *Chlamydia trachomatis* are not only arranged in tandem but also cotranscribed [77].

The appearance of new gene copies by amplification may have several possible outcomes: the copies may serve the same function or one copy may acquire a new function, turn into a regulatory element, or become a nonfunctional pseudogene. In each case, the features of the evolution of individual copies will depend on their functions. The most rapid evolution is undergone by the sequences of nonfunctional pseudogenes, experiencing no selection pressure. For example, the chloroplast genome of a brown alga was shown to contain a pseudogene that significantly diverged in its sequence from the two functional copies of the 16S rRNA gene [78]. Changes in the functions of ribosomal operons are presently not known, although their amplified copies in the same organism could in principle evolve independently at a rate comparable with the rate of divergence of homologous genes in different organisms. However, as discussed in the previous section, in most cases when multiple copies of ribosomal genes occur within a single genome, their sequences are identical or very similar. Further support of this view comes from the comparison of copies of ribosomal operons in the fully sequenced prokaryotic genomes, as demonstrated by the phylogenetic tree of their 16S rRNA genes (Fig. 2). To explain the situation when the sequences of ribosomal genes are maintained quite homogeneous while the rate of divergence between homologous genes of different organisms is normal, the concept of *concert evolution* was proposed [75, 79]. The concert evolution of ribosomal genes is believed to result from a high selection pressure, acting at the level of rRNA sequences, towards the conservation of the important functional role of these genes in the overall structure of the complex cellular protein-synthesizing machinery. For example, a study of chloroplast evolution in green algae that have lost their photosynthetic capacity demonstrated a relationship between functional relaxation and changes in the features of the mutation process of the 16S rRNA genes [80].

Despite the lack of definite conclusions on the exact nature of molecular mechanisms responsible for the homogenization of nucleotide sequences of multiple ribosomal genes and thus favoring their concert evolution in prokaryotes, gene conversion processes seem to be best suited to this role [75, 81]. Recombination between different ribosomal operons in the same chromosome was experimentally determined to maintain the homogeneity of their sequences in *E. coli* [82], *Salmonella typhi* [83], and *Vibrio cholerae* [84].

At the same time, the cases of significant heterogeneity or ribosomal genes within a single microbial genome, discussed in the previous section, testify to the existence of other mechanisms of ribosomal gene evolution. Two alternatives for the origins of heterogeneity are possible: mutations arising due to DNA polymerase errors during DNA replication and horizontal transfer of genetic material by conjugative plasmids, followed by recombination. The former mechanism would act against concert evolution, promoting the amplified genes in families to diverge. For example, the difference in the rate of mutation accumulation was proposed as a reason for the microheterogeneity (1.2%) between two copies of the 16S rRNA gene in the genome of a strain of *Mycobacterium* [52]. This was explained by the level of expression of one of these copies (the one evolving faster and thus less stable and less active), which was tenfold lower than for the other one. The second mechanism can also impede concert evolution of amplified genes by transfer and recombination of foreign genetic material into them. For example, it was suggested that both mechanisms act to generate heterogeneity of the sequences of variable regions of the multiple 16S rRNA genes in different streptomycete strains [60]. It was recently confirmed experimentally for a large number of actinomycete strains that small fragments of ribosomal genes may be actively involved in horizontal transfer [85]. The results of such transfer may become fixed in the recipient gene and later spread into other copies of this gene (if any) if they offer selective advantages such as drug resistance [86].

In addition, the second mechanism possibly operates to generate the heterogeneity of amplified genes due to the presence of IVS. It has been extensively argued that IVS are mobile elements and their dissemination by horizontal transfer is more likely than the vertical descent from a common ancestor. This model is supported by comparative analysis of phylogenetic trees of 23S rRNA IVS from the species of the genus *Leptospira* [87], by high similarity of IVS in the organisms from distantly related genera *Escherichia* and *Yersinia* [88], and significant deviation of the G+C content in IVS sequences of *Campylobacter jejuni* from the total G+C content in its genome [89].

At the same time, horizontal (interspecies) transfer of complete genes may provide an alternative to amplification in the formation of repetitive gene families consisting of orthologous genes. For instance, two alternative scenarios were suggested to explain a significant heterogeneity of ribosomal operons in the genome of *Haloarcula marismortui* [63–65]. The first one assumes that both operons are fully paralogous and appeared as a result of amplification; thus, their heterogeneity is explained by a breach in the mechanism of concert evolution. Accumulation of differences in the sequence, structure, and processing of the operons followed, but the differences did not disable the functionality of the operons. However, some of the features of the differences between the operons prompted the authors to lend more credibility to the other scenario, in which these operons are orthologous and coexist in the

