Is the 16S–23S rRNA Internal Transcribed Spacer Region a Good Tool for Use in Molecular Systematics and Population Genetics? A Case Study in Cyanobacteria

Sarah L. Boyer, Valerie R. Flechtner, and Jeffrey R. Johansen

Department of Biology, John Carroll University

We amplified, TA-cloned, and sequenced the 16S–23S internal transcribed spacer (ITS) regions from single isolates of several cyanobacterial species, *Calothrix parietina, Scytonema hyalinum, Coelodesmium wrangelii, Tolypothrix distorta,* and a putative new genus (isolates SRS6 and SRS70), to investigate the potential of this DNA sequence for phylogenetic and population genetic studies. All isolates carried ITS regions containing the sequences coding for two tRNA molecules ($tRNA^{Ile}$ and $tRNA^{Ala}$). We retrieved additional sequences without tRNA features from both *C. parietina* and *S. hyalinum.* Furthermore, in *S. hyalinum,* we found two of these non-tRNA-encoding regions to be identical in length but different in sequence. This is the first report of ITS regions from a single cyanobacterial isolate not only different in configuration, but also, within one configuration, different in sequence. The potential of the ITS region as a tool for studying molecular systematics and population genetics is significant, but the presence of multiple nonidentical rRNA operons poses problems. Multiple nonidentical rRNA operons may impact both studies that depend on comparisons of phylogenetically homologous sequences and those that employ restriction enzyme digests of PCR products. We review current knowledge of the numbers and kinds of 16S–23S ITS regions present across bacterial groups and plastids, and we discuss broad patterns congruent with higher-level systematics of prokaryotes.

Introduction

During the last decade, biologists have employed a variety of molecular techniques to address questions about phylogeny, evolution, and population diversity. Analysis of 16S (small subunit) rRNA and, more recently, the 16S–23S internal transcribed spacer (ITS) has figured heavily in these studies, particularly those involving prokaryotic and eukaryotic microorganisms.

Several phycologists have used the 16S rRNA gene to provide insight into the phylogenetic relationships of cyanobacterial genera within the orders proposed by Komárek and Anagnostidis (1986, 1989) (e.g., Giovannoni et al. 1988; Wilmotte and Golubic 1991; Wilmotte et al. 1992; Wilmotte, Neefs, and De Wachter 1994; Nelissen et al. 1996; Nübel et al. 1996; Turner 1997). However, Fox, Wisotzkey, and Jurtshuk (1992) concluded that identity in 16S rRNA sequence data was not sufficient grounds for establishing species identity and thus not appropriate for studies at the subgeneric level. As a result, researchers have increasingly turned to the more variable 16S–23S ITS.

Restriction enzyme digestion of the 16S–23S ITS region has been used for phylogenetic analyses of strains of nonphotosynthetic eubacterial genera (Navarro et al. 1992; Vinuesa et al. 1998). Restriction digests of this region have also been used to examine variability and phylogenetic relationships within orders of cyanobacteria (Lu et al. 1997), among genera of heterocystous cyanobacteria (West and Adams 1997), and among strains in a single filamentous genus (Scheldeman et al. 1999). Otsuka et al. (1999) were the first and are currently the only research group to use direct sequencing of the ITS to study subgeneric phylogenetic relationships in a cyanobacterial genus, Microcystis. They found that the phylogeny based on the ITS data did not correlate perfectly with established Microcystis morphospecies or phycoerythrin production, although concordance with microcystin production was evident.

Operons containing the genes coding for the three rRNAs (16S, 23S, 5S) and their associated ITS regions are normally present in multiple copies in the bacterial genome (7 in *Escherichia coli* and Salmonella, 10 in *Bacillus subtilis*). In *E. coli,* these copies are named rrnA–rrnE, rrnH, and rrnG. Antón, Martínez-Murcia, and Rodríguez-Valera (1998) have demonstrated that there are major heterogeneities among operons in terms of the type and number of tRNA genes present. In *E. coli* K12, within the 16S–23S ITS, operons rrnB, rrnC, rrnE, and rrnG contain a gene coding for $tRNA^{Glu-2}$, whereas operons rrnA, rrnD, and rrnH have genes for $tRNA^{Ile-1}$ and $tRNA^{Ala-1B}$. There are other major heterogeneities in various regions of the ITS within each of these two general ''types'' of ITS region. It has been suggested that the spacer sequence could reflect intraspecies phylogeny (García-Martínez et al. 1996), but that to target the same operon (rrnA, rrnB, rrnC, rrnD, rrnE, rrnH, or rrnG) in multiple strains, operon-specific primers must be used in PCR (Antón, Martínez-Murcia, and Rodríguez-Valera 1998).

In this study, we examined variability in the ITS among multiple rRNA operons in five species: *Scytonema hyalinum, Tolypothrix distorta, Calothrix parietina, Coelodesmium wrangelii,* and a putative new genus (isolates SRS6 and SRS70). These species represent three of the four families currently recognized in the order Nostocales. Most of the 16S rRNA data available for members of this order come from representatives of the fourth family (Nostocaceae), including sequences

Key words: Internal transcribed spacer, cyanobacteria, intergenic spacer region, rRNA, phylogenetics, population genetics.

Address for correspondence and reprints: Valerie R. Flechtner, Department of Biology, John Carroll University, 20700 North Park Boulevard, University Heights, Ohio 44118. E-mail: flechtner@jcu.edu.

Mol. Biol. Evol. 18(6):1057–1069. 2001

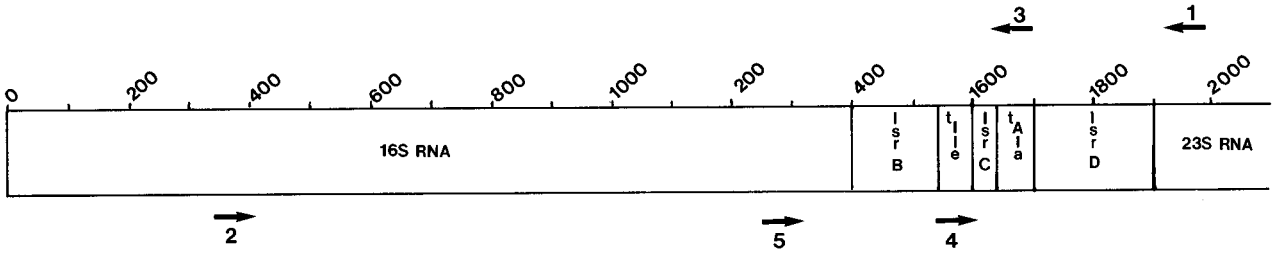

FIG. 1.—Map of primer sites in the 16S–23S rRNA operon that were used in this study. Primers 1 and 2 were used in the first PCR reaction, and that product was reamplified using primers 1 and 5 to obtain the ''short'' PCR product used in cloning. Primers 3 and 4 were used when it was not possible to obtain an ITS region containing both tRNA genes with primers 1 and 5. ISR-B, ISR-C, and ISR-D are noncoding intergenic spacer regions which surround the tRNA genes for isoleucine (t^{Ile}) and alanine (t^{Ala}).

from Nostoc and Anabaena. ITS sequences have also recently been reported for several species of Nostoc (Lu 1999; Iteman et al. 2000; Li 2000). Very little data exist from the three families that we examined: partial 16S rRNA sequences are available in GenBank from three isolates of Scytonema (in the family Scytonemataceae), but no sequence data exist for members of the families Microchaetaceae (represented in this study by *T. distorta* and *C. wrangelii*) or Rivulariaceae (represented by *C. parietina*). Therefore, the data presented here are of value to anyone interested in understanding relationships within the order Nostocales. In addition, our study is the first comparison of ITS sequence similarities among multiple members of a cyanobacterial order. We review the current knowledge of the numbers and kinds of ITS regions present across bacteria and plastids and discuss the potential utility of the ITS region as a tool for both broad- and fine-scale phylogeny reconstruction.

Materials and Methods

Collection and Isolation of Cyanobacterial Strains

Cyanobacterial isolates used in this study were all collected, isolated, and identified by V.R.F. and J.R.J. For those cyanobacteria isolated from soils, dry soil samples were crushed, subsampled, and dilution plated as described in Flechtner (1999). Cyanobacteria were isolated into unialgal culture from the plates and kept on agar slants of Z-8 medium (Carmichael 1986). *Coelodesmium wrangelii* was isolated directly from a stream sample onto agar slants of Z-8 medium. All isolates were examined on Olympus photomicroscopes with Nomarski DIC optics. Strains were kept in dim light (<50 $\mu E/cm^2/s$ illuminance) at 7°C on a 12:12 h light : dark cycle.

DNA Extraction

DNA was extracted from 20 mg of fresh unialgal tissue using the Cullings (1992) modification of the Doyle and Doyle (1987) technique. The resultant DNA was suspended in 50 μl TE and stored at −20°C.

