JOURNAL OF BACTERIOLOGY, Aug. 1988, p. 3584–3592
0021-9193/88/083584-09$02.00/0

Vol. 170, No. 8

Evolutionary Relationships among Cyanobacteria and Green Chloroplasts

STEPHEN J. GIOVANNONI,† SEÁN TURNER, GARY J. OLSEN,‡ SUSAN BARNS,§ DAVID J. LANE,§ AND NORMAN R. PACE*

Department of Biology and Institute for Molecular and Cellular Biology, Indiana University, Bloomington, Indiana 47405

Received 25 January 1988/Accepted 20 May 1988

The 16S rRNAs from 29 cyanobacteria and the cyanelle of the phytoflagellate *Cyanophora paradoxa* were partially sequenced by a dideoxynucleotide-terminated, primer extension method. A least-squares distance matrix analysis was used to infer phylogenetic trees that include green chloroplasts (those of euglenoids, green algae, and higher plants). The results indicate that many diverse forms of cyanobacteria diverged within a short span of evolutionary distance. Evolutionary depth within the surveyed cyanobacteria is substantially less than that separating the major eubacterial taxa, as though cyanobacterial diversification occurred significantly after the appearance of the major eubacterial groups. Three of the five taxonomic sections defined by Rippka et al. (R. Rippka, J. Deruelles, J. B. Waterbury, M. Herdman, and R. Y. Stanier, J. Gen. Microbiol. 111:1–61, 1979) (sections II [pleurocapsalean], IV [heterocystous, filamentous, nonbranching], and V [heterocystous, filamentous, branching]) are phylogenetically coherent. However, the other two sections (I [unicellular] and III [nonheterocystous, filamentous]) are intermixed and hence are not natural groupings. Our results not only support the conclusion of previous workers that the cyanobacteria and green chloroplasts form a coherent phylogenetic group but also suggest that the chloroplast lineage, which includes the cyanelle of *C. paradoxa*, is not just a sister group to the free-living forms but rather is contained within the cyanobacterial radiation.

The cyanobacteria are one of the most morphologically diverse and conspicuously successful procaryotic groups. It is generally believed that the cyanobacteria were the first major group of phototrophs to arise with a two-stage photosynthetic pathway capable of oxidizing water to produce molecular oxygen. Geochemical and fossil evidence indicates that in the Precambrian Era they caused the transition in the Earth's atmosphere from its primordial, anaerobic state to its current, aerobic condition (19, 36, 43). Moreover, molecular phylogenetic analysis of *c*-type cytochrome and rRNA sequences have established a relationship between cyanobacteria and the green (euglenoids, green algae, and higher plants) and red (rhodophyte) chloroplasts, thus supporting the procaryotic origins of chloroplasts.

Because of their ubiquity, rRNA sequences are particularly useful for establishing evolutionary relationships among diverse organisms. Woese and colleagues, using partial (RNase T_1-generated oligonucleotide catalogs) and complete 16S rRNA sequences, have defined about 10 major divisions (phyla) of eubacteria (45). The cyanobacteria are one of these phyla. However, too few strains (eight) of cyanobacteria had been investigated to develop a comprehensive overview of the diversity of the group.

We have used a method for directly sequencing 16S rRNA to explore the evolutionary relationships among 30 representatives of the diverse cyanobacterial groups, including the photosynthetic organelle of the phytoflagellate *Cyanophora paradoxa*. The results shed new light on the relative ages of the cyanobacteria and other eubacterial lineages and on the origins of chloroplasts.

MATERIALS AND METHODS

Organisms and cultivation. Cultures or cell paste of all Pasteur Culture Collection strains were provided by John B. Waterbury (Woods Hole Oceanographic Institution). All Castenholz Culture Collection strains were supplied as cultures by Richard H. Castenholz (University of Oregon). Cultures of *Oscillatoria limnetica* (Solar Lake) and *Microcoleus* sp. strain 10 mfx were from Yehuda Cohen (H. Steinitz Marine Biology Laboratory, Eilat, Israel). Cultures were maintained in BG-11 medium (34). *C. paradoxa* cyanelle RNA, provided by Stuart Maxwell (North Carolina State University) and Jessup M. Shively (Clemson University), was prepared as described by Starnes et al. (40).

Preparation of RNA for sequencing. High-molecular-weight cellular RNAs were prepared for sequencing as previously described (25). Briefly, cell pellets (0.3 to 3.0 g [wet weight]) were removed from storage at −70°C, thawed, and lysed by one or two passages through a French pressure cell. RNA was purified by extraction with phenol, precipitated with ethanol, and suspended. Overnight precipitation from 1.0 M NaCl at 0°C was used to prepare high-molecular-weight RNA.

RNA sequencing. Dideoxynucleotide-terminated sequencing, using reverse transcriptase and synthetic oligodeoxynucleotide primers complementary to conserved 16S rRNA sequences, was carried out as described by Lane et al. (25; D. J. Lane, K. G. Field, G. J. Olsen, and N. R. Pace, Methods Enzymol., in press). Three "universal" small subunit rRNA sequencing primers (complementary to *Escherichia coli* 16S rRNA sequence positions 519 to 536, 907 to 926, and 1392 to 1406) were used to determine ca. 1,000 nucleotides of each rRNA sequence. The sequences have

* Corresponding author.
† Present address: Department of Microbiology, Oregon State University, Corvallis, OR 97331.
‡ Present address: Department of Microbiology, University of Illinois, Urbana, IL 61801.
§ Present address: Gene-Trak Systems, Framingham, MA 01701.

been deposited with GenBank and are available from the authors upon request.

Sequence analysis. Sequences were manually aligned on the basis of conserved sequence and secondary structural elements. Regions of ambiguous sequence alignment were omitted from subsequent analyses. For each pair of sequences, the number of nucleotide differences was used to estimate the average number of fixed point mutations per sequence position that has accumulated since their divergence (22). This is called the "evolutionary distance" separating the contemporary sequences (Fig. 1 and 2, lower left). The statistical uncertainty of the pairwise evolutionary distance estimates was determined by the method of Kimura and Ohta (23) (Fig. 1 and 2, upper right). A least-squares method was used to infer the phylogenetic tree most consistent with the pairwise distance estimates and their statistical uncertainties (30).