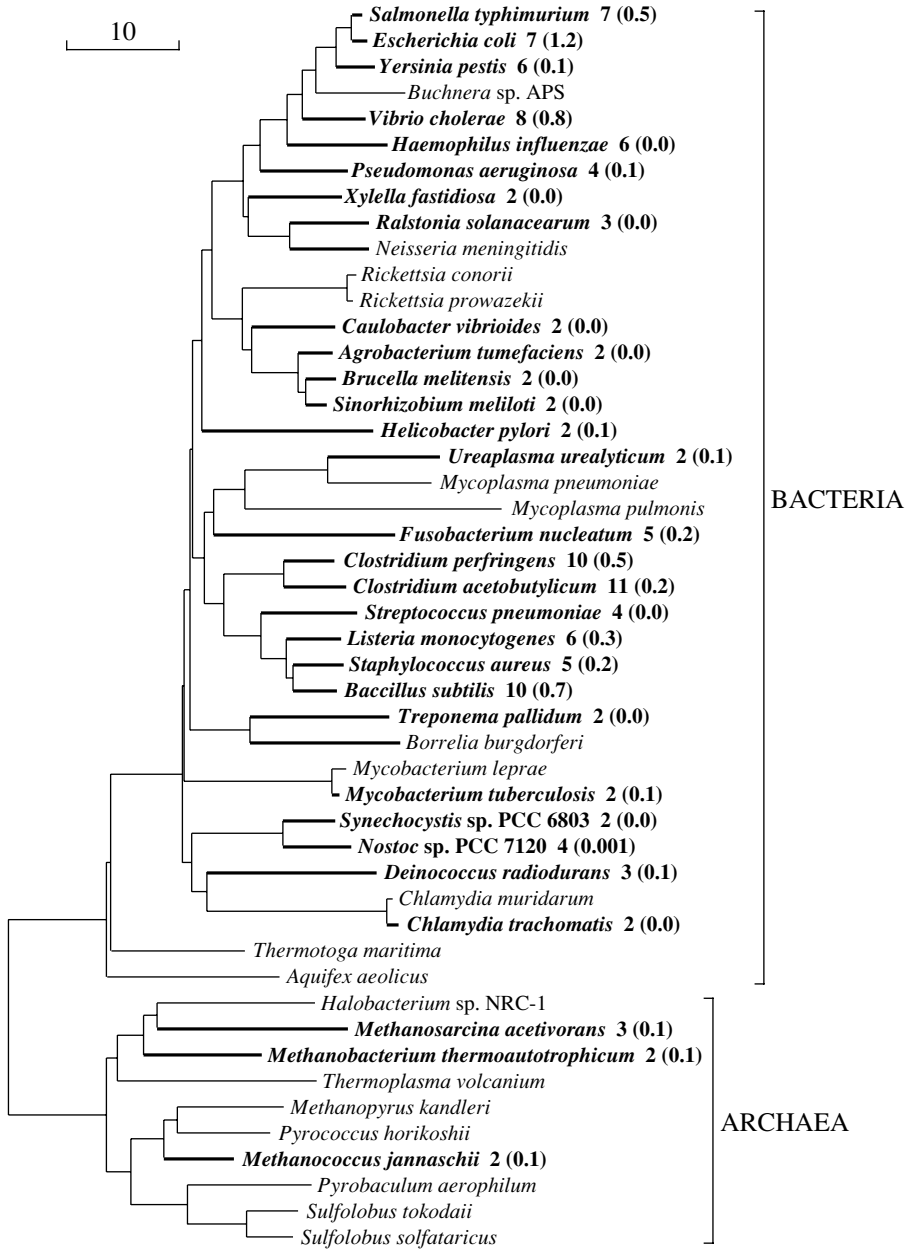

Fig. 2. Phylogenetic tree inferred from all 16S rRNA genes from fully sequenced genomes of 38 bacteria and 10 archaea. The sequences were aligned using CLUSTAL W software, and the tree was constructed by the neighbor-joining method implemented in the TREECONW software package [97]. Branch lengths are proportional to the number of nucleotide substitutions per 100 nucleotides (% homology). Names of species possessing more than one ribosomal operon are given in boldface. The numbers of 16S rRNA genes in their genomes are indicated, and the degree of heterogeneity (%) is given in parentheses.

same genome as a result of a horizontal gene transfer event. In this situation, some parts of these operons might have become more homogeneous later, during their concert evolution.

Similar ideas were put forth to explain significant heterogeneity between two types of 16S rRNA genes in the *Thermobispora bispora* genome [13]. To validate the possibility of a horizontal gene transfer, an extensive search for a possible donor of an orthologous gene was undertaken, but to no avail. In spite of that, the authors favor the appearance of two types of 16S rRNA genes by horizontal gene transfer into this organism, but divergent evolution still cannot be excluded.

The first experimental proof of ribosomal operon heterogeneity due to horizontal gene transfer is the case of *Thermomonospora chromogena* [14]. Comparative analysis revealed a high degree of homology between the 16S and 23S rRNA gene sequences of the *T. chromogena* rrnB operon and the *Thermobispora bispora* rrnA operon (Fig. 3). The authors believe that *T. chro-*

mogena acquired the *Thermobispora bispora* rrnB operon, or an operon of another organism closely related to both these species, by horizontal transfer. The transferred operon then underwent concert evolution, becoming more similar to other ribosomal operons of the recipient genome.

It should be noted that the role of horizontal interspecies gene transfer is one of the most hotly debated topics in molecular evolution. The opinion prevails that, unlike protein-coding, or operational genes, the informational genes of prokaryotes, including ribosomal genes, are transferred very infrequently in the evolution [2]. According to the proposed *complexity hypothesis*, this low frequency of transfer is due to the fact that informational genes usually function in gene complexes, whereas operational genes do not [90]. At the same time, the case of *Thermomonospora chromogena* proves that a naturally transferred foreign ribosomal operon can successfully work with the protein-synthesizing machinery of the recipient. In addition, it has recently been shown that a mutant strain of *E. coli* with all ribosomal operons inactivated could be made viable by introducting foreign ribosomal operons and incorporating of their products into hybrid ribosomes [38]. These results suggest that coevolution of ribosomal genes with a multitude of genes of the protein synthesis system can hamper, but not totally exclude, the possibility of their horizontal transfer.

POLYMORPHISM OF RIBOSOMAL GENES AS A FACTOR IN PHYLOGENETIC ANALYSIS

There are several possible ways in which the presence of multiple, possibly polymorphous copies or ribosomal genes may influence phylogenetic analysis based on their comparison.

To begin with, technical difficulties could appear, hindering the analysis. At present, the most widely used method of sequencing of ribosomal (especially 16S rRNA) genes includes PCR amplification with specially designed primers as its first step. The heterogeneity of individual gene copies may generate chimeric artifacts [91, 92], thus impeding taxonomic identification of the organism by seemingly detecting more copies of the rRNA genes than really exist. For example, 30 and 35 PCR cycles were shown to yield 11,1 and 15,5% of chimeric molecules, respectively, as a result of coamplification of two types of 16S rRNA genes in *Thermobispora bispora* [13]. Of course, the proportion of chimeric molecules produced during amplification depends on the number of gene copies in a genome and the level of heterogeneity between them. A good illustration of this situation is given by *Desulfotomaculum luciae*, described as a separate species. The 16S rRNA gene of this organism contains IVS in its 5'- and 3'-terminal variable regions [93]. Each of these IVS was later shown to be homologous to one of the IVS from the respective regions of two heterogeneous copies of the 16S rRNA gene of *D. kuznetsovii* [74]. Therefore it will require specially designed experiments to clarify the question of whether the 16S rRNA gene of *D. luciae* is indeed a unique phylotype of a new species or represents a chimeric artifact that arose in one of the *D. kuznetsovii* strains. To avoid erroneous interpretation of artifactual chimeric sequences, it is recommended that all newly sequenced rRNA genes be checked against dedicated databases using special tools such as CHECK_CHIMERA [http://rdp.cme.msu.edu/html/], although, as the example of *D. luciae* shows, such precautions do not fully exclude ambiguous interpretations of the results. In this case, the ambiguity can only be ruled out by sequencing and comparing all copies of the gene in the genome under study. Another possible technical difficulty caused by the presence of IVS is the possible exclusion of IVS-containing sequences during PCR because of their increased length, not corresponding to the standard size of a 16S rRNA gene [92].

As mentioned above, if the PCR product of a 16S rRNA gene is sequenced immediately, the possible heterogeneity of the multiple copies can be directly observed as sequence ambiguity. If, however, the amplified products are subcloned and only then sequenced, the probability of missing one or more of the multiple copies present in the genome is very high, especially if just one or two clones are sequenced. This can also occur if the number of PCR cycles is insufficient for the amplification of all present gene copies; in this case, one of several copies is amplified preferentially because of the stochastic nature of initiation by DNA polymerase. If all copies are identical, the missing sequences will not influence the conclusions of a comparative phylogenetic analysis, but if a nonidentical copy is missed, the results of the phylogenetic analysis on a subset of copies could be erroneous.