Polymerase Chain Reaction

Primers were designed after Wilmotte, Van der Auwera, and De Wachter (1993), Wilmotte (1994), and Nübel, Garcia-Pichel, and Muyzer (1997). They were designated primer 1 (5′-CTC TGT GTG CCT AGG TAT CC-3′; after Wilmotte, Van der Auwera, and De Wachter 1993), primer 2 (5′-GGG GGA TTT TCC GCA ATG GG-3′; after Nübel, Garcia-Pichel, and Muyzer 1997), primer 3 (5′-CGC TCT ACC AAC TGA GCT A-3′; after Wilmotte 1994), primer 4 (5′-ATT AGC TCA GGT GGT TAG-3′ after Wilmotte, Van der Auwera, and De Wachter 1993), and primer 5 (5′-TGT ACA CAC CGG CCC GTC-3′; after Wilmotte, Van der Auwera, and De Wachter 1993).

The positions of these primers with regard to the 16S RNA gene, the 23S RNA gene, and the transfer RNA genes that had previously been found between them are shown in figure 1. Primers (Midland Certified Reagent Company) were made up in 100 μM stock solutions. For use in PCR, a mix of 1.2 μl each of two primers and 7.6 μl sterile water was made.

Initially, each DNA sample was amplified using primers 1 and 2. This resulted in a product approximately 1,600 bp long (''long PCR''), which was then used as a template for a reamplification using primers 1 and 5, resulting in a product approximately 600 bp long (''short PCR''). Each 100-μl reaction contained 86 μl sterile water, 10μl 10 × buffer (Promega), 0.5 μl of each dNTP (G, A, T, C) at 10 mM, 0.5 μl of the primer mixture described above, 0.5 μl *Taq* polymerase (Promega), and, typically, 1.0 μl template DNA (5–10 ng).

PCR was optimized for each species, with the template DNA amount varying from 0.5 to 2.0 μl and the annealing temperatures varying from 55°C to 57°C. The most commonly used profile for the long-PCR reaction using primers 1 and 2 was 94°C for 1 min, 57°C for 1 min, and 72°C for 4 min (35 cycles), followed by a 10-min extension at 72°C. For the short-PCR reamplifications, the most commonly used profile was 94°C for 1 min, 56°C for 45 s, and 72°C for 2 min (20 cycles). Reactions were carried out using Thermolyne's Amplitron and Temptronic thermocyclers. Results were checked using a 1% agarose gel.

Sequencing

Short PCR product was cloned into plasmids containing the sites for the universal primers M13 forward and reverse on either side of the cloning site using Invitrogen's TOPO TA Cloning Kit for Sequencing, version A. Plasmid DNA was generally obtained from three of the resultant clones using Qiagen's QiaPrep Spin Kit.

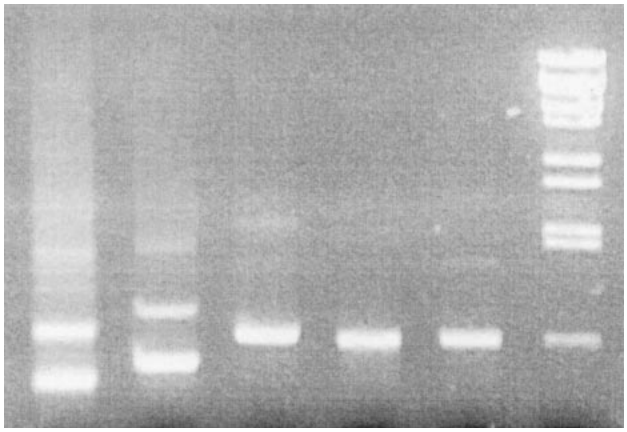

FIG. 2.—Digital photograph of an agarose gel with internal transcribed spacer (ITS) (short) PCR products from the following taxa (left to right): *Calothrix parietina, Syctonema hyalinum, Coelodesmium wrangelii*, putative new genus (isolate SRS70), and *Tolypothrix distorta.* To the right is λ *Bst*EII digest. Note that *S. hyalinum* and *C. parietina*, both of whose short-PCR products show two distinct bands, are the taxa from which we retrieved multiple ITS sequences of different lengths.

Automated sequencing was performed by Cleveland Genomics with the universal primers M13 forward and reverse.

Data Analysis

Forward and reverse primer sequences were checked against each other by generating the reverse complement of the ''reverse'' sequence using Oxford Molecular Group's Omiga and aligning it with the ''forward'' sequence using the CLUSTAL W Multiple Sequence Alignment Program, version 1.7 (Thompson, Higgins, and Gibson 1994) via the Baylor College of Medicine's Search Launcher (Smith et al. 1996) at http://dot.imgen.bcm.tmc.edu:9331/. This resulted in the longest possible read of the sequence, in addition to acting as a check on the sequencing. Sequences were aligned using CLUSTAL W. These alignments were checked by eye.

GenBank Numbers

ITS sequence data for species examined in this study were deposited with GenBank. Accession numbers are as follows: *C. parietina,* AF236642 (two tRNAs) and AF236643 (no tRNAs); *C. wrangelii*, AF236652; *S. hyalinum,* AF236650 (no tRNAs), AF236651 (two tRNAs), AY007688 (no tRNAs); *T. distorta,* AY007689; Tolypothrix field sample, AF236644–AF236649 (six sequences); putative new genus, AF2326659.

Results

Agarose Gel Electrophoretic Analysis of PCR Products

Agarose gel electrophoresis of ITS PCR products (short PCR) from five cyanobacterial genera (fig. 2) revealed obvious differences among the genera. We consistently obtained a single band from both isolates of the putative new genus (SRS6 and SRS70), *C. wrangelii,* and *T. distorta* and two or more bands from *S. hyalinum* and *C. parietina*. The presence of multiple bands in PCR reactions suggested that at least some of our isolates possessed multiple rRNA operons.

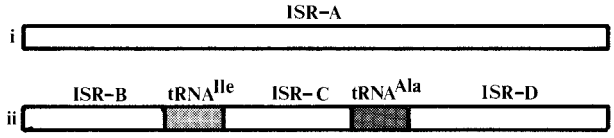

FIG. 3.—Stylized map showing two patterns observed in the 16S–23S rRNA ITS region within the cyanobacterial isolates examined in this study. The four intergenic spacer regions (ISRs) are labeled based on their relationship to the tRNA genes present.

Sequence Data from the 16S–23S ITS Region

To investigate the possibility that multiple rRNA operons existed in the genomes of at least some of our cyanobacterial species, we cloned our short PCR products into competent *E. coli* cells. At least three individual clones per isolate were selected for automated sequencing.

We sequenced the ITS regions from two isolates of the putative new genus (SRS6 and SRS70) and one isolate each for *C. wrangelii, T. distorta, S. hyalinum,* and *C. parietina*. The sizes of the 16S–23S ITS regions in these genera ranged from 347 in *C. parietina* to 648 bp in one sequence from *S. hyalinum.* Two sequence organization patterns were evident: some 16S–23S ITS sequences contained no tRNA molecule sequences (fig. 3, i) and some contained both $tRNA^{Ala}$ and $tRNA^{Ile}$ (fig. 3, ii). We have denoted the intergenic spacer regions (ISRs) located within the ITS and between the various coding sequences using the following designations: ISR-A (16S–23S, without interruption from tRNAs) (fig. 3, i), ISR-B (16S-$tRNA^{Ile}$), ISR-C ($tRNA^{Ile}$-$tRNA^{Ala}$), and ISR-D ($tRNA^{Ala}$-23S) (fig. 3, ii).

Variability in ITS and ISR Size and Sequence Among Species and Within Isolates

We found striking variability in the size of the ITS region among the various species examined (fig. 4). Even when sequences from multiple taxa showed the same overall pattern of configuration, the sizes of the individual 16S–23S ITS regions often differed tremendously. For example, while all species carry at least one copy of the 16S–23S ITS region containing both $tRNA^{Ile}$ and $tRNA^{Ala}$, the size of the entire 16S–23S ITS region for this operon is 491 bp in *C. wrangelii* and 648 bp in *S. hyalinum* (fig. 4). The size difference is particularly noticeable in ISR-C, which contains 83 bp in *S. hyalinum* but a mere 9 bp in *C. parietina.*

We found evidence for multiple nonidentical rRNA operons in *S. hyalinum* and *C. parietina. Calothrix parietina* carries at least one two-tRNA-containing and at least one tRNA-lacking operon. *Scytonema hyalinum* carries at least one two-tRNA-containing and at least two tRNA-lacking operons. The two tRNA-lacking ITS regions of *S. hyalinum* are both 400 bp long but non-

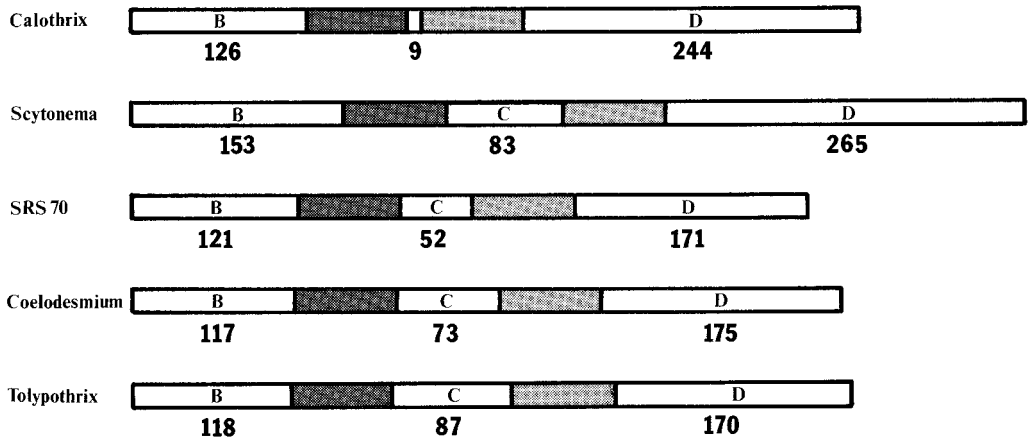

FIG. 4.—Map of 16S rRNA internal transcribed spacer regions containing both tRNA genes in five isolates of heterocystous cyanobacteria from soils. Letters in the sequence refer to intergenic spacer region (ISR) designation (B = ISR-B, etc.). Sequence length (in bp) is given below each ISR.

identical in sequence (fig. 5). We noted the presence of all structural features of the ITS identified by Iteman et al. (2000) in each of these three *S. hyalinum* sequences (fig. 5).

