orfy, Eds. [*The Hydrogen Bond*, vol. 1, *Theory*; vol. 2, *Structure and Spectroscopy*; vol. 3, *Dynamics, Thermodynamics, and Special Systems* (North-Holland, Amsterdam, 1976)].
5. M. Atoji and W. N. Lipscomb, *Acta Crystallogr.* **7**, 173 (1954).
6. I. Olovsson and D. H. Templeton, *ibid.* **12**, 832 (1959).
7. D. W. Davidson, in *Water, A Comprehensive Treatise*, F. Franks, Ed. (Plenum, New York, 1973), vol. 2, p. 115.
8. D. Eisenberg and W. Kauzmann, *The Structure and Properties of Water* (Oxford Univ. Press, New York, 1969), chap. 3.
9. T. R. Dyke, K. M. Mack, J. S. Muenter, *J. Chem. Phys.* **66**, 498 (1977).
10. The calculations are reviewed in F. H. Stillinger, *Adv. Chem. Phys.* **31**, 1 (1975); see also P. Schuster, in *The Hydrogen Bond*, P. Schuster, G. Zundel, C. Sandorfy, Eds. (North-Holland, Amsterdam, 1976), vol. 1, p. 25.
11. H. S. Frank and W. Y. Wen, *Discuss. Faraday Soc.* **24**, 133 (1957); H. S. Frank, *Proc. R. Soc. London Ser. A* **247**, 481 (1958).
12. H. Popkie, H. Kistenmacher, E. Clementi, *J. Chem. Phys.* **59**, 1325 (1973).
13. N. Bjerrum, *Science* **115**, 385 (1952); L. Pauling, *The Nature of the Chemical Bond* (Cornell Univ. Press, Ithaca, N.Y., ed. 3, 1960).
14. A. H. Narten and H. A. Levy, *J. Chem. Phys.* **55**, 2263 (1971).
15. A. H. Narten, *ibid.* **56**, 5681 (1972).
16. S. Swaminathan and D. L. Beveridge, *J. Am. Chem. Soc.* **99**, 8392 (1977).
17. F. H. Stillinger and A. Rahman, *J. Chem. Phys.* **60**, 1545 (1974).
18. F. H. Stillinger, *J. Phys. Chem.* **74**, 3677 (1970).
19. L. Pauling, in *Hydrogen Bonding*, D. Hadzi and H. W. Thompson, Eds. (Pergamon, London, 1959), p. 1.
20. D. P. Stevenson, *J. Phys. Chem.* **69**, 2145 (1965).
21. A Rahman and F. H. Stillinger, *J. Am. Chem. Soc.* **95**, 7943 (1973).
22. G. Nemethy and H. A. Scheraga, *J. Chem. Phys.* **36**, 3382 (1962).
23. A. Geiger, F. H. Stillinger, A. Rahman, *ibid.* **70**, 4185 (1979).
24. A. Rahman and F. H. Stillinger, *ibid.* **55**, 3336 (1971); W. L. Jorgensen, *Chem. Phys. Lett.*, in press.
25. C. A. Angell, in *Water, A Comprehensive Treatise*, F. Franks, Ed. (Plenum, New York, in press), vol. 7.
26. H. E. Stanley, *Introduction to Phase Transitions and Critical Phenomena* (Oxford Univ. Press, New York, 1971).
27. F. H. Stillinger, in *Water in Polymers*, S. P. Rowland, Ed. (American Chemical Society, Washington, D.C., in press).
28. J. H. Hildebrand, *J. Phys. Chem.* **72**, 1841 (1968).
29. F. Franks, in *Water, A Comprehensive Treatise*, F. Franks, Ed. (Plenun, New York, 1975), vol. 4, p. 1.
30. A. Ben-Naim, *Hydrophobic Interactions* (Plenum, New York, 1980).
31. V. G. Dashevsky and G. N. Sarkisov, *Mol. Phys.* **27**, 1271 (1974); J. C. Owicki and H. A. Sheraga, *J. Am. Chem. Soc.* **99**, 7413 (1977); S. Swaminathan, S. W. Harrison, D. L. Beveridge, *ibid.* **100**, 5705 (1978); A. Geiger, A. Rahman, F. H. Stillinger, *J. Chem. Phys.* **70**, 263 (1979); C. Pangali, M. Rao, B. J. Berne, *ibid.* **71**, 2982 (1979).
32. W. Kauzmann, *Adv. Protein Chem.* **14**, 1 (1959).
33. G. Wada and S. Umeda, *Bull. Chem. Soc. Jpn.* **35**, 646 (1962).