The next aspect of the problem is related to the construction of phylogenetic trees and their usefulness for drawing taxonomic conclusions about the microbial isolates under study. Here the influence of the multiple-copy nature of the main phylogenetic marker, the 16S rRNA gene, depends on the level of polymorphism between individual copies. If they are identical or display a very low microheterogeneity (0.1–0.5%, comparable to the experimental error), the presence of the multiple copies does not influence the comparison of even closely related organisms. An increase in heterogeneity of two copies of the 16S rRNA gene to 1.2%, as in the strain of *Mycobacterium*, can already lead to a change in the position of these copies in a phylogenetic tree [52]. The probability of a change in the tree topology, taking into account the representation of a respective strain in the general population, is about 2% per strain [60].

These rather small changes in the phylogenetic tree topology due to the microheterogeneity of multiple copies of the 16S rRNA gene are significant only for the positioning of closely related organisms in the tree. It is well known that the predictive value of analysis of the 16S rRNA gene decreases rapidly in more closely

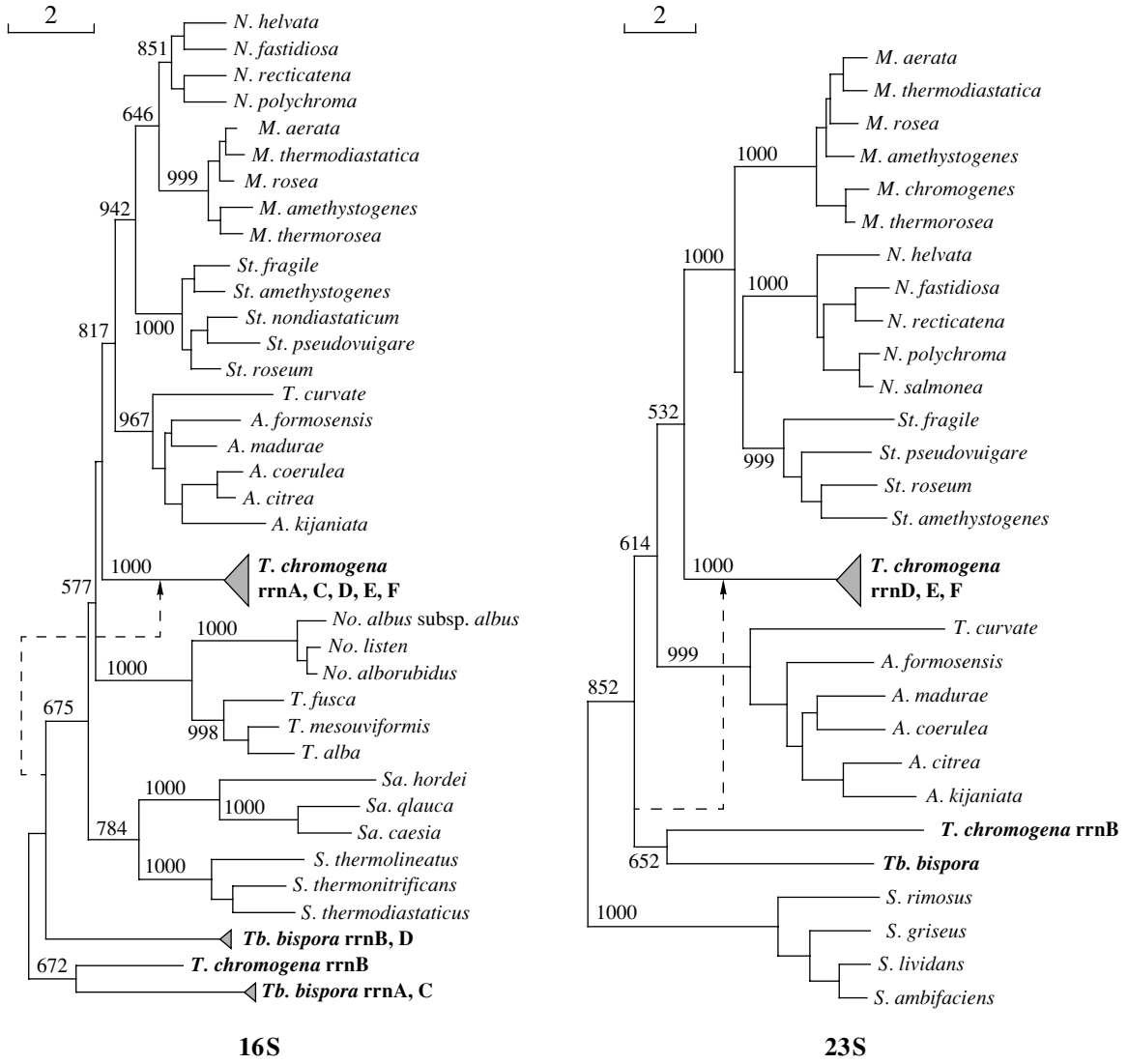

Fig. 3. Phylogenetic trees of actinomycetes inferred from the sequences of their 16S and 23S rRNA genes [14]. Multiple copies of *Thermomonospora chromogena* and *Thermobispora bispora* ribosomal genes are indicated in boldface. The genera are abbreviated as follows: *N*, *Nonomuria*; *M*, *Microbispora*; *St*, *Streptosporangium*; *A*, *Actinomadura*; *T*, *Thermomonospora*; *No*, *Nocardiopsis*; *Sa*, *Saccharopolyspora*; *S*, *Streptomyces*; *Tb*, *Thermobispora*. Branch lengths are proportional to the number of nucleotide substitutions per 100 nucleotides. The numbers at branching points indicate branching order confidence according to bootstrap analysis of 1000 trees. The dashed line shows changed positions of the respective branches when *Streptomyces* and *Saccharopolyspora* species were excluded from the analysis.

related organisms [94]. This method cannot be applied to detect the differences at the intraspecies, and sometimes even the interspecies, level, since many prokaryotic species may significantly differ phenotypically and at the level of DNA–DNA hybridization but possess nearly identical 16S rRNA sequences [95, 96]. The divergence of 16S rRNA gene sequences between strains of a single species was determined at a level of 3% [96]; therefore, the level of divergence due to microheterogeneity of multiple copies of these genes (2%) is on a par with the intraspecies divergence and cannot significantly hamper the identification of a microbial strain.