Discussion

During the last 15 years, our understanding of evolutionary relationships among microorganisms has expanded dramatically, due in large part to the use of 16S (small ribosomal subunit [SSU]) sequence data. This approach, pioneered by Carl Woese (Woese, Kandler, and Wheelis 1990), has become so widely used that 16S sequence data will be the basis for defining taxonomic groups in the second edition of *Bergey's Manual of Determinative Bacteriology.* However, some investigators (e.g., Fox, Wisotzkey, and Jurtshuk 1992) have questioned whether sufficient variability exists in 16S RNA to allow discrimination among species of a genus or strains of a species. It has been suggested that the ITS region separating 16S and 23S sequences in the rRNA operon might be useful for these fine levels of discrimination. In the discussion that follows, we review current knowledge of the number of rRNA operons and patterns of 16S–23S ITS configurations therein in bacteria and plastids, with special focus on cyanobacteria. We finish by discussing the usefulness of ITS data for phylogenetic studies.

ITS Patterns of Configuration Reflect Higher-Level Phylogenetic Groupings

Gürtler and Stanisich (1996) compared the sizes and makeups of the 16S–23S ITS regions of 44 species in 27 genera of bacteria. During the last four years, many new sequences have been added to the database; all are presented in table 1. Collectively, the data show three general patterns of sequence composition in 16S–23S ITS regions. The two most common patterns are the absence of tRNA sequences (as in fig. 3, i) or the presence of two tRNA sequences ($tRNA^{Ile}$ and $tRNA^{Ala}$, as in fig. 3, ii). In the two tRNA-containing ITS regions, in all but one taxon the $tRNA^{Ile}$ is just downstream of the 16S rRNA and the $tRNA^{Ala}$ is just upstream of the 23S rRNA; the order is reversed only in the recently sequenced plant pathogen *Xylella fastidiosa* (Simpson et al. 2000). A third pattern noted is the presence of a lone tRNA which may be $tRNA^{Ala}$, $tRNA^{Glu}$, or $tRNA^{Ile}$.

If one considers the 16S–23S ITS sequence data currently available across phylogenetic groups, certain patterns begin to emerge. In the Archaea, members of the kingdom Crenarchaeota have ITS regions lacking tRNA genes, while members of the Euryarchaeota carry $tRNA^{Ala}$ (table 1). Among the Eubacteria, members of the primitive Aquaficales and Thermotoga carry both $tRNA^{Ile}$ and $tRNA^{Ala}$. Proteobacteria display all varieties of ITS configuration, including $tRNA^{Glu}$ only, which has thus far been found exclusively in the γ-division of this group. Among the gram-positive bacteria (Firmicutes), there is a difference in ITS types observed between organisms of low G-C content and organisms of high G-C content. Those of low G-C content (the Bacillus/Clostridium group) display a variety of ITS configurations. Most taxa in this group lack tRNA sequences, and of those that carry tRNA sequences, two tRNAs are present in five species, $tRNA^{Ala}$ alone is present in three species, and $tRNA^{Ile}$ alone is present in a single species (*Staphylococcus aureus*) (table 1). In contrast, all of the gram-positive bacteria of high G-C content (Actinobacteria) have ITS regions lacking tRNA sequences. Among the cyanobacteria, by far the most common pattern of configuration is the presence of both $tRNA^{Ile}$ and $tRNA^{Ala}$; ITS regions in this group may also lack tRNA sequences or carry $tRNA^{Ile}$ alone (table 1). In plastids, all ITSs contain both $tRNA^{Ile}$ and $tRNA^{Ala}$, except in *Toxoplasma gondii,* which has a split rRNA operon with no ITS (Kohler et al. 1997), and in the plastids of the parasitic plants *Conopholis americana* (Wimpee, Morgan, and Wrobel 1992) and *Epifagus virginiana* (Wolfe et al. 1992), which have both undergone a deletion event eliminating most or all of the tRNA from the ITS region. It therefore appears, on the basis of the limited amount of data available, that 16S–23S ITS composition mirrors

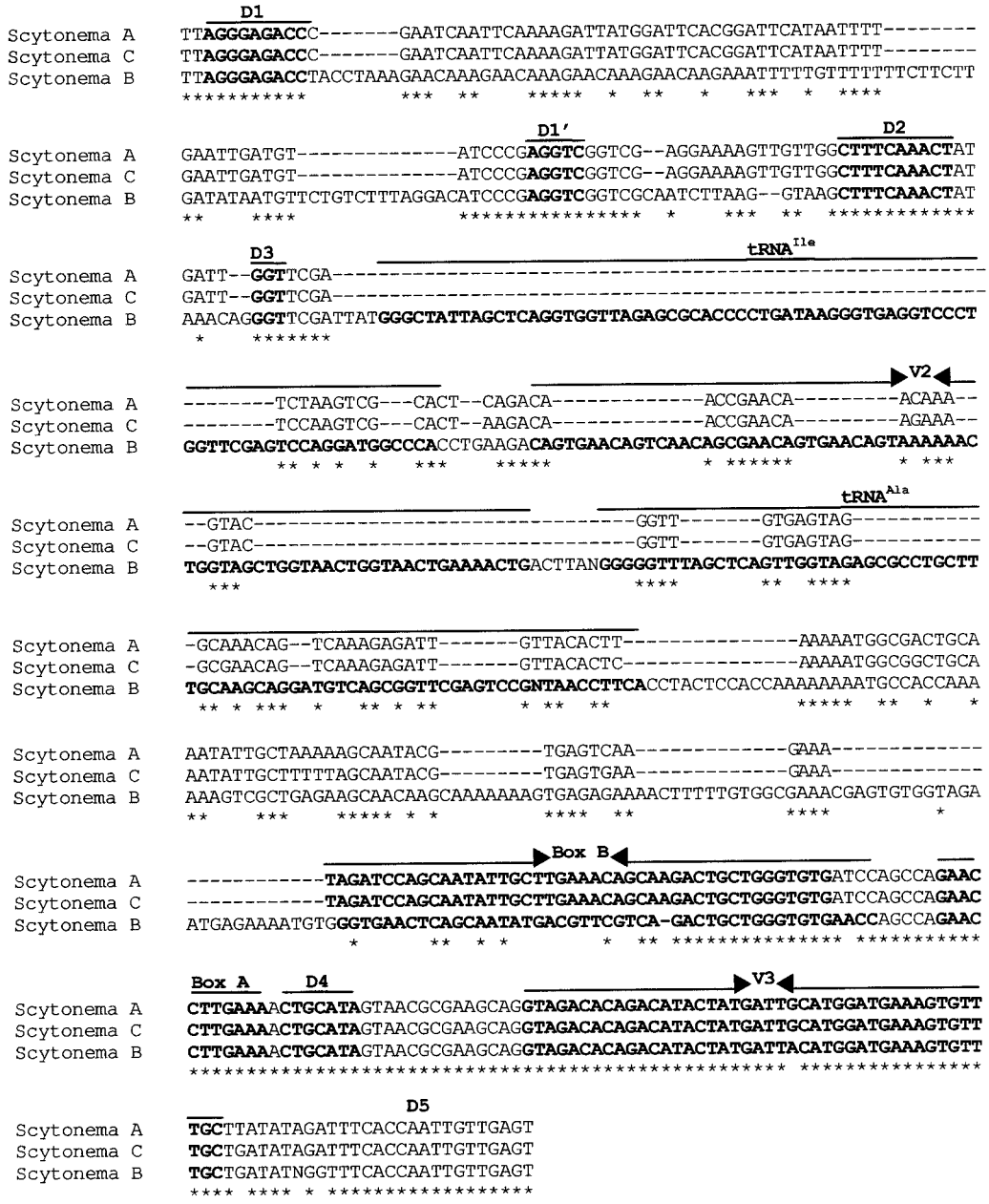

FIG. 5.—Alignment of the 16S–23S rRNA internal transcribed spacer (ITS) region from a single isolate of *Scytonema hyalinum*. One sequence (clone B) contains both $tRNA^{Ile}$ and $tRNA^{Ala}$; the other two (clones A and C) contain neither tRNA and are identical in length but nonidentical in sequence. Note that all three sequences contain all important structural elements of the ITS, as described by Iteman et al. (2000). Tips of stem structures are indicated between arrowheads. Region D5 lies between V3 and 23S but cannot be specifically identified in the absence of sequence data from the 23S–5S ITS.

higher phylogenetic groupings. It will be interesting to see if these patterns hold up as additional ITS regions are sequenced, and it is possible that the configuration of the ITS region could be a character useful for higher-level phylogeny reconstruction.