RESULTS AND DISCUSSION

Relationships among free-living cyanobacteria. The unrooted phylogenetic tree in Fig. 3 broadly depicts evolutionary relationships of oxygenic, phototrophic bacteria and organelles to other representatives of the three primary lines of descent: the archaebacteria, the eucaryotes, and the eubacteria. The diversity among cyanobacteria and chloroplasts is indicated by the depth of the hatching. The cyanobacteria and chloroplasts together make up 1 of approximately 10 major eubacterial taxa (not all of which are represented in this tree). The rRNA sequence diversity within the cyanobacterial lineage is substantially less than the separations between the major eubacterial taxa. In contrast, the sequence diversity observed within most other eubacterial rRNA phyla is nearly equal to interphylum evolutionary distances (45). Interphylum evolutionary distances for representative eubacteria vary from 0.19 to 0.30 fixed mutations per sequence position (Fig. 1). Intraphylum evolutionary distances among the cyanobacteria are typically much smaller, about 0.15 mutations per sequence position. Thus, relatively close phylogenetic relationships underlie the extraordinary morphological diversity of cyanobacteria.

The phylogenetic tree in Fig. 4 depicts inferred relationships among the nonheterocystous cyanobacteria inspected. A diverse representation of cyanobacteria was selected for sequencing on the basis of available taxonomic, morphological, and physiological information. Many of the lineages branch at similar depths in the cyanobacterial tree, lending the tree a fan-like aspect. This result indicates that many modern cyanobacterial lineages arose in an expansive evolutionary radiation. The nearly equivalent depths of many of these branchings render their detailed relationships (branching orders) unresolved. This is because the uncertainty in the evolutionary distances separating pairs of organisms increases with the distance, a consequence of uncertainty in the number of multiple mutations of individual residues. The statistical errors for the calculated evolutionary distances separating pairs of organisms are included in Fig. 2. The relative positions of the nodes in the tree are more accurate than suggested by Fig. 2, because the uncertainties are not independent and some of the uncertainty applies only to the lengths of peripheral branches, not to the lengths of the branchings deeper in the tree. Although the precise branching order of the short segments near the root of the cyanobacterial tree is uncertain, significant relationships emerge.

The morphological complexity of cyanobacteria served as a primary basis of the classification system of Rippka et al. (34), which divides the cyanobacteria into five sections. The unicellular cyanobacteria constitute section I. Members of section II, the pleurocapsalean cyanobacteria, share a common developmental feature, the capacity of large cells to subdivide internally, producing numerous smaller cells (baeocytes). Section III incorporates the filamentous, nonheterocystous organisms. Sections IV and V are, respectively, the nonbranching and branching filamentous forms that are capable of forming heterocysts (specialized, nitrogen-fixing cells).

The deepest branches so far in the cyanobacterial tree are represented by *Gloeobacter violaceus* (section I) and two closely related strains of the genus *Pseudanabaena* (section III) (Fig. 4). *G. violaceus* is unique among cyanobacteria because of its lack of thylakoids and the unusual structure of its phycobilisomes. The photosynthetic reaction centers are contained in the cytoplasmic membrane with the phycobilisomes arranged in an underlying cortical layer (17). The early divergence of this organism from the main cyanobacterial lineage is consistent with this remarkably different light-harvesting apparatus. The result suggests that the common ancestor of the modern cyanobacteria may have lacked thylakoids. Unlike *G. violaceus*, the *Pseudanabaena* strains inspected have no obvious characteristics to suggest an early evolutionary divergence from the main cyanobacterial line. Biochemical and ultrastructural studies of these strains may reveal additional features that distinguish them from the other cyanobacteria and may clarify the phylogenetic relations among them (15, 16).

The unicellular cyanobacteria of section I and the nonheterocystous, filamentous organisms of section III are dispersed throughout the tree (Fig. 4), indicating that these morphotypes have multiple evolutionary origins. Thus, taxonomic classifications based principally on morphology do not necessarily reflect phylogenetic relationships.

The orientation of cell division planes has been used as a further basis for subdividing the unicellular cyanobacteria into several genera. An earlier study of RNase T_1-generated 16S rRNA oligonucleotide catalogs appeared to support this definition of natural relationships (5). However, the additional sequence data gathered during the present study led us to conclude that the various modes of cell division are not well correlated with phylogenetic groups. An example in Fig. 4 is the local cluster that contains *Gloeothece* sp. strain PCC 6501 (cell division in one plane), *Synechocystis* sp. strain PCC 6308 (cell division in two planes), and *Synechocystis* (*Eucapsis*) sp. strain PCC 6906 and *Gloeocapsa* sp. strain PCC 73106 (cell division in three planes). This cluster also contains the section III filamentous cyanobacterium *Spirulina* sp. strain PCC 6313 (cell division in one plane). Other cellular properties are consistent with the conclusion that division planes do not define phylogenetic groups. For instance, the DNA base compositions of unicellular cyanobacteria that divide in one plane span a broad range (35 to 71 mol% G+C) (20). Phylogenetic studies that include additional genera are likely to uncover still greater diversity among unicellular cyanobacteria.

Even a conspicuous feature such as the oscillatorian morphotype may not indicate a relatedness group, to the exclusion of other organisms. Although some clustering of *Oscillatoria* spp. is indicated (e.g., *O. limnetica* and PCC 7515), other *Oscillatoria* spp. are scattered throughout the tree (e.g., *O. amphigranulata*). Some physiological traits are consistent with the rRNA sequence diversity seen in this group. For instance, both *O. limnetica* and *O. amphigranulata* are able to use sulfide, as well as water, as a source of