At the same time, macroheterogeneity of copies of the 16S rRNA gene may exceed the intraspecies level, thus greatly increasing the probability of multiple gene copies significantly influencing the phylogenetic tree topology and the outcome of the analysis. This can be illustrated by phylogenetic trees constructed for two prokaryotic genera, the bacterial genus *Desulfotomaculum* (Fig. 4a), and the archaeal genus *Haloarcula* (Fig. 4c), which include several organisms possessing heterogeneous copies of 16S rRNA genes in the genome. Consideration of either of these phylogenetic trees reveals that different copies of the genes from the same organism occupy notably different positions, pro-

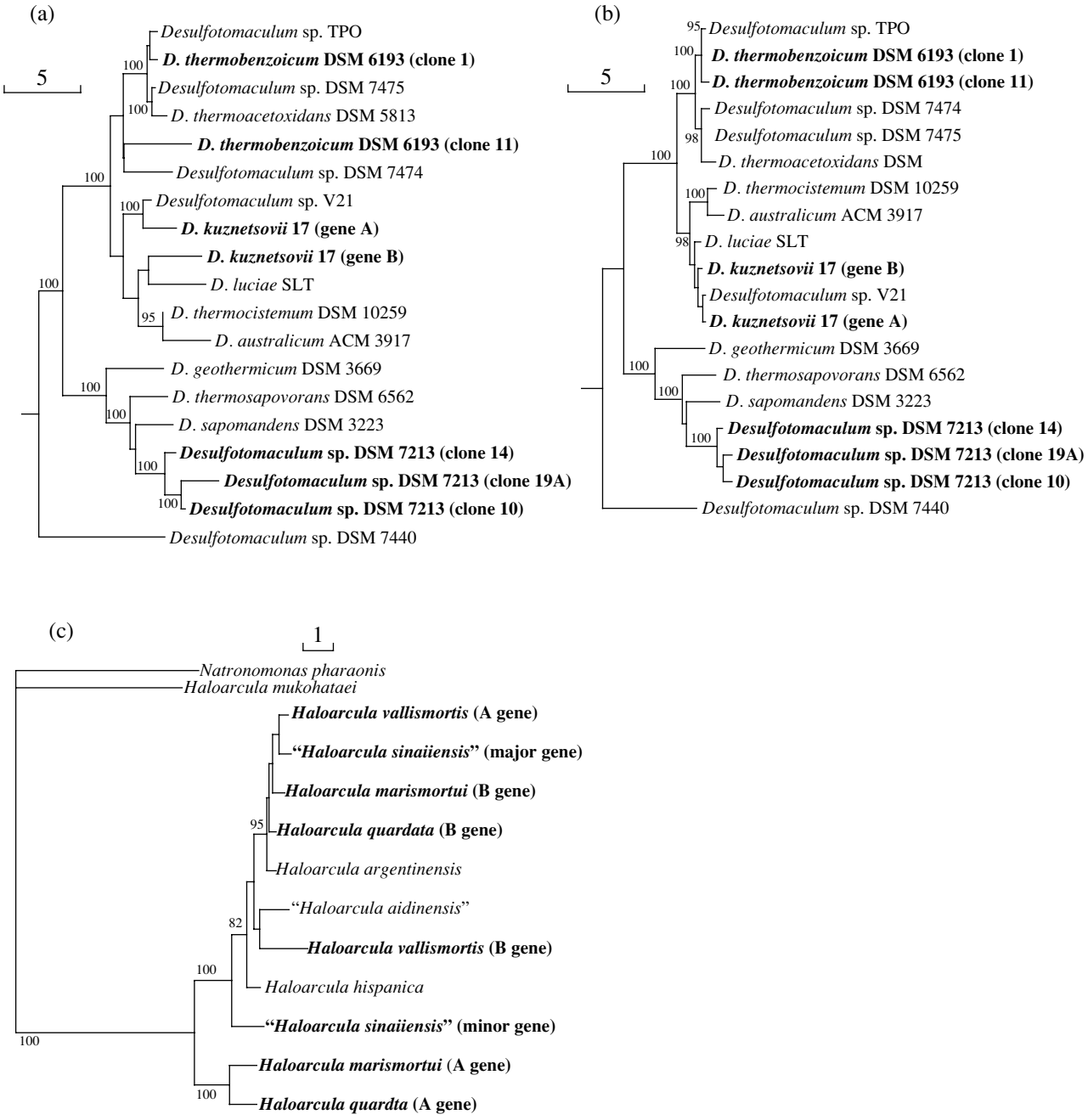

Fig. 4. Phylogenetic trees for prokaryotic genera for which the heterogeneity of 16S rRNA gene copies has been shown: (a) the tree for *Desulfotomaculum* inferred from the full sequences of the 16S rRNA genes; (b) the tree for *Desulfotomaculum* after the exclusion of variable 5'- and 3'-terminal regions of 16S rRNA gene from the analysis; (c) the tree for *Haloarcula* [67]. Names of the species possessing more than one ribosomal operon are given in boldface. Branch lengths are proportional to the number of nucleotide substitutions per 100 nucleotides. The numbers at branching points indicate branching order confidence according to bootstrap analysis of 100 trees.

viding ample grounds for erroneous taxonomic conclusions if only one copy is analyzed for the technical reasons described above.

Limiting the analysis to exclude hypervariable regions of the 16S rRNA gene, especially those containing IVS, may minimize the possibility of distortion in the phylogenetic data. Fig. 4b demonstrates how the tree topology changes when this procedure is applied to comparative analysis of 16S rRNA genes in the genus *Desulfotomaculum*. After the correction, copies of this gene from a single organism were grouped much tighter in the tree, forming monophyletic clusters, and the level of differences between the sequences did not exceed 0.8%, well within the limits of intraspecies

deviation and close to the experimental margin of error. Exclusion of variable regions of 16S rRNA gene sequences that are not susceptible to unambiguous alignment is a common practice in comparative analysis of distantly related organisms. Therefore, the heterogeneity of multiple copies cannot significantly impede the reliability of phylogenetic analysis at the level of genera and higher taxonomic groups.

Another aspect of the problem is associated with the use of the results of the ribosomal phylogenetics approach in the molecular ecology studies of prokaryote diversity in natural environments. The presence of multiple, possibly heterogeneous copies in the genome of a single organism cast doubts on the "one phylotype–one organism" interpretation, since some of the different phylotypes may actually be associated with sequences of different copies of 16S rRNA genes within a single genome. The possibility must be admitted that groups of uncultured environmental clones with a high degree of similarity between their 16S rRNA genes, often found in molecular ecology studies, come from heterogeneous copies of these genes in a single strain rather than from different closely related strains [19]. The exclusion of the variable regions of 16S rRNA genes from the analysis may also decrease the number of such mistakes. In addition, molecular ecology studies of the rRNA genes may be fruitfully complemented by analysis of functional rRNAs, which helps to minimize the errors caused by the presence of IVS [92].

CONCLUSIONS

Although prokaryotic genomes often contain multiple copies of ribosomal genes, this is hardly an obstacle to comparing these genes in phylogenetic analyses. In most cases the sequences of different copies of ribosomal genes are completely or nearly identical. The reason for this is the existence of several genetic processes (still poorly understood in molecular terms) that promote homogenization of the sequences of multiple copies of ribosomal genes in a prokaryotic genome and allow these copies to evolve in concert. The phenomenon of concert evolution is believed to be governed by strong stabilizing selection pressure, which prevents changes in the structure of functional ribosomal genes. If, against the odds of concert evolution, heterogeneity appears in the sequences of different copies of ribosomal genes, in most cases the differences are limited and do not affect the results of phylogenetic analyses. At the same time, there exist several possible mechanisms that could generate significantly polymorphic copies of ribosomal genes within one genome and thus influence the outcome of phylogenetic analyses. This possibility must be considered when applying methods of ribosomal phylogenetics to identification of prokaryotic isolates and, especially, to estimates of prokaryotic diversity in environmental samples.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project no. 02-04-48112.