The Phylogeny of Prokaryotes

G. E. Fox, E. Stackebrandt, R. B. Hespell, J. Gibson
J. Maniloff, T. A. Dyer, R. S. Wolfe, W. E. Balch
R. S. Tanner, L. J. Magrum, L. B. Zablen, R. Blakemore
R. Gupta, L. Bonen, B. J. Lewis, D. A. Stahl
K. R. Luehrsen, K. N. Chen, C. R. Woese*

A revolution is occurring in bacterial taxonomy. What had been a dry, esoteric, and uncertain discipline—where the accepted relationships were no more than officially sanctioned speculation—is becoming a field fresh with the excitement of the experimental harvest. For the most part the transition reflects the realization that molecular sequencing techniques permit a direct measurement of genealogical relationships.

Bacterial genealogies appear to be far more ancient than their eukaryotic counterparts. For this reason techniques that have worked so well in establishing eukaryotic phylogenies—such as comparative analysis of cytochrome c sequences (*1*)—are of limited value when applied to the bacteria. However, by the choice of the properly constrained molecule, the molecular approaches to phylogeny that have proved to be successful for eukaryotes can readily be extended to the bacterial domain. The molecules of choice in this instance appear to be the ribosomal RNA's (rRNA's). They are universally distributed, exhibit constancy of function, and appear to change in sequence very slowly—far more slowly than most proteins. Moreover, they are readily isolated.

Over the past decade, comparative analysis of the 16*S* ribosomal RNA sequence has been used to explore prokaryote phylogeny. For the most part these studies have been carried out at the University of Illinois. However, the cyanobacterial characterizations discussed below have been done at Dalhousie. To date more than 170 individual species have been characterized in this way (*2–16*). In this article we summarize these results and provide a first glimpse of the outline of prokaryotic phylogeny.

Experimental details of the approach and the methods of data analysis have been described (*3*, *4*, *17*). In summary: ^{32}P-labeled 16*S* rRNA from a given species is digested with ribonuclease T1, and the resulting oligonucleotides are resolved by two-dimensional paper electrophoresis (*17*) and then sequenced—thus producing a catalog of sequences characteristic of the organism. Comparisons of these catalogs—which is tantamount to comparing the original 16*S* rRNA sequences—reveals the genealogical relationships among the organisms considered. These relationships can be quantified in terms of an association coefficient, S_{AB}, which is calculated for each binary pair. Dendrograms are then obtained by cluster analysis (average linkage between the merged groups) (*18*). Unfortunately, this association coefficient is related to the actual number of nucleotide differences between the underlying 16*S* rRNA sequences in an unknown and nonlinear way. Hence the usual methods of tree construction (that is, the matrix method) have not been employed, since it is unclear how the branch lengths should be interpreted. Nevertheless the analysis appears to yield a good approximation to the true phylogenetic relationships provided that

The authors' scientific addresses past and present are as follows: G. E. Fox, Department of Genetics and Development (UI-G), University of Illinois, Urbana 61801 and the Department of Biophysical Sciences (UH-B), University of Houston, Houston, Texas 77004; E. Stackebrandt, UI-G and Department of Microbiology, Technische Universität München, 8 München, West Germany; R. B. Hespell, Department of Dairy Science, University of Illinois; J.Gibson, Section of Biochemistry, Molecular and Cell Biology, Cornell University, Ithaca, New York 14853; J. Maniloff, Department of Microbiology, University of Rochester Medical Center, Rochester, New York 14642; T.A. Dyer, Plant Breeding Institute, Maris Lane, Trumpington, Cambridge CB2 2LQ, England; R. S. Wolfe, Department of Microbiology (UI-M), University of Illinois; W. E. Balch, UI-M and Department of Biochemistry, Stanford University Medical Center, Palo Alto, California 94305; R. S. Tanner, UI-M; L. J. Magrum, UI-G; L. B. Zablen, UI-G; R. Blakemore, UI-M and Department of Microbiology, University of New Hampshire, Durham 03824; R. Gupta, UI-G; L. Bonen, UI-G and Department of Biochemistry, Dalhousie University, Halifax, Nova Scotia; B. J. Lewis, UI-G; D. A. Stahl, UI-M and National Jewish Hospital, Denver, Colorado 80206; K. R. Luehrsen, UI-G and UH-B; K. N. Chen, UH-B; C. R. Woese, Department of Genetics and Development, 515 Morrill Hall, University of Illinois, Urbana 61801.
This article is dedicated to Professor C. B. van Niel.