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1 Gloeobacter PCC 7421	---	15	15	17	18	16	16	16	16	15	16	17	17	17	18	21	23	21	21	21	20	24	32
2 Synechococcus PCC 6301	133	---	14	13	15	15	14	14	15	13	15	16	18	17	17	20	22	22	22	20	20	23	32
3 Oscillatoria "amphigr."	127	107	---	14	16	15	14	13	14	14	16	17	18	17	19	20	22	21	23	21	21	24	33
4 Phormidium PCC 7375	155	103	115	---	14	15	13	14	16	14	15	17	17	17	18	19	23	22	22	21	21	25	33
5 Synechocystis PCC 6308	175	140	142	121	---	15	14	13	17	14	15	18	19	18	18	20	25	22	22	21	22	24	34
6 Gloeothece PCC 6501	146	132	120	134	129	---	12	13	14	14	15	16	17	17	17	20	22	22	23	21	22	23	34
7 Gloeocapsa PCC 73106	146	124	113	104	116	89	---	13	15	14	16	16	17	16	18	20	22	21	21	19	20	24	32
8 Myxosarcina PCC 7312	142	119	104	113	105	96	103	---	15	14	14	16	17	17	17	19	23	21	22	21	21	24	32
9 Oscillatoria PCC 7515	147	135	118	144	156	110	137	132	---	15	15	16	17	16	19	21	24	22	23	22	21	24	35
10 Anabaena PCC 7122	130	104	120	112	120	117	110	117	127	---	13	15	17	16	17	20	24	22	22	21	21	24	32
11 Cyanophora cyanelle	147	141	147	135	127	128	141	119	131	110	---	14	16	15	16	18	24	20	23	22	22	24	32
12 liverwort chloroplast	155	157	165	165	177	143	149	147	146	131	121	---	9	8	15	18	22	22	23	22	21	24	33
13 maize chloroplast	163	180	172	167	191	171	157	171	163	161	143	55	---	7	17	19	24	23	24	23	23	25	34
14 tobacco chloroplast	158	165	159	160	183	158	148	158	147	143	127	38	33	---	16	18	23	22	24	22	22	24	34
15 Chlamydomonas chloropl.	182	167	185	180	188	171	175	169	201	166	149	134	162	149	---	18	22	22	24	22	23	25	33
16 Euglena chloroplast	226	209	206	206	214	222	208	191	226	218	178	177	200	189	175	---	25	25	27	24	24	29	34
17 Escherichia coli	257	257	248	263	296	257	257	267	283	280	293	259	294	273	251	312	---	18	20	19	21	24	36
18 Pseudomonas testosteroni	235	247	229	242	251	247	231	226	251	245	223	247	264	246	259	305	190	---	21	20	20	25	34
19 Desulfovibrio desulfur.	237	245	260	242	251	274	227	251	268	245	274	276	294	281	293	346	215	235	---	19	19	24	36
20 Agrobacterium tumefaci.	228	219	220	226	241	241	205	226	255	233	247	259	276	254	261	286	201	217	203	---	19	24	35
21 Bacillus subtilis	220	216	222	233	246	254	222	235	240	228	250	242	263	249	267	294	238	224	195	199	---	22	32
22 Thermomicrobium roseum	286	273	286	298	295	264	294	287	290	292	292	298	307	291	314	378	296	300	285	281	259	---	37
23 Methanococcus vannielii	418	428	441	443	459	453	420	421	468	428	434	451	468	469	449	470	492	461	498	486	429	510	---

FIG. 1. Estimated evolutionary distance (lower left) and corresponding statistical uncertainty (upper right) separating pairs of representative cyanobacterial, chloroplast, and bacterial 16S rRNA sequences (×1,000). The evolutionary distances (22) and uncertainties (23) are estimated from the sequence differences within regions (totaling about 670 nucleotides) for which sequence data were available from all the strains and the sequences could be unambiguously aligned. For example, the *Gloeobacter* sp. strain PCC 7421 and *Synechococcus* sp. strain PCC 6301 (*Anacystis nidulans*) sequences are estimated to be separated by 0.133 ± 0.015 fixed point mutations per sequence position. Sources of published 16S rRNA sequences are as follows: *Synechococcus* sp. strain PCC 6301 (*Anacystis nidulans*) (42); liverwort (*M. polymorpha*) chloroplast (29); maize (*Z. mays*) chloroplast (38); tobacco (*N. tabacum*) chloroplast (41); *Chlamydomonas reinhardii* chloroplast (chloropl.) (11); *Euglena gracilis* chloroplast (13); *E. coli* (6); *Pseudomonas testosteroni* (46); *Desulfovibrio desulfuricans* (desulfur.) (32); *A. tumefaciens* (tumefaci.) (46); *Bacillus subtilis* (14); *Thermomicrobium roseum* (31); and *Methanococcus vannielii* (21). Oscillatoria "amphigr.," *Oscillatoria amphigranulata*.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
1	Pseudanabaena PCC 6903	---	7	15	14	14	15	15	15	15	16	15	17	16	15	15	15	16	16	15	17	17	15	17	15	16	14	15	17	15	16	15
2	Pseudanabaena "galeata"	32	---	16	15	15	16	14	15	15	16	15	17	17	15	15	15	17	17	15	17	16	16	17	15	15	14	15	17	15	15	15
3	Gloeobacter PCC 7421	129	142	---	15	15	15	17	14	16	17	17	18	16	15	16	15	18	16	16	17	17	16	16	14	17	16	15	17	16	15	15
4	Synechococcus "lividus"	114	130	119	---	11	13	14	13	13	15	14	16	15	14	15	13	15	13	14	16	15	14	15	12	15	14	13	15	14	14	13
5	Synechococcus PCC 6301	118	134	133	79	---	14	14	14	15	13	14	15	15	14	14	14	15	14	15	16	15	14	15	13	15	14	14	15	15	14	13
6	Oscillatoria "amphigr."	129	141	127	93	107	---	12	14	13	14	14	16	15	13	14	13	15	13	16	17	17	15	14	13	16	15	15	16	16	15	14
7	Plectonema PCC 73110	126	112	156	114	121	86	---	15	13	14	13	17	16	14	15	14	15	14	17	18	17	16	17	15	16	15	15	15	15	14	14
8	Oscillatoria PCC 6304	123	131	119	98	112	107	128	---	13	14	14	15	14	13	14	11	14	13	15	16	15	14	13	11	14	14	13	14	14	14	14
9	Chamaesiphon PCC 7430	130	126	144	100	126	96	105	102	---	13	13	14	14	13	14	11	15	13	14	16	16	14	15	12	13	13	13	14	14	13	13
10	Phormidium PCC 7375	139	145	155	123	103	115	122	115	101	---	14	14	15	14	13	12	15	14	15	17	16	14	16	13	15	15	14	15	15	14	14
11	Synechocystis PCC 6906	127	129	163	119	115	113	107	123	105	120	---	13	12	12	12	12	14	13	15	17	15	14	15	13	15	14	14	14	14	14	13
12	Synechocystis PCC 6308	155	160	175	138	140	142	157	128	114	121	102	---	15	14	14	14	15	13	15	18	17	16	17	14	16	15	15	16	14	15	14
13	Gloeothece PCC 6501	141	163	146	121	132	120	140	113	113	134	91	129	---	11	12	12	14	13	15	16	16	13	14	13	15	14	13	15	14	15	14
14	Spirulina PCC 6313	122	134	125	107	116	101	117	105	101	110	89	113	75	---	13	11	14	12	14	15	14	13	15	12	14	13	12	14	13	14	12
15	Gloeocapsa PCC 7321	136	136	146	122	124	113	126	110	110	104	90	116	89	96	---	10	13	13	16	16	15	15	15	12	16	15	15	15	15	13	14
16	Dermocarpa PCC 7437	122	124	134	98	112	96	111	79	75	86	91	113	82	68	66	---	11	10	15	15	14	12	14	12	12	12	12	13	14	12	13
17	Pleurocapsa PCC 7321	144	164	167	119	125	118	132	108	120	129	112	125	117	106	102	72	---	9	16	18	17	15	17	14	15	14	14	15	15	15	14
18	Myxosarcina PCC 7312	143	159	142	101	119	104	117	99	101	113	98	105	96	87	103	59	50	---	14	16	16	13	15	13	15	14	14	15	15	14	14
19	Cyanophora cyanelle	130	137	147	118	141	147	161	134	120	135	125	127	128	118	141	124	146	119	---	14	16	14	15	13	15	14	14	16	14	15	13
20	liverwort chloroplast	161	163	155	148	157	165	178	149	148	165	163	177	143	130	149	125	168	147	121	---	17	15	16	15	17	16	15	17	16	16	15
21	Oscillatoria "limnetica"	154	149	153	131	134	156	159	134	139	141	134	153	143	113	127	111	161	137	144	156	---	14	15	13	16	16	15	17	16	16	15
22	Microcoleus 10 mfx	133	139	141	116	118	122	140	119	114	113	114	144	96	96	121	84	124	105	115	128	110	---	14	12	15	15	14	16	16	16	15
23	Oscillatoria PCC 7515	157	163	147	123	135	118	160	105	120	144	132	156	110	124	137	118	149	132	131	146	135	111	---	12	15	15	14	16	15	15	15
24	Lyngbya PCC 7419	129	129	112	90	102	101	127	79	85	101	102	114	103	82	88	81	111	99	102	126	102	86	92	---	14	12	12	14	13	13	12
25	Scytonema PCC 7110	146	135	161	121	135	139	146	112	105	135	137	150	135	109	140	90	130	125	132	160	151	127	136	108	---	11	12	13	12	13	12
26	Fischerella PCC 7414	122	120	142	116	124	129	125	113	105	128	118	127	113	97	124	85	113	112	111	142	141	123	128	92	72	---	9	12	10	12	10
27	Chlorogloeopsis PCC 6718	124	128	134	104	118	127	132	97	95	119	114	127	100	85	126	85	116	114	113	137	133	111	113	82	84	52	---	13	10	11	10
28	Calothrix PCC 7102	161	156	158	128	137	142	137	114	110	129	123	150	124	109	133	98	121	126	154	160	167	138	138	112	96	83	96	---	12	11	11
29	Nostoc PCC 73102	130	135	144	115	125	137	138	109	115	134	118	121	123	99	125	108	133	134	116	146	152	137	126	99	91	63	65	86	---	10	8
30	Nodularia PCC 73104	142	137	138	115	120	131	115	115	106	122	120	131	130	108	106	90	122	123	133	148	141	143	135	98	98	86	79	81	67	---	8
31	Anabaena PCC 7122	131	133	130	104	104	120	120	110	105	112	109	120	117	85	110	96	118	117	110	131	133	122	127	83	90	68	60	75	46	37	---