REFERENCES

1. Woese, C.R., Bacterial Evolution, *Microbial. Rev.*, 1987, vol. 51, pp. 221–271.
2. Woese, C.R., The Universal Ancestor, *Proc. Natl. Acad. Sci. USA*, 1998, vol. 95, pp. 6854–6859.
3. Stackebrandt, E., Frederiksen, W., Garrity, G.M., Grimont, P.A.D., Kampfer, P., Maiden, M.C.J., Nesme, X., Rossello-Mora, R., Swings, J., Truper, H.G., Vauterin, L., Ward, A.C., and Whitman, W.B., Report of the Ad Hoc Committee for the Re-Evaluation of the Species Definition in Bacteriology, *Int. J. Syst. Evol. Microbiol.*, 2002, vol. 52, pp. 1043–1047.
4. Golovlev, E.L., On the Old Problems of the New Bacterial Systematics, *Mikrobiologiya*, 1998, vol. 67, pp. 281–286.
5. Taschke, C. and Herrmann., R., Analysis of Transcription and Processing Signals of the 16S-23S rRNA Operon of *Mycoplasma hyopneumoniae*, *Mol. Gen. Genet.*, 1986, vol. 205, pp. 4705–4707.
6. Chen, X. and Finch, L.R., Novel Arrangement of rRNA Genes in *Mycoplasma gallisepticum*: Separation of the 16S Gene of One Set from the 23S and 5S Genes, *J. Bacteriol.*, 1989, vol. 171, pp. 2876–2878.
7. Andersson, S.G.E., Stothard, D.R., Fuerst, P., and Kurland, C.G., Molecular Phylogeny and Rearrangement of rRNA Genes in *Rickettsia* Species, *Mol. Biol. Evol.*, 1999, vol. 16, pp. 987–995.
8. Hartmann, R.K. and Erdmann, V.A., *Thermus thermophilus* 16S rRNA Is Transcribed from an Isolated Transcription Unit, *J. Bacteriol.*, 1989, vol. 171, pp. 2933–2941.
9. Hartmann, R.K., Ulbrich, N., and Erdmann, V.A., An Unusual rRNA Operon Constellation: In *Thermus thermophilus* HB8 the 23S/5S rRNA Operon Is a Separate Entity from the 16S rRNA Operon, *Biochimie*, 1987, vol. 69, pp. 1097–104.
10. Bensaadi-Merchermek, N., Salvado, J.C., Cagnon, C., Karama, S., and Mouches, C., Characterization of the Unlinked 16S rDNA and 23S-5S rRNA Operon of *Wolbachia pipientis*, a Prokaryotic Parasite of Insect Gonads, *Gene*, 1995, vol. 165, pp. 81–86.
11. Ree, H.K., Cao, K.M., Thurlow, D.L., and Zimmermann, R.A., The Structure and Organization of the 16S Ribosomal RNA Gene from the Archaebacterium *Thermoplasma acidophilum*, *Can. J. Microbiol.*, 1989, vol. 35, pp. 124–133.
12. Ree, H.K. and Zimmermann, R.A., Organization and Expression of the 16S, 23S and 5S Ribosomal RNA Genes from Archaebacterium *Thermoplasma acidophilum*, *Nucleic Acids Res.*, 1990, vol. 18, pp. 4471–4478.
13. Wang, Y., Zhang, Z., and Ramanan, N., The Actinomycete *Thermobispora bispora* Contains Two Distinct Types of Transcriptionally Active 16S rRNA Genes, *J. Bacteriol.*, 1997, vol. 179, pp. 3270–3276.
14. Yap, W.H., Zhang, Z., and Warg, Y., Distinct Types of rRNA Operons Exist in the Genome of the Actinomycete

Thermomonospora chromogena and Evidence for Horizontal Transfer of an Entire rRNA Operon, *J. Bacteriol.*, 1999, vol. 181, pp. 5201–5209.