ITS Configurations in Cyanobacteria and Plastids

As a group, cyanobacteria show three types of 16S–23S ITS region sequence composition. ITS regions containing both $tRNA^{Ile}$ and $tRNA^{Ala}$ are most common and have been previously reported in members of the genera Anabaena (Lu 1999), Arthrospira (Nelissen et al. 1994), Nostoc (Lu 1999; Iteman et al. 2000), Synechococcus (''Anacystis'') (Tomioka and Sugiura 1984), and Trichodesmium (Wilmotte, Neefs, and De Wachter 1994). Our study adds the genera Calothrix, Coelodesmium, Scytonema, Tolypothrix, and a putative new genus (isolates SRS6 and SRS70) to this group (table 1). ITS regions with $tRNA^{Ile}$ only have been reported in the genera Microcystis (Otsuka et al. 1999), Spirulina (Nelissen et al. 1994), and Synechocystis (Kaneko et al. 1996). It appears that this latter configuration is more frequently seen in the cyanobacteria than in other groups

Table 1
16S–23S rRNA ITS Data Currently Available from Archaea, Eubacteria, and Plastids

Organism	Operon No.	tRNA Type	Size (bp)	Reference
Archaea				
Crenarchaeota				
Sulfolobus solfataricus	1	N	237	Olsen et al. (1985)
Thermoproteus tenax	1	N	60	Kjems et al. (1987)
Euryarchaeota				
Archaeoglobus fulgidus	1	A	236	Klenk et al. (1997)
Methanobacterium thermoautotrophicum delta H	2	A	328	Smith, et al. (1997)
Methanococcus jannaschii	2	A	337–346	Bult et al. (1996)
Pyrococcus horikoshii OT3	1	A	288	Kawarabayasi et al. (1998)
Eubacteria				
Aquificales				
Aquifex aeolicus	2	IA	315	Deckert et al. (1998)
Thermotoga				
Thermotoga maritima	1	IA	240	Nelson et al. (1999)
Thermus/Deinococcus group				
Deinococcus radiodurans	?	S		White et al. (1999)
Proteobacteria				
α division				
Bartonella (five species)	?	IA	906–1028	Minnick, Strange, and Williams (1994); Roux and Raoult (1995)
Caulobacter crescentus	?	IA	653	Feingold et al. (1995); Anderson et al. (1998)
Rickettsia prowazekii Madrid E	1	S		
β division				
Burkholderia cepacia	?	IA	530–580	#
		N	394	
Thiobacillus ferooxidans	2	IA	440	Sagredo, Jedlicki, and Orellana (1992)
Neisseria meningitidis	4	IA	664	Tettelin et al. (2000)
		N	664	
γ division				
Aeromonas hydrophila	?	G	511	East and Collins (1993)
Escherichia coli	7	G	355–437	Harvey et al. (1988)
		IA	441	Young, Macklis, and Steitz (1979)
Haemophilus influenzae	6	G	478	Fleischmann et al. (1995)
		IA	723	
Plesiomonas shigelloides	?	G	264	East, Allaway, and Collins (1992)
Pseudomonas (five species)	?	IA	542–642	Gill et al. (1994)
Rhodobacter sphaerodies	3	IA	667	Dryden and Kaplan (1990)
Salmonella typhimurium	7	G	~400	Lehner, Harvey, and Hill (1984)
Xylella fastidiosa	2	AI	466	Simpson et al. (2000)
ε division				
Helicobacter pylori 26695	2	S		Tomb et al. (1997)
Firmicutes (gram-positive)				
Low G-C content (Bacillus/Clostridium group)				
Bacillus subtilis	10	IA	344	Loughney, Lund, and Dahlberg (1982); Green et al. (1985)
		N	164	
Clavibacter michiganensis subsp. *michiganensis*	?	N	485	#
Clostridium botulinum	?	N	365–437	Campbell et al. (1993)
Enterococcus faecalis	?	A	327	Hall (1994)
		N	225	
Enterococcus hirae	6	A	318	Sechi and Daneo-Moore (1993)
		N	138	
Lactobacillus (four species)	?	N	205–219	Nakagawa et al. (1994)
Listeria monocytogenes	?	N	243	Emond, Fliss, and Pandian (1993)
Mycoplasma pneumoniae	1	N	253	Harasawa et al. (1993)
Staphylococcus aureus	9	IA	303–551	Gürtler and Barrie (1995)
		I	335–460	
		N	330–469	
Streptococcus (three species)	?	IA	270–391	Whiley et al. (1995)
Streptococcus pneumoniae	6	A	294	Bacot & Reeves (1991)
Urealpasma urealyticum	2	N	348	#
High G-C content (Actinobacteria)				
Acholeplasma laidlawii	2	N	167	Nakagawa et al. (1992)
Frankia strain ORS020606	2	N	411	Normand et al. (1992)

Table 1
Continued

Organism	Operon No.	tRNA Type	Size (bp)	Reference
Mycobacterium phlei	1	N	363	Van Der Giessen et al. (1994)
Mycobacterium tuberculosis	1	N	276	Van Der Giessen et al. (1994)
Streptomyces ambofaciens	4	N	303	Pernodet et al. (1989)
Streptomyces griseus subsp. *griseus*	6	N	278–303	Kim et al. (1993)
Streptomyces lividans TK21	6	N	278	Suzuki et al. (1988)
Spirochaetales				
Treponema pallidum	2	I	294	Fraser et al. (1998)
		A	294	
Borrelia (three species)	1	S		Ojaimi et al. (1994)
Cyanobacteria				
Anabaena (four species)	?	IA	?	Lu (1999)
Arthrospira PCC 7345	?	IA	?	Nelissen et al. (1994)
Calothrix parietina	?	IA	529	
		N	329	
Coelodesmium wrangelii	?	IA	515	
Microcystis (47 strains)	?	I	~360	Otsuka et al. (1999)
Nodularia BCNOD 9427	?	N	354	Hayes and Barker (1997)
Nostoc PCC 7120	4(?)	IA	512	Iteman et al. (2000)
		N	283	
Nostoc (three species)	?	IA	?	Lu (1999)
Scytonema hyalinum	?	IA	651	
		N	400	
Spirulina PCC 6313	?	I	?	Nelissen et al. (1994)
Synechococcus PCC 6301	2	IA	545	Tomioka and Sugiura (1984), Williamson and Doolittle (1983)
(''*Anacystis nidulans*'')		I^{ps}?		
Synechococcus PCC 7002	2	?	?	Chen and Widger (1993)
Synechocystis PCC 6803	2	I	?	Kaneko et al. (1996)
Tolypothrix distorta	?	IA	525	
Trichodesmium NIBB1067	?	IA	547	Wilmotte, Neefs and De Wachter (1994)
Isolates SRS6, SRS70	?	IA	494	
Eucaryotic plastids				
Euglenozoa				
Euglena gracilis	3	IA	267	Hallick et al. (1993)
Alveolata				
Toxoplasma gondii	?	S		Kohler et al. (1997)
Cryptophyta				
Gullardia theta	2	IA	276	Douglas and Penny (1999)
Glaucocystophyceae				
Cyanophora paradoxa	2	IA	288	Helmchen, Bhattacharya, and Melkonian (1995)
Rhodophyta (red algae)				
Chondrus crispus	?	IA	277	Leblanc, Boyen, and Louiseaux-de Goer (1995)
Cyanidium caldarium	1	IA	278	#
Bacillariophyta (diatoms)				
Odontella sinensis	2	IA	282	Kowallik et al. (1995)
Chlorophyta (green algae)				
Chlorella vulgaris	1	IA	1,006	Wakasugi et al. (1997)
Mesostigma viride	2	IA	337	Lemieux, Otis, and Turmel (2000)
Nephroselmis olivacea	2	IA	182	Turmel, Otis, and Lemieux (1999)
Charophyta				
Coleochaete orbicularis	1	IA	1,585	Manhart and Palmer (1990)
Embryophyta				
Hepatophyta (liverwort)				
Marchantia polymorpha	2	IA	2,308	Kochi et al. (1988)
Spermatophyta (seed plants)				
Arabidopsis thaliana	2	IA	2,189	Sato et al. (1999)
Conopholis americana	?	N	398	Wimpee, Morgan, and Wrobel (1992)
Epifagus virginiana	?	IA^{ps}	2,010	Wolfe et al. (1992)
Nicotiana tabacum	2	IA	2,079	Shinozaki et al. (1986)
Oryza sativa	2	IA	2,424	Shimada and Sugiura (1991)
Pinus thunbergii	1	IA	2,519	Tsudzuki et al. (1992)
Spinacea oleracea	2	IA	2,218	#
Zea mays	2	IA	2,411	Maier et al. (1995)

NOTES.—A = $tRNA^{Ala}$, G = $tRNA^{Glu}$; I = $tRNA^{Ile}$; IA = $tRNA^{Ile}$ and $tRNA^{Ala}$; AI = $tRNA^{Ala}$ and $tRNA^{Ile}$; I^{ps} = a partial sequence of $tRNA^{Ile}$; IA^{ps} = a partial sequence of $tRNA^{Ile}$ and $tRNA^{Ala}$; N = no tRNA present in the ITS; S = split operon; ? = unknown; # = sequences reported only in GenBank.

of eubacteria (table 1). Spacer regions lacking tRNA sequences altogether had been previously reported in Nostoc (Iteman et al. 2000) and Nodularia (Hayes and Barker 1997); we found similar regions in Calothrix and Scytonema.