FIG. 2. Estimated evolutionary distance (lower left) and corresponding statistical uncertainty (upper right) separating each pair of cyanobacterial rRNA sequences (×1,000). Derivation of the values and sources of sequences are as described in the legend to Fig. 1. Oscillatoria "amphigr.," *Oscillatoria amphigranulata*.

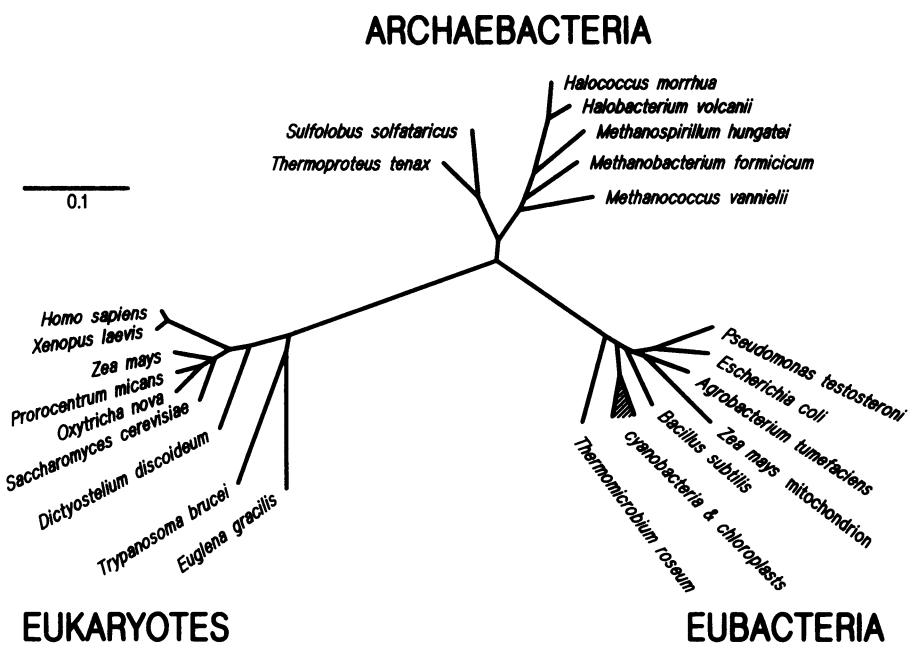

FIG. 3. Unrooted phylogenetic tree illustrating evolutionary relationships among diverse 16S-like rRNAs (adapted from data in reference 30). The hatched area depicts evolutionary diversity among oxygenic, phototrophic bacteria and organelles. Segment lengths are proportional to evolutionary distances. The scale bar corresponds to 0.1 fixed point mutations per sequence position. This topology is based on an analysis of 920 nucleotides from complete sequences.

electrons for noncyclic photosynthesis. However, the details of sulfide adaptation in these two strains differ (8, 9), and they are not close relatives (Fig. 4). The LPP subgroup of filamentous cyanobacteria (section III) was provisionally created to encompass organisms which did not fit strictly within the morphological definitions of the genera *Oscillatoria*, *Pseudanabaena*, and *Spirulina* (34). Here the group is represented by *Lyngbya* sp. strain PCC 7419, *Plectonema* sp. strain PCC 73110, and *Phormidium* sp. strain PCC 7375, which do not cluster phylogenetically (Fig. 4). A *Lyngbya* sp. (LPP group A; sheathed, nonmotile except for hormogonia) branches deeply but appears to be related to a phylogenetic subgroup containing other filamentous genera, including two *Oscillatoria* spp. and a *Microcoleous* sp.

The pleurocapsalean organisms (section II) studied so far constitute a coherent, phylogenetic subgroup, to the exclusion of other cyanobacteria (Fig. 4). Although only three of six genera of this section have been analyzed, all the strains in the section have very similar DNA base compositions (16). This and their common mode of reproduction distinguish them from all other cyanobacteria, suggesting that reproduction by multiple fission (baeocyte formation) is of monophyletic origin.