15. Fogel, G.B., Collins, C.R., Li, J., and Brunk, C.F., Prokaryotic Genome Size and SSU rDNA Copy Number: Estimation of Microbial Relative Abundance from a Mixed Population, *Microb. Ecol.*, 1999, vol. 38, pp. 93–113.
16. Garnier, T., Canard, B., and Cole, S.T., Cloning, Mapping, and Molecular Characterization of the rRNA Operons of *Clostridium perfringens*, *J. Bacteriol.*, 1991, vol. 173, pp. 5431–5438.
17. Yap, W.H. and Wang, Y., Molecular Cloning and Comparative Sequence Analyses of rRNA Operons in *Streptomyces nodosus* ATCC 14899, *Gene*, 1999, vol. 232, pp. 77–85.
18. Johansen, T., Carlson, C.R., and Kolsto, A.R., Variable Numbers of rRNA Gene Operons in *Bacillus cereus* Strains, *FEMS Microbiol. Lett.*, 1996, vol. 136, pp. 325–328.
19. Rainey, F.A., Ward-Rainey, N.L., Janssen, P.H., Hippe, H., and Stackebrandt, E., *Clostridium paradoxum* Contains Multiple 16S rRNA Genes with Heterogeneous Intervening Sequences, *Microbiology* (UK), 1996, vol. 142, P. 2087-2095.
20. Klappenbach, J.A., Dunbar, J.M., and Schmidt, T.M., rRNA Operon Copy Number Reflects Ecological Strategies of Bacteria, *Appl. Environ. Microbiol.*, vol. 66, pp. 1328–1333.
21. Neumann, H., Gierl, A., Tu, J., Leibrock, J., Staiger, D., and Zillig, W., Organization of the Genes for Ribosomal RNA in Archaebacteria, *Mol. Gen. Genet.*, 1983, vol. 192, pp. 66–72.
22. Shao, Z., Mages, W., and Schmitt, R., A Physical Map of the Hyperthermophilic Bacterium *Aquifex pyrophilus* Chromosome, *J. Bacteriol.*, 1994, vol. 176, pp. 6776–6780.
23. Sanz, J.L., Marin, I., Ramirez, L., Abad, J.P., Smith, C.L., and Amils, R., Variable rRNA Gene Copies in Extreme Halobacteria, *Nucleic Acids Res.*, 1988, vol. 16, pp. 7827–7832.
24. Widom, R.L., Jarvis, E.D., LaFauci, G., and Runder, R., Instability of rRNA Operons in *Bacillus subtilis*, *J. Bacteriol.*, 1988, vol. 170, pp. 605–610.
25. Ellwood, M. and Nomura, M., Deletion of a Ribosomal Ribonucleic Acid Operon in *Escherichia coli*, *J. Bacteriol.*, 1980, vol. 143, pp. 1077–1080.
26. Condon, C., Philips, J., Fu, Z.-Y., Squires, C., and Squires, C.L., Comparison of the Expression of the Seven Ribosomal RNA Operons in *Escherichia coli*, *EMBO J.*, 1992, vol. 11, pp. 4175–4185.
27. Gottlieb, P., LaFauci, G., and Rudner, R., Alterations in the Number of rRNA Operons within the *Bacillus subtilis* Genome, *Gene*, 1985, vol. 33, pp. 259–268.
28. Nubel, U., Engelen, B., Felske, A., Snaidr, J., Wieshuber, A., Amann, R., Ludwig, W., and Backhaus, H., Sequence Heterogeneities of Genes Encoding 16S rRNAs in *Paenibacillus polymyxa* Detected by Temperature Gradient Gel Electrophoresis, *J. Bacteriol.*, 1996, vol. 178, pp. 5636–5643.
29. Bentleya, R.W. and Leigha, J.A., Determination of 16S Ribosomal RNA Gene Copy Number in *Streptococcus uberis, S. agalactiae, S. dysgalactiae* and *S. parauberis*, *FEMS Immunol. Med. Microbiol.*, 1995, vol. 12, pp. 1–7.
30. Bremer, H. and Dennis, P.P., Modulation of Chemical Composition and Other Parameters of the Cell Growth Rate, *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, Neidhardt, F.C. *et al.*, Eds., Washington: Am. Soc. Microbiol., 1987, pp. 1527–1542.
31. Bremer, H., Parameters Affecting the Synthesis of Ribosomes and RNA Polymerase in Bacteria, *J. Theor. Biol.*, 1975, vol. 53, pp. 115–124.
32. Fegatella, F., Lim, J., Kjelleberg, S., and Cavicchioli, R., Implications of rRNA Operon Copy Number and Ribosome Content in the Marine Oligotrophic Ultramicrobacterium *Sphingomonas* sp. Strain RB2256, *Appl. Environ. Microbiol.*, 1998, vol. 64, pp. 4433–4438.
33. Pang, H. and Winkler, H.H., Copy Number of the 16S rRNA Gene in *Rickettsia prowazekii*, *J. Bacteriol.*, 1993, vol. 175, pp. 3893–3896.
34. Menendez, M.C., Garcia, M.J., and Navarro, M.C., Gonzalez-y-Merchand, J.A., Rivera-Gutierrez, S., Garcia-Sanchez, L., and Cox, R.A., Characterization of an rRNA Operon (rrnB) of *Mycobacterium fortuitum* and Other Mycobacterial Species: Implications for the Classification of Mycobacteria, *J. Bacteriol.*, 2002, vol. 184, pp. 1078–1088.
35. Ji, Y.E., Colston, M.J., and Cox, R.A., The Ribosomal RNA (rrn) Operons of Fast-Growing Mycobacteria: Primary and Secondary Structures and Their Relation to rrn Operons of Pathogenic Slow-Growers, *Microbiology* (UK), 1994, vol. 140, pp. 2829–2840.
36. Springer, B., Kidan, Y.G., Prammananan, T., Elliot, K., Bottger, E.C., and Sander, P., Mechanisms of Streptomycin Resistance: Selection of Mutations in the 16S rRNA Gene Conferring Resistance, *Antimicrob. Agents Chemother.*, 2001, vol. 45, pp. 2877–2884.
37. Zillig, W., *The Order Thermococcales*, Balows, A. *et al.*, Eds., New York: Springer, 1992, pp. 702–706.
38. Asai, T., Zaporojets, D., Squires, C., and Squires, C.L., An *Escherichia coli* Strain with All Chromosomal rRNA Operons Inactivated: Complete Exchange of rRNA Genes between Bacteria, *Proc. Natl. Acad. Sci. USA*, 1999, vol. 96, pp. 1971–1976.
39. Amikam, D., Razin, S., and Glaser, G., Mycoplasmas (*Mollicutes*) Have a Low Number of rRNA Genes, *J. Bacteriol.*, 1984, vol. 158, pp. 376–378.
40. Suzuki, Y., Mori, T., Miyata, Y., and Yamada, T., The Number of Ribosomal RNA Genes in *Mycobacterium lepraemurium*, *FEMS Microbiol. Lett.*, 1987, vol. 44, pp. 73–76.
41. Wolfe, C.J. and Haygood, M.G., Bioluminescent Symbionts of the Caribbean Flashlight Fish (*Kryptophanaron alfredi*) Have a Single rRNA Operon, *Mol. Mar. Biol. Biotechnol.*, 1993, vol. 2, pp. 189–197.
42. Condon, C., Liveris, D., Squires, C., Schwartz, I., and Squires, C.L., rRNA Operon Multiplicity in *Escherichia coli* and the Physiological Implications of rrn Inactivation, *J. Bacteriol.*, 1995, vol. 177, pp. 4152–4156.
43. Clayton, R.A., Sutton, G., Hinkle, P.S., Bult, C., Jr., and Fields, C., Intraspecific Variation in Small-Subunit rRNA Sequences in GenBank: Why Single Sequences

May Not Adequately Represent Prokaryotic Taxa, *Int. J. Syst. Bacteriol.*, 1995, vol. 45, pp. 595–599.