Sequence data from rRNA-encoding genes (Giovannoni et al. 1988) provided early support for the theory that all plastids arose in an endosymbiotic event from a common cyanobacterial ancestor. All plastids except for those found in two parasitic plants have 16S–23S ITS regions containing both $tRNA^{Ile}$ and $tRNA^{Ala}$ (table 1). There are two rRNA operons in most cases; four taxa have one operon, and *Euglena gracilis* has three operons (table 1). In many algae and the plastids of the liverwort *Marchantia polymorpha* and nonparasitic angiosperms, the operons appear as inverted repeats. In all land plants, the operon has two unique features: an additional 4.5S rRNA molecule and large (650–1,050-bp) inserts in each of the two tRNA sequences in the 16S–23S ITS. The identification of similar inserts in the tRNA sequences of plastids from the charophycean alga Coleochaete led Manhart and Palmer (1990) to propose this alga to be the sister group of land plants.

What kind of 16S–23S ITS region might have existed in the rRNA operon of the cyanobacterial ancestor of chloroplasts? It seems likely that the order of the components of the rRNA operon(s) of the ancestor was 16S-$tRNA^{Ile}$-$tRNA^{Ala}$-23S-5S, with the two tRNA sequences separated by an ISR of only a few ($<$10) nucleotides. Evidence supporting this supposition is that this arrangement exists (1) in *Aquifex aeolicus* and *Thermotoga maritima,* the two oldest extant bacteria; (2) in most of the extant cyanobacteria for which we currently have data; (3) in the cyanelle of *Cyanophora paradoxa*; (4) in the chloroplasts of all algae in the Charophyta, Chlorophyta, Rhodophyta, and Bacillariophyta; and (5) in the chloroplast of *E. gracilis* (table 1). The oldest extant cyanobacteria are thought to be members of the genera Gloeobacter and Pseudanabaena (Turner 1997). No information is currently available on the number or configuration of rRNA operons of these organisms; such information could prove valuable in reconstructing the phylogeny of cyanobacteria and plastids.

Split rRNA Operons

The presence of split rRNA operons has been detected primarily when an entire genome has been sequenced. In most of the mitochondrial genomes for which sequences are available in GenBank, 16S- and 23S-encoding genes exist as separate, intact single copies. Boer and Gray (1988) reported a bizarre organization of rRNA genes in the mitochondrial DNA of *Chlamydomonas reinhardtii* where each gene is discontinuous and dispersed throughout the genome. More recently, split rRNA operons have also been identified in Eubacteria and Archaea (table 1). In the most common pattern, the 23S and 5S genes form an operon, and the 16S rRNA gene is separate; there may be one or more than one copy of each of the genes. In two Archaea (*Archaeoglobus fulgidus* and *Pyrococcus horikoshii* OT3), the 16S and 23S genes are linked and the 5S gene is separate. In the Eubacteria, *Helicobacter pylori* 26695 carries two 23S–5S operons, two single 16S genes, and one separate 5S gene. *Rickettsia prowazekii* Madrid E and *Wolbachia pipientis,* three species of Borrelia, and *Deinococcus radiadurans* carry one or more 5S–23S operons and one or more single 16S genes (table 1). Split rRNA operons are also present in the chloroplast of *T. gondii,* which carries single, unlinked 16S and 23S genes and no 5S gene. To date, the complete genome of only one cyanobacterium, Synechocystis PCC6803, has been determined (Kaneko et al. 1996); no split operons exist in this taxon. Perhaps as more complete genome sequences from this group of eubacteria are determined, more split operons will be identified.

Multiple Nonidentical rRNA Operons in Cyanobacteria

The number of rRNA operons present is known for relatively few bacterial species and ranges from 1 or 2 operons in the members of Archaea and several nonoxygenic eubacterial genera (Mycobacteria, Mycoplasma, Rhodothermus, and Thiobacillus) to 6–10 in some members of the proteobacteria and gram-positive eubacteria (table 1). Until recently, the presence of multiple rRNA operons in cyanobacteria has received little attention, so data relating to the number and sequence diversity of rRNA operons in cyanobacterial species are sparse.

The existence of multiple rRNA operons in cyanobacteria was first reported by Nichols et al. (1982), who used Southern hybridization to detect rRNA genes in three species of Anabaena and one species of Nostoc. Chen and Widger (1993) identified two operons on the physical map of Synechococcus strain PCC 7002. Two groups reported that Synechococcus PCC6301 (''*Anacystis nidulans*'') carries two rRNA operons, one of whose ITS regions contain both $tRNA^{Ile}$ and $tRNA^{Ala}$ (Tomioka, Shinozaki, and Sugiura 1981; Williamson and Doolittle 1983), and one of which contains a $tRNA^{Ile}$ pseudogene but has not been completely sequenced (Williamson and Doolittle 1983). Genome sequencing of Synechocystis (Kaneko et al. 1996) has shown two rRNA operons, both containing $tRNA^{Ile}$ only, present in inverted repeats. In this taxon, sequence data for the ISR regions flanking the tRNA are not available, so information about the sequence identity of the operons is unavailable at this time. In a brief note, Ligon et al. (1991) reported they had identified four rRNA operons in Nostoc (Anabaena) strain PCC 7120 but provided no methods or results to support their statement.

The presence of single rRNA operons or multiple identical operons in some cyanobacterial taxa was suggested by Scheldeman et al. (1999), who reported finding only one band of a consistent length when they amplified the 16S plus ITS region of multiple isolates of Arthrospira. Otsuka et al. (1999) recovered unambiguous ITS sequences from each of 47 clonal isolates of Microcystis without using a cloning step. Similarly, we sequenced the ITS regions from nine separate clones of the PCR product from two different isolates (SRS 6 and

SRS 70) of the putative new genus and always obtained sequences identical in structure (containing both $tRNA^{Ile}$ and $tRNA^{Ala}$, fig. 3, ii), length, and nucleotide sequence.

Evidence for multiple nonidentical rRNA operons has been presented by Wilmotte, Neefs, and De Wachter (1994), who described problems in obtaining unambiguous rDNA sequences using direct sequencing of ITS PCR products in Trichodesmium NIBB 1067. After cloning these PCR products, they discovered differences in ITS sequences (but not configurations) among clones. More recently, Iteman et al. (2000) found two ITS sequences with completely different configurations in Nostoc PCC 7120.

We identified, based on sequence analysis of individual clones of PCR-amplified ITS regions, two types of ITS regions in a unialgal isolate of *C. parietina* and three different types of ITS regions from a unialgal isolate of *S. hyalinum.* This means that *C. parietina* carries at least two rRNA operons and that *S. hyalinum* carries at least three different rRNA operons, two of which are identical in size (400 bp) but nonidentical in sequence (fig. 4). One significant difference between our results and the findings of Iteman et al. (2000) concerns the number and/or identity of the operons lacking tRNA. While the former investigators found sequence identity between the tRNA-lacking regions of all operons in a single species, we identified sequence differences between two 400-bp tRNA-lacking operons in *S. hyalinum.* We also identified rRNA operons of similar configurations but different sequence compositions in individual clones of Microcoleus and Nostoc (unpublished data).

Iteman et al. (2000) determined that both the tRNA-lacking and the tRNA-containing 16S–23S ITS regions of Nostoc PCC 7120 contain the sequences necessary for proper folding and processing of this region to release functional rRNA and tRNA. These regions included the D1, D1′, D2, and D3 short consensus sequences and the antiterminator box A. A box-B loop-and-stem structure just upstream of box A and a V3 stem-loop structure have variable primary structures but consistent secondary structures. We identified these same regions in the five species examined in this study (e.g., in Scytonema; fig. 4). Their presence suggests that all of the ITS regions we sequenced are involved in the folding and processing of rRNA.