Similarly, the members of sections IV and V, the heterocystous, filamentous forms, constitute a distinct phylogenetic group. The organisms of section V are a subgroup arising from within the lines of section IV (Fig. 5). This confirms similar conclusions drawn from DNA-DNA hybridization data (24). The organisms of sections IV and V are among the most morphologically diverse of the cyanobacteria. Aside from heterocysts, many strains also produce akinetes, "resting" cells that develop as a result of insufficient light or other factors (35). The strains shown in Fig. 5 that produce hormogonia as part of their developmental cycle (*Scytonema* sp. strain PCC 7110, *Chlorogloeopsis* sp. strain PCC 6718, *Fischerella* sp. strain PCC 7414, and *Calothrix* sp. strain PCC 7102) branch more deeply in the heterocystous forms than do strains of those genera that do not produce hormogonia (*Anabaena* sp. strain PCC 7122 and *Nodularia* sp. strain PCC 73104).

Origin of photosynthetic organelles. The present sequence comparisons confirm the conclusion of Bonen et al. (3–5), based on RNase T_1-generated oligonucleotide catalogs, that the cyanobacteria and green chloroplasts form a natural group (Fig. 6). However, in contrast to the earlier cluster analysis of oligonucleotide catalogs, which indicated that the progenitor of the green chloroplasts diverged before the radiation that gave rise to the diversity seen in modern free-living cyanobacteria (3–5), the present analysis of continuous sequences included the green chloroplasts well within the cyanobacterial radiation. That is, the green chloroplast line of descent should be viewed as one of the cyanobacterial sublines. While the present study included a greater diversity of cyanobacteria than previous analyses, this alone does not explain the different placement of the green chloroplasts. The interpretation of the origin of chloroplasts on the basis of their 16S rRNA sequences is complicated by their large amount of evolutionary divergence relative to the other cyanobacterial sublines. Because of undercompensation for multiple mutations, rapidly evolving lineages can branch spuriously deeply in inferred phylogenetic tree topologies (30). Green chloroplast 16S rRNA sequences from the following organisms were analyzed: *Zea mays* (maize), *Nicotiana tabacum* (tobacco), *Marchantia polymorpha* (liverwort), *Chlamydomonas reinhardii* (green alga), and *Euglena gracilis* (euglenoid). The analysis indicated that these sequences are specifically related. The sequence from *M. polymorpha* showed the lowest average rate of sequence change for the group and hence was chosen for comparison in Fig. 6. Cluster analysis, as previously used to establish the close relationship of chloroplasts and cyanobacteria, is particularly susceptible to artifacts in branching order that arise from different rates of change among compared sequences (10). The least-squares distance matrix

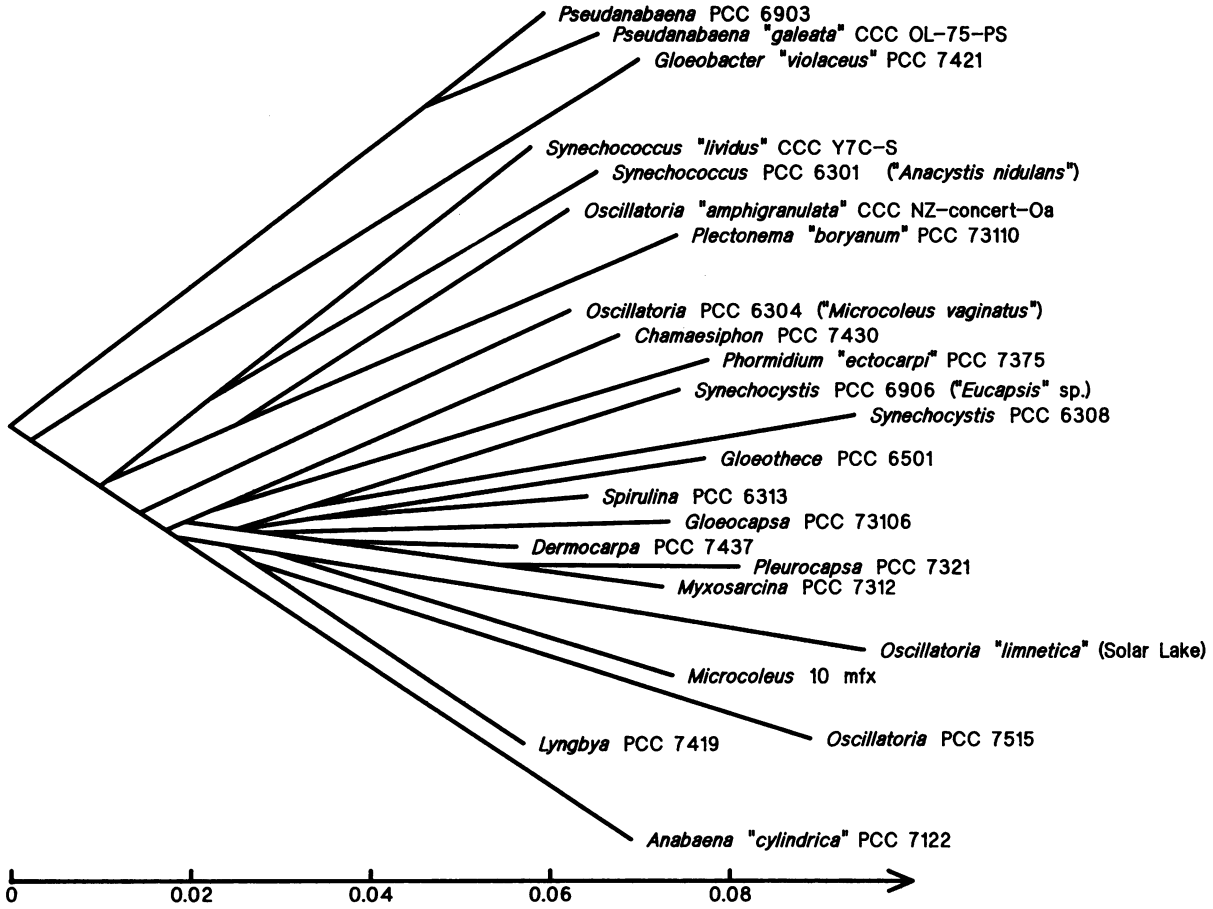

FIG. 4. Rooted-tree topology illustrating evolutionary relationships among 16S rRNAs from cyanobacteria. Evolutionary distances are proportional to the horizontal component of segment length in this representation. *A. tumefaciens*, *B. subtilis*, and *P. testosteroni* 16S rRNA sequences were used to locate the root. Only one heterocystous cyanobacterium, *Anabaena* sp. strain PCC 7122, is included in this tree. The scale is in units of fixed point mutations per sequence position.

method used here is less sensitive to this problem, and it includes the chloroplast ancestry within the cyanobacterial radiation. Thus, it seems necessary to include the chloroplasts, if the cyanobacteria are to be considered a holophyletic group, i.e., a group consisting of a common ancestor and all lineages derived from it (1).