0036-8075/80/0725-0457$01.75/0

Summary. For the first time a single experimental approach, 16*S* ribosomal RNA sequence characterization, has been used to develop an overview of phylogenetic relationships in the bacterial world. The technique permits the tracing of relationships back to the common ancestor of all extant life. This first glimpse of bacterial phylogeny reveals a world whose roots appear to span more than 3 billion years. A deep phylogenetic split exists among the bacteria, which necessitates their division into two major lines of descent, the archaebacteria and the true bacteria (or eubacteria). It is a general finding that the most ancient bacterial phenotypes are anaerobic, and that aerobic phenotypes have arisen a number of times. Photosynthetic phenotypes are also extremely ancient. Many nonphotosynthetic groups appear to have arisen from photosynthetic ancestry, which is reason to question the generally held belief that the first bacteria were anaerobic heterotrophs. The two ultimate lines of bacterial descent are no more closely related to one another than either is to the cytoplasmic aspect of the eukaryotic cell. However, in that the eukaryotic cell is a phylogenetic chimera, it itself cannot be seen as a line of descent comparable to the two bacterial lines—although some of its individual parts can be so viewed. In this way, the chloroplast and perhaps the mitochondrion are each eubacterial, and at least one ribosomal protein is archaebacterial. A third line of descent that is neither eubacterial nor archaebacterial is represented in the 18*S* ribosomal RNA.

all the lines of descent considered have evolved at similar rates, which is generally the case (*4*, *15*). This problem is discussed in (*15*).

The Basic Phylogenetic Categories

The general outline of phylogeny seen in 16*S* rRNA comparisons is shown in Fig. 1. Three lines of descent emerge from a common ancestor. These define the primary kingdoms, the ultimate groupings on the prokaryotic organizational level (*5*, *6*). The S_{AB} values separating organisms in one kingdom from those in another are in the range of 0.10, a factor of 2 less than the S_{AB} values that separate distantly related species within any kingdom (*5*). There is no indication in the data that any two of the kingdoms are specifically related to one another to the exclusion of the third (*5*). The better characterized of the bacterial kindgoms contains most of the commonly recognized species and is therefore designated the true bacteria or eubacteria. A second kingdom contains methanogenic bacteria, the extreme halophiles, and certain thermoacidophiles, and is referred to collectively as the archaebacteria. The remaining kingdom, is defined solely in terms of one aspect of the eukaryotic cell, its cytoplasmic rRNA (*5*).

This tripartite division of extant life is incompatible with the conventionally accepted view in which living systems are divided into two basic phylogenetic categories, prokaryotes and eukaryotes (*19*, *20*). However, the eukaryotic cell is now recognized to be a genetic chimera, whose evolutionary origins we do not yet understand (*20*, *21*). It is no longer permissible to take these two categories—or their equivalents as others define them (*19*, *22*)—as phylogenetically comparable, much less to see them as exclusive. It is only the genealogically distinct components of the eukaryotic cell that can be compared individually to prokaryotes. Let us now consider in some detail the two bacterial kingdoms, the archaebacteria and the eubacteria.

The Archaebacteria

Although there is a superficial resemblance between them, the archaebacteria are quite distinct from the eubacteria in phenotype—sufficiently so to justify a separation of the two at the highest taxonomic level (*23*). *Archaebacteriae* should be considered a separate kingdom of prokaryotes that possess (i) a variety of cell walls, none of which contain muramic acid (the hallmark of eubacterial walls) (*24*), (ii) membranes whose major component is a branched chain (phytanyl), ether-linked lipid (*25*), (iii) transfer RNA's (tRNA) devoid of ribothymidine in the TΨC loop (T, thymine; Ψ, pseudouridine; C, cytidine) (*26*), (iv) distinctive RNA polymerase subunit structures (*27*, *28*), and perhaps (v) an unusual but still not fully elaborated spectrum of coenzymes (*11*, *29*). (All archaebacteria demonstrate the first four of these properties; the fifth so far is confined largely to methanogens.)