44. Sriprakash, K.S. and Gardiner, D.L., Lack of Polymorphism within the rRNA Operons of Group a Streptococci, *Mol. Gen. Genet.*, 1997, vol. 255, pp. 125–130.
45. Grayson, T.H., Alexander, S.M., Cooper, L.F., and Gilpin, M.L., *Renibacterium salmoninarum* Isolates from Different Sources Possess Two Highly Conserved Copies of the rRNA Operon, *Antonie van Leeuwenhoek*, 2000, vol. 78, pp. 51–61.
46. Yoon, J.-H., Yim, D.K., Goodfellow, M., and Yong-Ha Park, Y.-H., Sequence Analysis of 16S rRNA Genes Amplified from Two Ribosomal RNA Gene Clusters of *Bifidobacterium bifidum*, *Antonie Van Leeuwenhoek*, 1999, vol. 75, pp. 329–333.
47. Mygind, T., Birkelund, S., and Christiansen, G., DNA Sequencing Reveals Limited Heterogeneity in the 16S rRNA Gene from the rrnB Operon among Five *Mycoplasma hominis* Isolates, *Int. J. Syst. Bacteriol.*, 1998, vol. 48, pp. 1067–1071.
48. Dryden, S.C. and Kaplan, S., Localization and Structural Analysis of the Ribosomal RNA Operons of *Rhodobacter sphaeroides*, *Nucleic Acids Res.*, 1990, vol. 18, pp. 7267–7277.
49. Peterka, M, and Avgustin, G., Ribosomal RNA Genes from *Prevotella bryantii*: Organization and Heterogeneity, *Folia Microbiol.* (Praha), 2001, vol. 46, pp. 67–70.
50. Moreno, C., Romero, J., and Espejo, R.T., Polymorphism in Repeated 16S rRNA Genes Is a Common Property of Type Strains and Environmental Isolates of the Genus *Vibrio*, *Microbiology* (UK), 2002, vol. 148, pp. 1233–1239.
51. Carbon, P., Ehresmann, C., Ehresman, B., and Ebel, J.P., The Complete Nucleotide Sequence of the Ribosomal 16S RNA from *Escherichia coli*. Experimental Details and Cistron Heterogeneities, *Eur. J. Biochem.*, 1979, vol. 100, pp. 299–410.
52. Ninet, B., Monod, M., Emler, S., Pawlowski, J., Metral, C., Rohner, P., Auckenthaler, R., and Hirschel, B., Two Different 16S rRNA Genes in a Mycobacterial Strain, *J. Clin. Microbiol.*, 1996, vol. 34, pp. 2531–2536.
53. Reischl, U., Feldmann, K., Naumann, L., Gaugler, B.J.M., Ninet, B., Hirschel, B., and Emlers, S., 16S rRNA Sequence Diversity in *Mycobacterium celatum* Strains Caused by Presence of Two Different Copies of 16S rRNA Gene, *J. Clin. Microbiol.*, 1998, vol. 36, pp. 1761–1764.
54. Liefting, L.W., Andersen, M.T., Beever, R.E., Gardner, R.C., and Forster, R.S., Sequence Heterogeneity in the Two 16S rRNA Genes of *Phormium yellow* Leaf Phytoplasma, *Appl. Environ. Microbiol.*, 1996, vol. 62, pp. 3133–3139.
55. Pettersson, B., Leitner, T., Ronaghi, M., Bolske, G., Uhlen, M., and Johansson, K.E., Phylogeny of the *Mycoplasma mycoides* Cluster as Determined by Sequence Analysis of the 16S rRNA Genes from the Two rRNA Operons, *J. Bacteriol.*, 1996, vol. 178, pp. 4131–4142.
56. Pettersson, B.L., Bolske, G., Thiaucourt, F., Uhlen, M., and Johansson, K.E., Molecular Evolution of *Mycoplasma capricolum* subsp. *capripneumoniae* Strains, Based on Polymorphisms in the 16S rRNA Genes, *J. Bacteriol.*, 1998, vol. 180, pp. 2350–2358.
57. Wezel, G.P., Van., Vijgenboom E., Bosch L. A Comparative Study of the Ribosomal RNA Operons of *Streptomyces coelicolor* A3 (2) and Sequence Analysis of rrnA, *Nucleic Acids Res.*, 1991, vol. 19, pp. 4399–4403.
58. Ogasawara, N., Nakai, S., and Joshikawa, H., Systematic Sequencing of the 180 Kilobase Region of the *Bacillus subtilis* Chromosome Containing the Replication Origin, *DNA Res.*, 1994, vol. 1, pp. 1–14.
59. Shimizu, T., Ohshima, S., Ohtani, K., Hoshino, K., Honjo, K., Hayashi, H., and Shimizu, T., Sequence Heterogeneity of the Ten rRNA Operons in *Clostridium perfringens*, *Syst. Appl. Microbiol.*, 2001, vol. 24, pp. 149–156.
60. Ueda, K., Seki, T., Kudo, T., Yoshida, T., and Kataoka, M., Two Distinct Mechanisms Cause Heterogeneity of 16S rRNA, *J. Bacteriol.*, 1999, vol. 181, pp. 78–82.
61. Cilia, V., Lafay, B., and Christen, R., Sequence Heterogeneities among 16S Ribosomal RNA Sequences and Their Effect on Phylogenetic Analyses at the Species Level, *Mol. Biol. Evol.*, 1996, vol. 13, pp. 451–461.
62. Mevarech, M., Hirsch-Twizer, S., Goldman, S., Jakobson, E., Eisenberg, H., and Dennis, P.P., Isolation and Characterization of the rRNA Gene Clusters of *Haloarcula marismortui*, *J. Bacteriol.*, 1989, vol. 171, pp. 3479–3485.
63. Mylvaganam, S. and Dennis, P.P., Sequence Heterogeneity between the Two Genes Encoding 16S rRNA from the Halophilic Archaebacterium *Haloarcula marismortui*, *Genetics*, 1992, vol. 130, pp. 399–410.
64. Dennis, P.P., Ziesche, S., and Mulvaganam, S., Transcription Analysis of Two Disparate rRNA Operons in the Halophilic Archaeon *Haloarcula marismortui*, *J. Bacteriol.*, 1998, vol. 180, pp. 4804–4813.
65. Amann, G., Stetter, K.O., Llobet-Brossa, E., Amann, R., and Anton, J., Direct Proof for the Presence and Expression of Two 5% Different 16S rRNA Genes in Individual Cells of *Haloarcula marismortui*, *Extremophiles*, 2000, vol. 4, pp. 373–376.
66. Kamekura, M., Diversity of Members of the Family *Halobacteriaceae*, *Microbiology and Biogeochemistry of Hypersaline Environments*, Oren, A., Ed., Boca Raton: CRC, 1999, pp. 13–25.
67. Oren, A., Ventosa, A., Gutierrez, M.C., and Kamekura, M., *Haloarcula quadrata* sp. nov., a Square, Motile Archaeon Isolated from a Brine Pool in Sinai (Egypt), *Int. J. Syst. Bacteriol.*, 1999, vol. 49, pp. 1149–1155.
68. Belfort, M., Reaban, M.E., Coetzee, T., and Dalgaard, J.Z., Prokaryotic Introns and Inteins: A Panoply of Form and Function, *J. Bacteriol.*, 1995, vol. 177, pp. 3897–3903.
69. Springer, N., Ludwig, W., Amann, R., Schmidt, H.J., Goertz, H.D., and Schleifer, K.H., Occurrence of Fragmented 16S rRNA in an Obligate Bacterial Symbiont of *Paramecium caudatum*, *Proc. Natl. Acad. Sci. USA*, 1993, vol. 90, pp. 9892–9895.
70. Linton, D., Dewhirst, F.E., Clewley, J.P., and Owen, R.J., Two Types of 16S Ribosomal RNA Genes Are Found in *Campylobacter helveticus*: Analysis, Application and Characterization of the Intervening Sequence Found in Some Strains, *Microbiology* (UK), 1994, vol. 140, pp. 847–855.

71. Linton, D., Clewely, P.J., Burnens, A., Owen, J.R., and Stanley, J., An Intervening Sequence (IVS) in the 16S rRNA Gene of the Eubacterium *Helicobacter canis*, *Nucleic Acids Res.*, 1994, vol. 22, pp. 1954–1958.

72. Rainey, F.A., Ward, N.L., Morgan, H.W., Toalster, R., and Stackebrandt, E., Phylogenetic Analysis of Anaerobic Thermophilic Bacteria: Aid for Their Reclassification, *J. Bacteriol.*, 1993, vol. 175, pp. 4772–4779.