In many respects the photosynthetic organelle (cyanelle) of the flagellate *C. paradoxa* resembles a cyanobacterium more than a green chloroplast: it contains phycobilin antenna pigments, has a rudimentary peptidoglycan wall, lacks chlorophyll *b*, and lacks the double membrane and thylakoid arrangement of chloroplasts. Thus, it is frequently supposed that cyanelles arose independently of the chloroplasts as a relatively recent endosymbiosis between a eucaryote and a cyanobacterium (18). Our results are consistent with such a view but nevertheless indicate a specific relationship between the *C. paradoxa* cyanelle and the green chloroplasts (Fig. 6) (28). We offer two alternative explanations for this relationship. (i) The common ancestor of the cyanelle and green chloroplast lineages may have been a free-living cyanobacterium that independently initiated the two endosymbioses. In this case, the chlorophyll *b* light-harvesting mechanism would have arisen in the chloroplast lineage, in either the free-living or symbiotic state. (ii) Both lineages may derive from a single endosymbiotic event, in which case chlorophyll *b* would have originated in the chloroplast progenitor during its endosymbiotic state. The former theory might suggest that certain procaryotic lines are predisposed toward symbioses, a possibility that finds a parallel in the lineage that contains mitochondria, plant parasites such as *Agrobacterium tumefaciens* (46), and a rickettsia (44). The latter possibility implies multiple, independent origins for the chlorophyll *b* light-harvesting mechanism, which is also known to exist in symbiotic and free-living procaryotes (prochlorophytes) (7, 27).

Timing of events in cyanobacterial evolution. Although the cyanobacteria often are cited as a particularly ancient group, the sequence similarities of their rRNAs to one another and to those of other eubacteria show that the other major eubacterial taxa diverged significantly before the diversification of the modern cyanobacteria. Among these other eubacterial taxa are the family *Chloroflexaceae* (Fig. 3, *Thermomicrobium*), which diverge from the main eubacterial lineage substantially more deeply than do the cyanobacteria (31). Obligately anaerobic, phototrophic *Chloroflexus* spp. are known to form laminated microbial mats and are morphologically similar to microfossils in the earliest known stromatolites (2, 12). These considerations caution against the interpretation of the earliest microbial fossils as cyanobacterial in origin (37). Interpretations of microfossil evidence are frequently based upon the assumption that morphology is phylogenetically conserved. As seen in the extant

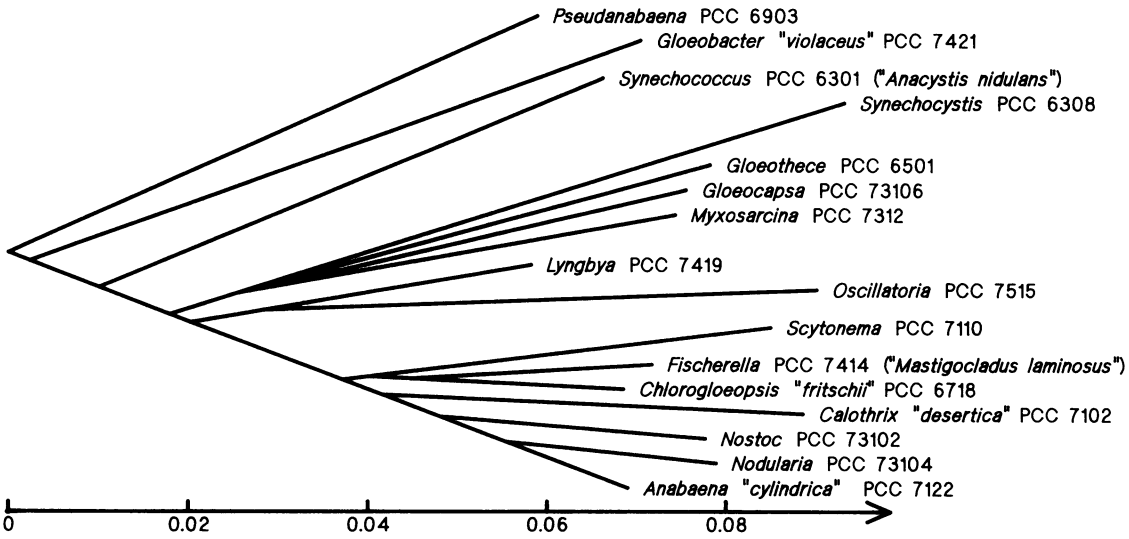

FIG. 5. Rooted-tree topology illustrating evolutionary relationships among 16S rRNAs from heterocystous cyanobacteria. The presentation is as described in the legend to Fig. 4.

procaryotes, morphology correlates imperfectly with phylogeny, suggesting that convergent evolution of morphological characteristics among procaryotes is a common theme (33, 39).

Because the 16S rRNAs of different organisms accumulate mutations at different rates, evolutionary distances cannot be accurately calibrated in terms of time. Hence, we do not attempt to infer the length of time that passed between the divergence of the major eubacterial phyla and the flowering of cyanobacterial diversity. It is clear, however, that cyanobacterial diversification occurred within a relatively short span of molecular evolutionary distance. Although the events responsible for this apparent burst of evolution in the cyanobacterial line of descent are uncertain, as the first organism to exploit water as an electron donor for photosynthesis, the common ancestor of the cyanobacteria and chloroplasts had available a novel and profoundly fertile physiological niche. We suggest that rapid diversification ensued, leaving its record in the molecular relationships observed here.

The molecular phylogenetic data suggest that the heterocystous cyanobacteria arose significantly after the appearance of other cyanobacterial lines. The oxygenic nature of cyanobacterial photosynthesis would have provided selective pressure for the evolution of a mechanism for sequestering the oxygen-sensitive process of N_2 reduction, i.e., the heterocyst. Aerobic nitrogen fixation is not limited to the heterocystous cyanobacteria; it also occurs in vegetative cells of some unicellular cyanobacteria (*Synechococcus* spp. [26] and *Gloeothece* spp. [34]) and in other eubacterial phyla. These lineages separate earlier in the inferred phylogeny than do the heterocystous forms, so it is likely that the latter had no role in the origin of biological nitrogen fixation.