The 16*S* rRNA data (Fig. 2) indicate that the archaebacteria break into two or three major subclusters: the methanogens and extreme halophiles in one group, and the thermoacidophile *Sulfolobus* in the other. It is uncertain from these data whether the wall-less thermoacidophile, *Thermoplasma*, is best grouped specifically with one or the other of these or must be considered to represent a third line. The subunit structure of RNA polymerase (*27*) has led Zillig and co-workers to prefer two major subclusters; the methanogens and halophiles in one group and the two thermoacidophiles in the other.

Figure 2 also shows that the methanogenic phenotype encompasses three major subgroups, the orders *Methanobacteriales*, *Methanococcales*, and *Methanomicrobiales* (*11*). The cluster of extreme halophiles falls among these three methanogenic groups. Although an S_{AB} analysis suggests that the halophiles may be most closely related to the *Methanobacteriales*, patterns of posttranscriptional modification of nucleotides in 16*S* rRNA (*11*) and the structure of 5*S* rRNA (*30*) point to a closer, specific relationship of the halophiles to the *Methanomicrobiales*. It should be mentioned also that the 16*S* rRNA catalogs for representative strains of *Halobacterium halobium*, *H. cutirubrum*, and *H. salinarium* are all identical, while that for *Amoebobacter morrhuae* differs from these by a single base change (*16*). These four organisms then must be considered at best strains of a single species. The specific groupings of methanogens indicated in Fig. 2 are in excellent agreement with various phenotypic characteristics of the various members—for example, cell wall and lipid compositions. This topic has been reviewed by Balch *et al*. (*11*).

In spite of the small number of known species, the diversity of the archaebacteria as measured by phylogenetic depth and phenotypic variety appears comparable to that encountered in the true bacteria. For example, at least four major cell wall types are encountered in the archaebacteria, compared to only one among all the eubacteria (*24*). DNA base compositions, a standard measure of phylogenetic diversity, range in guanine + cytosine (G + C) content from 27 to 68 moles percent (that is, per 100 moles) (*11*, *31*) among archaebacteria. Yet the archaebacteria appear restricted to an unusual class of niches. It is not unreasonable that the archaebacterial ancestral phenotype evolved at a period in the earth's history when what was a typical niche was very different from what it was later—that is, with respect to its physical parameters and the extent to which these varied. If so, basic archaebacterial metabolism and control mechanisms could have evolved to reflect conditions that no longer prevail.

The True Bacteria (Eubacteria)

All the bacterial strains examined to date by 16*S* rRNA cataloging that do not cluster with the archaebacteria, cluster instead in the eubacterial kingdom. The designation true bacteria reflects the fact that it is these organisms that have shaped our conception of bacteria. Eight major eubacterial divisions have so far been identified (Fig. 3), although only four of these have been explored to a significant extent by the present technique. It appears that all the major eubacterial groups diverged from one another over a relatively short period of time, so that their exact order of branching is difficult to determine.

The eight recognized groups are as follows:

1) A group (Fig. 4) containing the purple photosynthetic bacteria and various nonphotosynthetic genera such as *Escherichia*, *Pseudomonas*, *Rhizobium*, *Alcaligenes*, and *Desulfovibrio*.

2) A group (Figs. 5 and 6) which includes all Gram-positive eubacteria so far examined except for *Micrococcus radiodurans* and its immediate relatives, and in addition contains all mycoplasmas so far studied (except for *Thermoplasma*, which is a member of the archaebacteria) (*7*).

3) An apparently small group of nonsporeforming Gram-positive bacteria that includes *Micrococcus radiodurans* and relatives but no other *Micrococcus* species (*14*, *32*).

4) Members of the genus *Spirochaeta*.

5) *Leptospira*, which do not cluster with *Spirochaeta*.

6) The cyanobacteria (Fig. 7), a group which likely encompasses the chloroplast of *Porphyridium* (*8*) and shares a common ancestry with the chloroplasts of *Lemna* (*9*) and *Euglena* (*10*).

7 and 8) Two distinct lines of green photosynthetic bacteria, one represented by *Chlorobium*, the other by *Chloroflexus*. There is no indication that these eight categories cover all of the true bacteria. In fact, preliminary characterizations of some other species, such as myxobacteria, indicate that one or more categories will ultimately have to be added to the list.