The Use of 16S–23S ITS Sequence Data in Fine-Scale Phylogenetic Studies

The use of the 16S–23S ITS region in studies of phylogeny (reviewed Gürtler and Stanisich 1996; Lu 1999; Li 2000), molecular evolution (Antón, Martinez-Murcia, and Rodríguez-Valera 1998), or population genetics (Navarro et al. 1992; Otsuka et al. 1999; Scheldeman et al. 1999) is a potentially powerful tool. However, investigators need to be aware of potential problems imposed by the possibility of multiple nonidentical rRNA operons. Any study that depends on PCR amplification of the ITS region may run into the problem of preferential amplification of some operons. Amplification of cyanobacterial 16S–23S ITS regions depends on cyanobacterial-specific primers developed by Wilmotte, Van der Auwera, and De Wachter (1993), Nelissen et al. (1994), Wilmotte (1994), and Nübel, Garcia-Pichel, and Muyzer (1997). One of these primers sits at the 5′ end of the 23S rRNA molecule. Gürtler and Barrie (1995) found significant variation in this region of the 23S rRNA gene of the gram-positive bacterium *S. aureus* and cautioned that this primer site variability might mean that some rRNA operons might go unamplified or be underamplified in this species. In cyanobacteria, the existing data on 23S sequence composition are insufficient to ascertain whether a similar situation exists. Still, our experience with one field isolate of Tolypothrix suggests that such heterogeneity may be present. Attempts to amplify the ITS region of this organism with primers 1 and 5, which lie in the flanking 16S and 23S sequences (fig. 1), consistently produced one very bright band and one or two faint bands. When we sequenced three different clones of that PCR product, we observed only one sequence, which consistently contained no tRNA sequence (fig. 3, i). However, when we amplified the region with primers 1 and 4 or primers 3 and 5, primer sets that target $tRNA^{Ile}$ and $tRNA^{Ala}$ sequences (fig. 1), we obtained multiple nonidentical sequences containing both tRNA sequences. We interpret these results to indicate that the organism contains both tRNA-containing and tRNA-lacking rRNA operons and that the operon without tRNAs was preferentially amplified during PCR reactions using primers 1 and 5. This phenomenon may impact analyses employing restriction enzyme digests of ITS PCR products (e.g., Navarro et al. 1992; Lu et al. 1997; West and Adams 1997; Vinuesa et al. 1998; Scheldeman et al. 1999). That is, if the ITS from one operon is amplified preferentially in one isolate, while the ITS from a different operon is preferentially amplified in another isolate, comparisons of digests of PCR product from those two isolates may be flawed. Experimental designs should therefore contain safeguards against complications imposed by the presence of multiple nonidentical rRNA operons.

Acknowledgments

We thank Alison K. Dillon for developing our molecular protocols and Marisa DeNoble for helping with extractions and initial PCR. Jim Lissemore provided continual advice and support. Louise Lewis provided helpful criticism of early drafts of the manuscript. This work was supported in part by contract DACA88-98-Q-0181 from the U.S. Army Construction Engineering Research Laboratory in Champaign, Ill., and in part by funds from John Carroll University, including a Faculty Summer Research Fellowship to V.R.F. Publication costs were covered by a grant from the University Committee for Research, Service, and Faculty Development.

LITERATURE CITED

ANDERSON, S. G. E., A. ZOMORODIPOUR, J. A. ANDERSSON, T. SICHERITZ-PONTÉN, U. C. M. ALSMARK, R. M. PODOWSKI, A. K. NÄSLUND, A.-S. ERIKSSON, H. H. WINKLER, and C.

G. KURLAND. 1998. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. Nature **396**:133–139.

ANTÓN, A. I., A. J. MARTÍNEZ-MURCIA, and F. RODRÍGUEZ-VALERA. 1998. Sequence diversity in the 16S–23S intergenic spacer region (ISR) of the rRNA operons in representatives of the *Escherichia coli* ECOR collection. J. Mol. Evol. **47**:62–72.

BACOT, C. M., and R. H. REEVES. 1991. Novel tRNA gene organization in the 16S–23S intergenic spacer of the *Streptococcus pneumoniae* rRNA gene cluster. J. Bacteriol. **173**: 4234–4236.

BOER, P. H., and M. W. GRAY. 1988. Scrambled ribosomal RNA gene pieces in *Chlamydomonas reinhardtii* mitochondrial DNA. Cell **55**:399–411.

BULT, C. J., O. WHITE, G. J. OLSEN et al. (40 co-authors). 1996. Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. Science **273**:1058–1073.

CAMPBELL, K. D., A. K. EAST, D. E. THOMPSON, and M. D. COLLINS. 1993. Studies on the large subunit ribosomal RNA genes and intergenic spacer regions of non-proteolytic *Clostridium botulinum* types B, E, and F. Res. Microbiol. **144**: 171–180.

CARMICHAEL, W. W. 1986. Isolation, culture, and toxicity testing of toxic freshwater cyanobacteria (blue-green algae). Pp. 1249–1262 *in* V. SHILOV, ed. Fundamental research in homogenous catalysis. Vol. 3. Gordon and Breach, New York.

CHEN, X., and W. R. WIDGER. 1993. Physical genome map of the unicellular cyanobacterium *Synechococcus* sp. strain PCC 7002. J. Bacteriol. **175**:5106–5116.

CULLINGS, K. W. 1992. Design and testing of a plant-specific PCR primer for ecological and evolutionary studies. Mol. Ecol. **1**:233–240.

DECKERT, G., P. V. WARREN, T. GASSTERLAND et al. (15 co-authors). 1998. The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus.* Nature **392**:353–358.

DOUGLAS, S. E., and S. L. PENNY. 1999. The plastid genome of the cryptophyte alga, *Guillardia theta*: complete sequence and conserved synteny groups confirm its common ancestry with red algae. J. Mol. Evol. **48**:236–244.

DOYLE, J. J., and J. L. DOYLE. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. **19**:11–15.

DRYDEN, S. C., and S. KAPLAN. 1990. Localization and structural analysis of the ribosomal RNA operons of *Rhodobacter spaeroides.* Nucleic Acids Res. **18**:7267–7277.

EAST, A. K., D. ALLAWAY, and M. D. COLLINS. 1992. Analysis of DNA encoding 23S rRNA and 16S–23S rRNA intergenic spacer regions from *Plesiomonas shigellloides.* FEMS Microbiol. Lett. **74**:57–62.

EAST, A. K., and M. D. COLLINS. 1993. Molecular characterization of DNA encoding 23S rRNA and 16S–23S rRNA intergenic spacer regions from *Plesiomonoas shigelloies.* FEMA Microbiol. Lett. **74**:57–62.

EMOND, E., I. FLISS, and S. PANDIAN. 1993. A ribosomal DNA fragment of *Listeria monocytogenes* and its use as a genus-specific probe in an aqueous-phase hybridization assay. Appl. Environ. Microbiol. **59**:2690–2697.

FEINGOLD, J., V. BELLOFATTO, L. SHAPIRO, and K. AMEMIYA. 1985. Organization and nucleotide sequence analysis of an rRNA and tRNA gene cluster from *Caulobacter cresentus.* J. Bacteriol. **163**:155–166.

FLECHTNER, V. R. 1999. Enigmatic desert soil algae. Pp. 231–241 *in* J. SECKBACH, ed. Enigmatic microorganisms and life in extreme environments. Kluwer Academic Publishers, Amsterdam.

FLEISCHMANN, R. D., M. D. ADAMS, O. WHITE et al. (40 co-authors). 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science **269**:496–512.

FOX, G. E., J. D. WISOTZKEY, and P. JURTSHUK JR. 1992. How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. Int. J. Syst. Bacteriol. **42**:166–170.

FRASER, C. M., S. J. NORRIS, G. M. WEINSTOCK et al. (33 co-authors). 1998. Complete genome sequence of *Treponema pallidum,* the syphilis spirochete. Science **281**:375–388.

GARCÍA-MARTÍNEZ, J., A. J. MARTÍNEZ-MURCIA, A. I. ANTÓN, and F. RODRÍGUEZ-VALERA. 1996. Comparison of the small 16S to 23S intergenic spacer region (ISR) of the RNA operons of some *Escherichia coli* strains of the ECOR collection and *E. coli K12.* J. Bacteriol. **178**:6374–6377.

GILL, S., K. BELLES-ISLES, G. BROWN, S. GAGNE, C. LEMIEUX, J. P. MERCIER, and P. DION. 1994. Identification of variability of ribosomal DNA spacer from *Pseudomonas* soil isolates. Can. J. Microbiol. **40**:541–547.

GIOVANNONI, S. J., S. TURNER, G. J. OLSEN, S. BARNS, D. J. LANE, and N. R. PACE. 1988. Evolutionary relationships among cyanobacteria and green chloroplasts. J. Bacteriol. **170**:3584–3592.

GREEN, C. J., G. C. STEWART, M. A. HOLLIS, B. S. VOLD, and K. F. BOTT. 1985. Nucleotide sequence of the *Bacillus subtilis* ribosomal RNA operon, *rrnB.* Gene **37**:261–266.

GÜRTLER, V., and H. D. BARRIE. 1995. Typing of *Staphylococcus aureus* strains by PCR-amplification of variable length 16S–23S rDNA spacer regions: characterization of spacer sequences. Microbiology **140**:197–204.

GÜRTLER, V., and V. A. STANISICH. 1996. New approaches to typing and identification of bacteria using the 16S–23S rDNA spacer region. Microbiology **142**:3–16.

HALL, L. M. C. 1994. Are point mutations or DNA rearrangements responsible for the restriction fragment length polymorphisms that are used to type bacteria? Microbiology **140**:197–204.

HALLICK, R. B., L. HONG, R. G. DRAGER, M. R. FAVREAU, A. MONFORT, B. ORSAT, and A. SPIELMANN. 1993. Complete sequence of *Euglena gracilis* chloroplast DNA. Nucleic Acids Res. **21**:3537–3544.

HARASAWA, R., H. MIZUSAWA, K. NOZAWA, T. NAKAGA, K. ASADA, and I. KATO. 1993. Detection and tentative identification of dominant mycoplasma species in cell cultures by restriction analysis of he 16S–23S rRNA intergenic spacer regions. Res. Microbiol. **144**:489–493.

HARVEY, S., C. W. HILL, C. SQUIRES, and C. K. SQUIRES. 1988. Loss of the spacer loop sequence from the *rrnB* operon in the *Escherichia coli* K-12 subline that bears the *relA1* mutation. J. Bacteriol. **170**:1235–1238.