This study provides the most detailed phylogenetic data available on the evolution of oxygenic photosynthesis. Fundamental questions remain to be answered. Do green chlo-

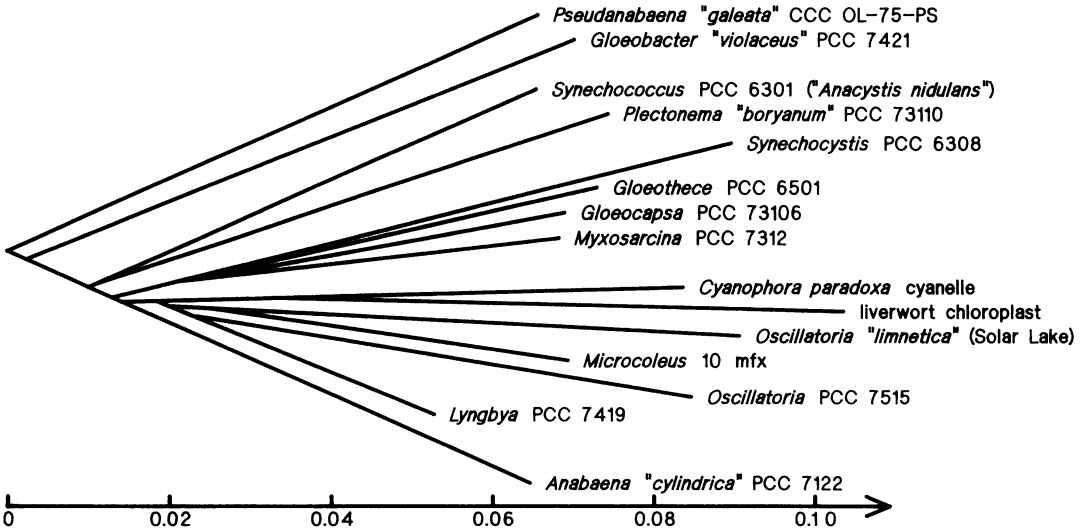

FIG. 6. Rooted-tree topology illustrating the relationships of green chloroplasts and cyanelle 16S rRNAs to cyanobacterial 16S rRNAs. The presentation is as described in the legend to Fig. 4. See text for discussion.

roplasts share a common ancestor with modern prochlorophytes? Do rhodophyte, chlorophyte, chrysophyte, and cryptomonad chloroplasts have monophyletic or polyphyletic origins? As further phylogenetic and biochemical data accumulate, an integrated view of this complex evolutionary history will emerge.

ACKNOWLEDGMENTS

We thank John Waterbury, Richard Castenholz, and Yehuda Cohen for the strains used in this study and Jessup Shively and Stewart Maxwell for *C. paradoxa* cyanelle RNA.

This work was supported by Public Health Service grant GM34527 to N.R.P. from the National Institutes of Health, National Science Foundation grant BSR 8600170 to S.J.G., and Office of Naval Research grants N14-86-K-268 to G.J.O. and N14-87-K-0813 to N.R.P.