73. Patel, B.K., Love, C.A., and Stackebrandt, E., Helix 6 of the 16S rRNA of the Bacterium *Desulfotomaculum australicum* Exhibits an Unusual Structural Idiosyncrasy, *Nucleic Acids Res.*, 1992, vol. 20, p. 5483.

74. Tourova, T.P., Kuznetsov, B.B., Novikova, E.V., Poltaraus, A.B., and Nazina, T.N., Heterogeneity of the Nucleotide Sequences of the 16S rRNA Genes of the Type Strain of *Desulfotomaculum kuznetsovii*, *Mikrobiologiya*, 2001, vol. 70, pp. 788–795.

75. Liao, D., Gene Conversion Drives within Genic Sequences: Concerted Evolution of Ribosomal RNA Genes in Bacteria and Archaea, *J. Mol. Evol.*, 2000, vol. 51, pp. 305–317.

76. Jarvis, E.D., Widom, R.L., LaFauci, G., Setoguchi, Y., Richter, I.R., and Rudner, R., Chromosomal Organization of rRNA Operons in *Bacillus subtilis*, *Genetics*, 1988, vol. 120, pp. 625–635.

77. Engel, J.N. and Ganem, D., Chlamydial rRNA Operons: Gene Organization and Identification of Putative Tandem Promoters, *J. Bacteriol.*, 1987, vol. 169, pp. 5678–5685.

78. Markowicz, Y., Loiseaux-de Goer, S., and Mache, R., Presence of a 16S rRNA Pseudogene in the Bi-Molecular Plastid Genome of the Primitive Brown Alga *Pylaiella littoralis.* Evolutionary Implications, *Curr. Genet.*, 1988, vol. 14, pp. 599–608.

79. Arnheim, N., Krystal, M., Schmickel, R., Wilson, G., Ryder, O., and Zimmer, E., Molecular Evidence for Genetic Exchanges among Ribosomal Genes on Nonhomologous Chromosomes in Man and Apes, *Proc. Natl. Acad. Sci. USA*, 1980, vol. 77, pp. 7323–7327.

80. Nedelcu, A.M., Complex Patterns of Plastid 16S rRNA Gene Evolution in Nonphotosynthetic Green Algae, *J. Mol. Evol.*, 2001, vol. 53, pp. 670–679.

81. Hillis, D.M., Moritz, C., Porter, C.A., and Baker, R.J., Evidence for Biased Gene Conversion in Concerted Evolution of Ribosomal DNA, *Science*, 1991, vol. 251, pp. 308–310.

82. Kowalczykowski, S.C., Dixon, D.A., Eggleston, D.A., Lauder, S.D., and Rehrauer, W.M., Biochemistry of Homologous Recombination in *Escherichia coli*, *Microbiol. Rev.*, 1994, vol. 58, pp. 401–465.

83. Liu, S.-L. and Sanderson, K.E., Highly Plastic Chromosomal Organization in *Salmonella typhi*, *Proc. Natl. Acad. Sci. USA*, 1996, vol. 93, pp. 10303–10308.

84. Lan, R. and Reeves, P.R., Recombination between rRNA Operons Created Most of the Ribotype Variation Observed in the Seventh Pandemic Clone of *Vibrio cholerae*, *Microbiology* (UK), 1998, vol. 144, pp. 1213–1221.

85. Wang, Y. and Zhang, Z., Comparative Sequence Analyses Reveal Frequent Occurrence of Short Segments Containing an Abnormally High Number of Nonrandom Base Variations in Bacterial rRNA Genes, *Microbiology* (UK), 2000, vol. 146, pp. 2845–2854.

86. Mankin, A.S., Pactamycin Resistance Mutations in Functional Sites of 16S rRNA, *J. Mol. Biol.*, 1997, vol. 274, pp. 8–15.

87. Ralph, D. and McClelland, M., Phylogenetic Evidence for Horizontal Transfer of an Intervening Sequence between Species in a Spirochete Genus, *J. Bacteriol.*, 1994, vol. 176, pp. 5982–5987.

88. Bercovier, H., Brenner, D.J., Ursing, J., Steigerwalt, A.G., Fanning, G.R., Alonso, J.M., Carter, G.P., and Mollaret, H.H., Characterization of *Yersinia enterocolitica* Sensu Stricto, *Curr. Microbiol.*, 1980, vol. 4, pp. 201–206.

89. Konkel, M.E., Marconi, R.T., Mead, D.J., and Cieplak, W., Jr., Identification and Characterization of an Intervening Sequence within the 23S Ribosomal Genes of *Campilobacter jejuni*, *Mol. Microbiol.*, 1994, vol. 14, pp. 235–241.

90. Jain, R., Rivera, M.C., and Lake, J.A., Horizontal Gene Transfer among Genomes: the Complexity Hypothesis, *Proc. Natl. Acad. Sci. USA*, 1999, vol. 96, pp. 3801–3806.

91. Wang, G.C.Y. and Wang, Y., The Frequency of Chimeric Molecules as a Consequence of PCR Co-Amplification of 16S rRNA Genes from Different Bacterial Species, *Microbiology* (UK), 1996, vol. 142, pp. 1107–1114.

92. Wintzingerode, F., Gobel, U.B., and Stackebrandt, E., Determination of Microbial Diversity in Environmental Samples: Pitfalls of PCR-Based rRNA Analysis, *FEMS Microbiol. Rev.*, 1997, vol. 21, pp. 213–229.

93. Liu, Y., Karnauchow, T.M., Jarrell, K.F., Balkwill, D.L., Drake, G.R., Ringelberg, D., Clarno, R., and Boone, D.R., Description of Two New Thermophilic *Desulfotomaculum* spp., *Desulfotomaculum putei* sp. nov., from a Deep Terrestrial Subsurface, and *Desulfotomaculum luciae* sp. nov., from a Hot Spring, *Int. J. Syst. Bacteriol.*, 1997, vol. 47, pp. 615–621.

94. Tourova, T.P., The Role of DNA–DNA Hybridization and 16S rRNA Gene Sequencing in Solving Taxonomic Problems by the Example of the Order *Haloanaerobiales*, *Mikrobiologiya*, vol. 69, pp. 741–752.

95. Fox, G.E., Wisotzkey, J.D., and Jurtshuk, P., Jr., How Close Is Close: 16S rRNA Sequence Identity May Not Be Sufficient to Guarantee Species Identity, *Int. J. Syst. Bacteriol.*, 1992, vol. 42, pp. 166–170.

96. Stackebrandt, E. and Goebel, B.M., Taxonomic Note: A Place for DNA–DNA Reassociation and 16S rRNA Sequence Analysis in the Present Species Definition in Bacteriology, *Int. J. Syst. Bacteriol.*, 1994, vol. 44, pp. 846–849.

97. 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. Applic. Biosci.*, 1994, vol. 10, pp. 569–570.