HAYES, P. K., and G. L. A. BARKER. 1997. Genetic diversity within Baltic Sea populations of *Nodularia* (cyanobacteria). J. Phycol. **33**:919–923.

HELMCHEN, T. A., D. BHATTACHARYA, and M. MELKONIAN. 1995. Analyses of ribosomal RNA sequences from glaucocystophyte cyanoelles provide new insights into the evolutionary relationships of plastids. J. Mol. Evol. **41**:203–210.

ITEMAN, I., R. RIPPKA, N. TANDEAU DE MARSAC, and M. HERDMAN. 2000. Comparison of conserved structural and regulatory domains within divergent 16S rRNA–23S rRNA spacer sequences of cyanobacteria. Microbiology **146**: 1275–1286.

KANEKO, T., S. SATO, H. KOTANI et al. (24 co-authors). 1996. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis sp.* strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. DNA Res. **3**:109–136.

KAWARABAYASI, Y., M. SAWADA, H. HORIKAWA et al. (31 co-authors). 1998. Complete sequence and gene organization of the genome of a hyper-thermophilic archaebacterium, *Pyrococcus horikoshii* OT3 (supplement). DNA Res. **5**: 147–155.

KIM, E., H. KIM, S. P. HONG, K. H. KANG, Y. H. KHO, and Y. H. PARK. 1993. Gene organization and primary structure of a ribosomal RNA gene cluster from *Streptomyces griseus* subsp. *griseus.* Gene **132**:21–31.

KJEMS, J., H. LEFFERS, R. A. GARRETT, G. WICH, W. LEINFELDER, and A. BÖCK. 1987. Gene organization, transcription signals and processing of the single ribosomal RNA operon of the archaebacterium *Thermoproteus tenax.* Nucleic Acids Res. **15**:4821–4835.

KLENK, H. P., R. A. CLAYTON, J. TOMB et al. (52 co-authors). 1997. The complete genome sequence of the hyperthermophilic, sulphate-reducing arachaeon *Archaeoglobus fulgidus.* Nature **390**:364–370.

KOCHI, T., H. SHIRAI, H. FUKUZAWA, T. SANO, T. KOMANO, K. UMESONO, H. INOKUCHI, H. OZEKI, and K. OHYAMA. 1988. Structure and organization of *Marchantia polymorpha* chloroplast genome. IV. Inverted repeat and small single copy regions. J. Mol. Biol. **203**:353–372.

KOHLER, S., C. F. DELWICHE, P. W. DENNY, L. G. TILNEY, P. WEBSTER, R. J. WILSON, and J. D. PALMER. 1997. A plastid of probable green algal origin in Apicomplexan parasites. Science **275**:1485–1489.

KOMÁREK, J., and K. ANAGNOSTIDIS. 1986. Modern approach to the classification system of cyanophytes. 2: Chroococcales. Algol. Stud. **43**:157–226.

———. 1989. Modern approach to the classification system of cyanophytes. 4: Nostocales. Algol. Stud. **56**:247–345.

KOWALLIK, K. V., B. STOEBE, I. SCHFFRAN, P. KROTH-PANCIC, and U. FREIER. 1995. The chloroplast genome of a chlorophyll a+c-containing alga, *Odontella sinensis.* Plant Mol. Biol. Rep. **13**:336–342.

LEBLANC, C., C. BOYEN, and S. LOISEAUX-DE GOER. 1995. Sequence and phylogeny of the plastidial 16S rRNA gene from the red alga *Chondrus crispus.* Eur. J. Phycol. **30**:133–140.

LEHNER, A. F., S. HARVEY, and C. W. HILL. 1984. Mapping and spacer identification of rRNA operons of *Salmonella typhimurium.* J. Bacteriol. **160**:682–686.

LEMIEUX, C., C. OTIS, and M. TURMEL. 2000. Ancestral chloroplast genome in *Mesostigma viride* reveals an early branch of green plant evolution. Nature **403**:649–652.

LI, X. 2000. Phenotypic and genotypic characterization of *Nostoc* species from 6 different sites in the Mojave desert. Master's thesis, John Carroll University, Cleveland, Ohio.

LIGON, P., K. G. MEYER, J. A. MARTIN, and S. E. CURTIS. 1991. Nucleotide sequence of a 16S rRNA gene from *Anabaena* sp. strain PCC 7120. Nucleic Acids Res. **19**:4553.

LOUGHNEY, K., E. LUND, and J. E. DAHLBERG. 1982. tRNA genes are found between 16S and 23S rRNA genes in *Bacillus subtilis.* Nucleic Acids Res. **10**:1607–1624.

LU, W. 1999. Application of molecular biological techniques to the identification of cyanobacteria. Master's thesis, Liverpool John Moores University, Liverpool, England.

LU, W., E. H. EVANS, S. K. MCCOLL, and V. A. SAUNDERS. 1997. Identification of cyanobacteria by polymorphisms of PCR-amplified ribosomal DNA spacer region. FEMS Microbiol. Let. **153**:141–149.

MAIER, R. M., K. NECKERMANN, G. IGLOI, and H. KOSSEL. 1995. Complete sequence of the maize chloroplast genome: gene content, hotspots of divergence and fine tuning of genetic information by transcript editing. J. Mol. Biol. **251**: 614–628.

MANHART, J. R., and J. D. PALMER. 1990. The gain of two chloroplast tRNA introns marks the green algal ancestors of land plants. Nature **345**:268–270.

MINNICK, M. F., J. C. STRANGE, and K. F. WILLIAMS. 1994. Characterization of the 16S23S rRNA intergenic spacer of *Bartonella bacilliformis.* Gene **143**:149–150.

NAKAGAWA, T., M. SHIMADA, H. MUKAI, K. ASADA, I. KATO, K. FUJINO, and T. SATO. 1994. Detection of alcohol-tolerant hiochi bacteria by PCR. Appl. Environ. Microbiol. **60**:637–640.

NAKAGAWA, T., T. UEOMORI, K. ASADA, I. KATO, and R. HARASAWA. 1992. *Achopeplasma laidlawii* has tRNA genes in the 16S–23S spacer of the rRNA operon. J. Bacteriol. **174**: 8163–8165.

NAVARRO, E., P. SIMONET, P. NORMAND, and R. BARDIN. 1992. Characterization of natural populations of *Nitrobacter* spp. using PCR/RFLP analysis of the ribosomal intergenic spacer. Arch. Microbiol. **157**:107–115.

NELISSEN, B., R. DE BAERE, A. WILMOTTE, and R. DE WACHTER. 1996. Phylogenetic relationships of nonaxenic filamentous cyanobacterial strains based on 16S rRNA sequence analysis. J. Mol. Evol. **42**:194–200.

NELISSEN, B., A. WILMOTTE, J.-M. NEEFS, and R. DE WACHTER. 1994. Phylogenetic relationships among filamentous helical cyanobacteria investigated on the basis of 16S ribosomal RNA gene sequence analysis. Syst. Appl. Microbiol. **17**:206–210.

NELSON, K. E., R. A. CLAYTON, S. R. GILL et al. (28 co-authors). 1999. Evidence of lateral transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima.* Nature **399**:323–329.

NICHOLS, J. M., I. J. FOULDS, D. H. CROUCH, and N. G. CARR. 1982. The diversity of cyanobacterial genomes with respect to ribosomal RNA cistrons. J. Gen. Microbiol. **128**:2739–2746.

NORMAND, P., B. COURNOYER, P. SIMONET, and S. NAZARET. 1992. Analysis of a ribosomal RNA operon in the actinomycete *Frankia.* Gene **111**:119–124.

NÜBEL, U., R. DE BAERE, A. WILMOTTE, and R. DE WACHTER. 1996. Phylogenetic relationships of nonaxenic filamentous cyanobacterial strains based on the 16S rRNA sequence analysis. J. Mol. Evol. **42**:194–200.

NÜBEL, U., F. GARCIA-PICHEL, and G. MUYZER. 1997. PCR primers to amplify 16S rRNA genes from cyanobacteria. Appl. Environ. Microbiol. **63**:3327–3332.

OJAIMI, C., B. E. DAVIDSON, I. SAINT GIRONS, and I. G. OLD. 1994. Conservation of gene arrangement and an unusual organization of rRNA genes in the linear chromosomes of the Lyme disease spirchaetes *Borrelia burgdorferi, B. garinii,* and *B. afzelii.* Microbiology **140**:2931–2940.

OLSEN, G. J., N. R. PACE, M. NUELL, B. P. KAINE, R. GUPTA, and C. R. WOESE. 1985. Sequence of the 16S rRNA gene from the thermoacidophilic archaebacterium *Sulfolobus solfataricus* and its evolutionary implications. J. Mol. Evol. **22**:301–307.

OTSUKA, S., S. SUDA, R. LI, M. WATANABE, H. OYAIZU, S. MATSUMOTO, and M. M. WATANABE. 1999. Phylogenetic relationships between toxic and non-toxic strains of the genus Microcystis based on 16S to 23S internal transcribed spacer sequence. FEMS Microbiol. Lett. **172**:15–21.

PERNODET, J.-L., F. BOCCARD, M.-T. ALEGRE, J. GAGNAT, and M. GUERINEAU. 1989. Organization and nucleotide se-

quence analysis of a ribosomal RNA gene cluster from *Streptomyces ambofaciens.* Gene **79**:33–46.