LITERATURE CITED

1. **Ashlock, P. D.** 1979. An evolutionary systematist's view of classification. Syst. Zool. **28**:441–450.
2. **Awramik, S. M., J. W. Schopf, and M. R. Walter.** 1983. Filamentous fossil bacteria 3.5×10^9 years old from the archean of western Australia. Precambrian Res. **20**:357–374.
3. **Bonen, L., and W. F. Doolittle.** 1975. On the prokaryotic nature of red algal chloroplasts. Proc. Natl. Acad. Sci. USA **72**:2310–2314.
4. **Bonen, L., and W. F. Doolittle.** 1976. Partial sequences of 16S rRNA and the phylogeny of blue-green algae and chloroplasts. Nature (London) **261**:669–673.
5. **Bonen, L., W. F. Doolittle, and G. E. Fox.** 1979. Cyanobacterial evolution: results of 16S ribosomal ribonucleic acid sequence analysis. Can. J. Biochem. **57**:879–888.
6. **Brosius, J., T. J. Dull, D. D. Sleeter, and H. F. Noller.** 1981. Gene organization and primary structure of a ribosomal RNA operon from *Escherichia coli*. J. Mol. Biol. **148**:107–127.
7. **Burger-Wiersma, T., M. Veenhuis, H. J. Korthals, C. C. M. Van de Wiel, and L. R. Mur.** 1986. A new prokaryote containing chlorophylls *a* and *b*. Nature (London) **320**:262–264.
8. **Castenholz, R. W., and H. C. Utikilin.** 1984. Physiology of sulfide tolerance in a thermophilic *Oscillatoria*. Arch. Microbiol. **138**:299–305.
9. **Cohen, Y., B. B. Jorgenson, E. Padan, and M. Shilo.** 1975. Sulfide-dependent anoxygenic photosynthesis in the cyanobacterium *Oscillatoria limnetica*. Nature (London) **257**:489–491.
10. **Colless, D. H.** 1970. The phenogram as an estimate of phylogeny. Syst. Zool. **19**:352–362.
11. **Dron, M., M. Rahire, and J.-D. Rochaix.** 1982. Sequence of the chloroplast 16S rRNA gene and its surrounding regions of *Chlamydomonas reinhardii*. Nucleic Acids Res. **10**:7609–7620.
12. **Giovannoni, S. J., D. M. Ward, N. P. Revsbech, and R. W. Castenholz.** 1987. Obligately phototrophic *Chloroflexus*: primary production in anaerobic, hot spring microbial mats. Arch. Microbiol. **147**:80–87.
13. **Graf, L., E. Roux, E. Stutz, and H. Kössel.** 1982. Nucleotide sequence of a *Euglena gracilis* DNA coding for the 16S rRNA: homologies to *E. coli* and *Zea mays* chloroplast 16S rRNA. Nucleic Acids Res. **10**:6369–6381.
14. **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.
15. **Guglielmi, G., and G. Cohen-Bazire.** 1984. Étude taxonomique d'un genre de cyanobactérie Oscillatoriacée: le genre *Pseudanabaena* Lauterborn. I. Étude ultrastructurale. Protistologica **20**:377–391.
16. **Guglielmi, G., and G. Cohen-Bazire.** 1984. Étude taxonomique d'un genre de cyanobactérie Oscillatoriacée: le genre *Pseudanabaena* Lauterborn. II. Analyse de la composition moléculaire et de la structure des phycobilisomes. Protistologica **20**:393–413.
17. **Guglielmi, G., G. Cohen-Bazire, and D. A. Bryant.** 1981. The structure of *Gloeobacter violaceus* and its phycobilisomes. Arch. Microbiol. **129**:181–189.
18. **Hall, W. T., and G. Claus.** 1963. Ultrastructural studies on the blue-green algal symbiont in *Cyanophora paradoxa* Korschikoff. J. Cell Biol. **19**:551–563.
19. **Hayes, J. M.** 1983. Geochemical evidence bearing on the origin of aerobiosis, a speculative hypothesis, p. 291–300. *In* J. W. Schopf (ed.), The Earth's earliest biosphere, its origins and evolution. Princeton University Press, Princeton, N.J.
20. **Herdman, M., M. Janvier, J. B. Waterbury, R. Rippka, R. Y. Stanier, and M. Mandel.** 1979. Deoxyribonucleic acid base compositions of cyanobacteria. J. Gen. Microbiol. **111**:63–71.
21. **Jarsch, J., and A. Böck.** 1985. Sequence of the 16S ribosomal RNA gene from *Methanococcus vannielii*: evolutionary implications. Syst. Appl. Microbiol. **6**:54–59.
22. **Jukes, T. H., and C. R. Cantor.** 1969. Evolution of protein molecules, p. 21–132. *In* H. N. Munro (ed.), Mammalian protein metabolism, vol. 3. Academic Press, Inc., New York.
23. **Kimura, M., and T. Ohta.** 1972. On the stochastic model for estimation of mutational distance between homologous proteins. J. Mol. Evol. **2**:87–90.
24. **Lachance, M.-A.** 1981. Genetic relatedness of heterocystous cyanobacteria by deoxyribonucleic acid-deoxyribonucleic acid reassociation. Int. J. Syst. Bacteriol. **31**:139–147.
25. **Lane, D. J., B. Pace, G. J. Olsen, D. A. Stahl, M. L. Sogin, and N. R. Pace.** 1985. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analysis. Proc. Natl. Acad. Sci. USA **82**:6955–6959.
26. **León, C., S. Kumazawa, and A. Mitsui.** 1986. Cyclic appearance of aerobic nitrogenase activity during synchronous growth of unicellular cyanobacteria. Curr. Microbiol. **13**:149–153.
27. **Lewin, R. A.** 1981. The prochlorophytes, p. 257–266. *In* M. P. Starr, H. Stolp, H. G. Trüper, A. Balows, and H. G. Schlegel (ed.), The prokaryotes, a handbook on habitats, isolation, and identification of bacteria. Springer-Verlag, New York.
28. **Maxwell, E. S., J. Liu, and J. M. Shively.** 1986. Nucleotide sequences of *Cyanophora paradoxa* cellular and cyanelle-associated 5S ribosomal RNAs: the cyanelle as a potential intermediate in plastid evolution. J. Mol. Evol. **23**:300–304.
29. **Ohyama, K., H. Fukuzawa, T. Kohchi, H. Shirai, T. Sano, S. Sano, K. Umesono, Y. Shiki, M. Takeuchi, Z. Chang, S.-I. Aota, H. Inokuchi, and H. Ozeki.** 1986. Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. Plant Mol. Biol. Reporter **4**:148–175.
30. **Olsen, G. J.** 1987. The earliest phylogenetic branchings: comparing rRNA-based evolutionary trees inferred with various techniques. Cold Spring Harbor Symp. Quant. Biol. **52**:825–838.
31. **Oyaizu, H., B. Debrunner-Vossbrinck, L. Mandelco, J. A. Studier, and C. R. Woese.** 1987. The green non-sulfur bacteria: a deep branching in the eubacterial line of descent. Syst. Appl. Microbiol. **9**:47–53.
32. **Oyaizu, H., and C. R. Woese.** 1985. Phylogenetic relationships among the sulfate respiring bacteria, myxobacteria and purple bacteria. Syst. Appl. Microbiol. **6**:257–263.
33. **Reichenbach, H., W. Ludwig, and E. Stackebrandt.** 1986. Lack of relationship between gliding cyanobacteria and filamentous gliding heterotrophic eubacteria: comparison of 16S rRNA catalogues of *Spirulina*, *Saprospira*, *Vitreoscilla*, *Leucothrix*, and *Herpetosiphon*. Arch. Microbiol. **145**:391–395.
34. **Rippka, R., J. Deruelles, J. B. Waterbury, M. Herdman, and R. Y. Stanier.** 1979. Generic assignments, strain histories and properties of pure cultures of cyanobacteria. J. Gen. Microbiol. **111**:1–61.
35. **Rippka, R., and M. Herdman.** 1985. Divisional patterns and cellular differentiation in cyanobacteria. Ann. Inst. Pasteur (Paris) **136A**:33–39.
36. **Schopf, J. W., J. M. Hayes, and M. R. Walter.** 1983. Evolution of the Earth's earliest ecosystem: recent progress and unsolved problems, p. 361–384. *In* J. W. Schopf (ed.), The Earth's earliest biosphere, its origins and evolution. Princeton University Press, Princeton, N.J.
37. **Schopf, J. W., and M. R. Walter.** 1987. Early archean (3.3-billion to 3.5-billion-year-old) microfossils from the Warrawoona Group, Australia. Science **237**:70–73.

38. **Schwarz, Z., and H. Kössel.** 1980. The primary structure of 16S rDNA from *Zea mays* chloroplast is homologous to *E. coli* 16S rRNA. Nature (London) **283**:739–742.
39. **Stahl, D. A., D. J. Lane, G. J. Olsen, D. J. Heller, T. M. Schmidt, and N. R. Pace.** 1987. Phylogenetic analysis of certain sulfide-oxidizing and related morphologically conspicuous bacteria by 5S ribosomal ribonucleic acid sequences. Int. J. Syst. Bacteriol. **37**:116–122.
40. **Starnes, S. M., D. H. Lambert, E. S. Maxwell, S. E. Stevens, R. D. Porter, and J. M. Shively.** 1985. Cotranscription of the large and small subunit genes of ribulose-1,5-bisphosphate carboxylase/oxygenase in *Cyanophora paradoxa*. FEMS Microbiol. Lett. **28**:165–169.
41. **Tohdoh, N., and M. Sugiura.** 1982. The complete nucleotide sequence of a 16S ribosomal RNA gene from tobacco chloroplasts. Gene **17**:213–218.
42. **Tomioka, N., and M. Sugiura.** 1983. The complete nucleotide sequence of a 16S ribosomal RNA gene from a blue-green alga, *Anacystis nidulans*. Mol. Gen. Genet. **191**:45–50.
43. **Walter, M. R.** 1987. Archean stromatolites: evidence of the Earth's earliest benthos, p. 187–212. *In* J. W. Schopf (ed.), The Earth's earliest biosphere, its origins and evolution. Princeton University Press, Princeton, N.J.
44. **Weisburg, W. G., C. R. Woese, M. E. Dobson, and E. Weiss.** 1985. A common origin of Rickettsiae and certain plant pathogens. Science **230**:556–558.
45. **Woese, C. R.** 1987. Bacterial evolution. Microbiol. Rev. **51**:221–271.
46. **Yang, D., Y. Oyaizu, H. Oyaizu, G. J. Olsen, and C. R. Woese.** 1985. Mitochondrial origins. Proc. Natl. Acad. Sci. USA **82**: 4443–4447.