ROUX, V., and D. RAOULT. 1995. The 16S–23S rRNA intergenic spacer region of *Bartonella* (*Rochalimaea*) species is longer than usually described in other bacteria. Gene **156**: 107–111.

SAGREDO, B., E. JEDLICKI, and O. ORELLANA. 1992. Organization of the 16S–23S intergenic spacer region of the two rRNA operons from *Thiobaclilus ferooxidans.* Geomicrobiol. J. **10**:239–247.

SATO, S., Y. NAKAMURA, T. KANEKO, E. ASAMUZI, and S. TABATA. 1999. Complete structure of the chloroplast genome of *Arabidopsis thaliana.* DNA Res. **6**:283–290.

SCHELDEMAN, P. B., D. BAURAIN, R. BOUHY, M. SCOTT, M. MUHLING, B. A. WHITTON, A. BELAY, and A. WILMOTTE. 1999. *Arthrospira* ('*Spirulina*') strains from four continents are resolved into only two clusters, based on amplified ribosomal DNA restriction analysis of the internally transcribed spacer. FEMS Microbiol. Lett. **172**:213–222.

SECHI, L. A., and L. DANEO-MOORE. 1993. Characterization of intergenic spacers in two *rrn* operons of *Enterococcus hirae* ARCC 9790. J. Bacteriol. **175**:3213–3219.

SHIMADA, H., and M. SUGIURA. 1991. Fine structural features of the chloroplast genome: comparison of the sequenced chloroplast. Nucleic Acids Res. **19**:983–995.

SHINOZAKI, K., M. OHME, M. TANAKA et al. (23 co-authors). 1986. The complete nucleotide sequence of tobacco chloroplast genome: its gene organization and expression. EMBO J. **5**:2043–2049.

SIMPSON, A. J. G., F. C. REINARCH, P. ARRUDA et al. (116 coauthors). 2000. The genome sequence of the plant pathogen *Xylella fastidiosa.* Nature **406**:151–157.

SMITH, D. R., L. A. DOUCETTE-STAMM, C. DELOUGHERY et al. (38 co-authors). 1997. Complete genome sequence of *Methoanobacterium thernoautotrophicum* deltaH: functional analysis and comparative genomics. J. Bacteriol **179**:7135–7155.

SMITH, R. F., B. A. WIESE, M. K. WOJZYNSKI, D. B. DAVISON, and K. C. WORLEY. 1996. BCM Search Launcher—an integrated interface to molecular biology data base search and analysis services available on the World Wide Web. Genome Res. **6**:454–462.

SUZUKI, Y., Y. ONO, A. NAGATA, and T. YAMADA. 1988. Molecular cloning and characterization of an rRNA operon in *Streptomyces lividans* TK21. J. Bacteriol. **170**:1631–1636.

TETTELIN, H., N. J. SAUNDERS, J. HEIDELBERG et al. (42 coauthors). 2000. Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58. Science **287**:1809–1815.

THOMPSON, J. D., D. G. HIGGINS, and T. J. GIBSON. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**:4673–4680.

TOMB, J.-F., O. WHITE, A. R. KERLAVAGE et al. (34 co-authors). 1997. The complete genome sequence of the gastric pathogen *Helicobacter pylori.* Nature **388**:539–547.

TOMIOKA, N., K. SHINOZAKI, and M. SUGIURA. 1981. Molecular cloning and characterization of ribosomal RNA genes from a blue-green algae, *Anacystic nidulans.* Mol. Gen. Genet. **184**:359–363.

TOMIOKA, N., and M. SUGIURA. 1984. Nucleotide sequence of the 16S–23S spacer region in the *rrnA* operon from a blue-green alga, *Anacystis nidulans.* Mol. Gen. Genet. **193**:427–430.

TSUDZUKI, J., K. NAKASHIMA, T. TSUDZUKI, J. HIRATSUKA, M. SHIBATA, T. WAKASUGI, and M. SUGIURA. 1992. Chloroplast DNA of black pine retains a residual inverted repeat lacking rRNA genes: nucleotide sequences of trnQ, trnK, psbA, trnI, and trnH and the absence of rps16. Mol. Gen. Genet. **232**:206–214.

TURMEL, M., C. OTIS, and C. LEMIEUX. 1999. The complete chloroplast DNA sequence of the green alga *Nephroselmi olivacea*: insights into the architecture of ancestral chloroplast genomes. Proc. Natl. Acad. Sci. USA **96**:10248–10253.

TURNER, S. 1997. Molecular systematics of oxygenic photosynthetic bacteria. Plant Syst. Evol. **11**(Suppl.):13–52.

VAN DER GIESSEN, J. W. B., R. M. HARING, and B. A. M. VAN DER ZEIJST. 1994. Comparison of the 23S ribosomal RNA genes and the spacer region between the 16S and 23S rRNA genes of the closely related *Mycobacterium avium* and *Mycobacterium paratuberculosis* and the fast-growing *Mycobacterium phlei.* Microbiology **140**:1103–1108.

VINUESA, P., J. L. W. RADEMAKER, F. J. DE BRUIJN, and D. WERNER. 1998. Genotypic characterization of *Bradyrhizobium* strains nodulating endemic woody legumes of the Canary Islands by PCR-restriction fragment length polymorphism analysis of genes encoding 16S rRNA (16S rDNA) and 16S–23S rDNA intergenic spacers, repetitive extragenic palindromic PCR genomic fingerprinting, and partial 16S rDNA sequencing. Appl. Environ. Microbiol. **64**:2096–2104.

WAKASUGI, T., T. NAGAI, M. KAPOOR, M. SUGITA, M. ITO, S. ITO, J. TSUDZUKI, and K. NAKASHIMA. 1997. Complete nucleotide sequence of the chloroplast genome from the green alga *Chlorella vulgaris*: the existence of genes possibly involved in chloroplast division. Proc. Natl. Acad. Sci. USA **94**:5967–5972.

WEST, N. J., and D. G. ADAMS. 1997. Phyenotypic and genotypic comparison of symbiotic and free-living cyanobacteria from a single field site. Appl. Environ. Micriobiol. **63**: 4479–4484.

WHILEY, R. A., B. DUKE, J. M. HARDIE, and L. M. C. HALL. 1995. Heterogeneity among 16S–23S rRNA intergenic spacers of species within the '*Streptococcus milleri* group'. Microbiology **140**:1983–1987.

WHITE, O., J. A. EISEN, J. F. HEIDELBERG et al. (35 co-authors). 1999. Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1. Science **286**:1575–1577.

WILLIAMSON, S. E., and W. F. DOOLITTLE. 1983. Genes for $tRNA^{Ile}$ and $tRNA^{Ala}$ in the spacer between the 16S and 23S rRNA genes of a blue-green alga: strong homology to chloroplast tRNA genes and tRNA genes of the *E. coli rrnD* gene cluster. Nucleic Acids Res. **11**:225–235.

WILMOTTE, A. 1994. Molecular evolution and taxonomy of the cyanobacteria. Pp. 1–25 *in* D. A. BRYANT, ed. The molecular biology of cyanobacteria. Kluwer, Dordrecht, The Netherlands.

WILMOTTE, A., and S. GOLUBIC. 1991. Morphological and genetic criteria in the taxonomy of Cyanophyta/Cyanobacteria. Arch. Hydrobiol. Suppl. **64**:1–24.

WILMOTTE, A., J. M. NEEFS, and R. DE WACHTER. 1994. Evolutionary affiliation of the marine nitrogen-fixing cyanobacterium *Trichodesmium* sp strain NIBB 1067, derived by 16S ribosomal RNA sequence analysis. Microbiology **140**: 2195–2164.

WILMOTTE, A., S. TURNER, Y. VAN DER PEER, and N. R. PACE. 1992. Taxonomic study of marine Oscillatorian strains (Cyanobacteria) with narrow trichomes. II. Nucleotide sequence analysis of the 16S ribosomal RNA. J. Phycol. **28**:828–838.

WILMOTTE, A., C. VAN DER AUWERA, and R. DE WACHTER. 1993. Structure of the 16S ribosomal RNA of the thermophilic cyanobacteria *Chlorogloeopsis* HTF (*Mastigocladus*

laminosus HTF) strain PCC7518 and phylogenetic analysis. FEBS Lett. **317**:96–100.

WIMPEE, C. F., R. MORGAN, and R. L. WROBEL. 1992. Loss of transfer RNA genes from the plastid 16S–23S ribosomal RNA gene spacer in a parasitic plant. Curr. Genet. **21**:417–422.

WOESE, C. R., O. KANDLER, and M. L. WHEELIS. 1990. Toward a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc. Natl. Acad. Sci. USA **87**:4576–4579.

WOLFE, K. H., C. W. MORDEN, S. C. EMS, and J. D. PALMER. 1992. Rapid evolution of the plastid translational apparatus in a nonphotosynthetic plant: loss or accelerated sequence evolution of tRNA and ribosomal protein genes. J. Mol. Evol. **35**:304–317.

YOUNG, R. A., R. MACKLIS, and J. A. STEITZ. 1979. Sequence of the 16S–23S spacer region in two ribosomal RNA operons of *Escherichia coli*. J. Biol. Chem. **253**:3264–3271.

PAMELA SOLTIS, reviewing editor

Accepted February 13, 2001