General Taxonomic Considerations

Several general observations regarding bacterial systematics can be made at this point. While in many instances the present genealogies are consistent with traditional classifications, it is clear that in others the classical taxa are not phylo-

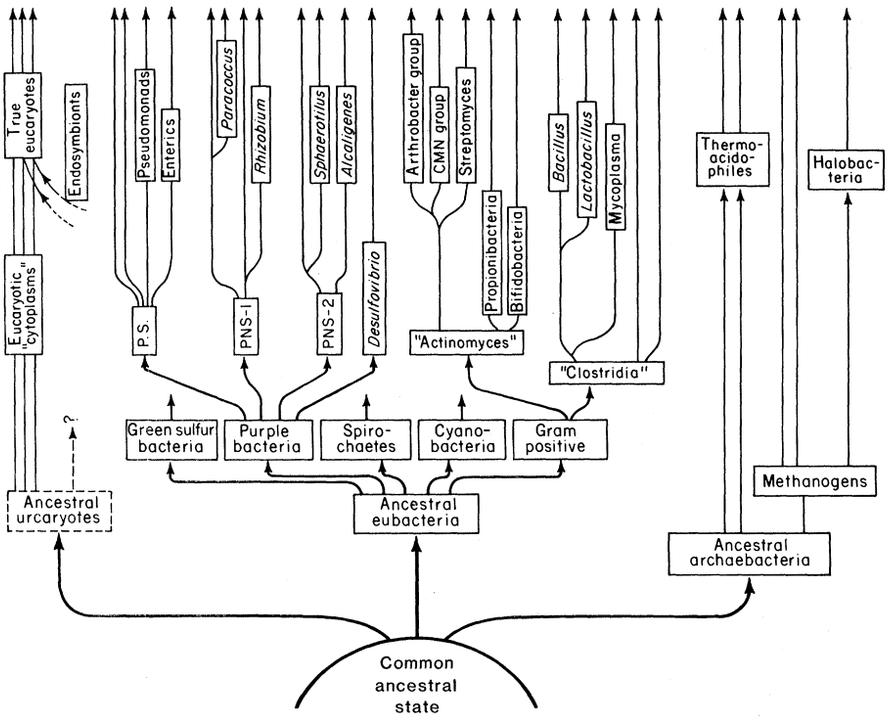


Fig. 1. Schematic representation of the major lines of prokaryotic descent.

genetically valid units. The heavy emphasis traditionally placed on morphological characteristics is seen not to be justified. Spherical shape is a principal offender. All spherical bacteria so far examined fall into phylogenetic categories defined in terms of nonspherical organisms. Examples are *Paracoccus* (a specific relative of *Rhodopseudomonas capsulata*) (*12*), *Sporosarcina* (a specific relative of *Bacillus pasteurii*) (*13*), and *Micrococcus* (the species of which are genealogically intermixed with those of *Arthrobacter*) (*14*).

Two other morphological characters that are phylogenetically deceptive are mode of cell division and lack of a cell wall. The fact that *Rhodomicrobium* divides by budding does not distinguish it strongly from nonbudding representatives of the purple nonsulfur bacteria (*12*). The mycoplasmas, which some would accord the lofty status of a separate class, division, or kingdom, are genealogically wall-less relatives of a particular subgroup of clostridia (*15*). In fact, the data can be interpreted as indicating that *Mycoplasma* and *Acholeplasma* did not even share a common ancestor that was itself wall-less. In other words, the mycoplasma condition may have arisen more than once from clostridial ancestry (*15*). Spore formation is in one sense a good phylogenetic indicator; all sporeformers are related, albeit sometimes at a deep level. However, lack of sporeforming capacity does not necessarily exclude an organism from phylogenetic groupings defined by sporeformers. *Eubacterium*, *Lactobacillus*, and *Streptococcus* are all examples.

The various taxonomic levels defined by traditional criteria bear little relationship to the phylogenetic levels defined by the association coefficient, S_{AB}. What is traditionally defined as a genus varies in evolutionary depth from very "shallow" groupings, such as the genera in the family *Enterobacteriaceae* (minimal S_{AB} values between 0.7 and 0.8), to rather substantial genera such as *Bacillus* (minimal S_{AB} values between 0.5 and 0.55), to "genera" such as *Clostridium*, the evolutionary depth of which compares to that of a "major" grouping like the cyanobacteria (minimal S_{AB} values between 0.03 and 0.35). An entirely comparable situation occurs among eukaryotes, where what is an intrageneric distinction among amphibians corresponds—an immunological distance measure—to intergeneric distinctions or higher among mammals (*33*). The S_{AB} values and related methods measure evolutionary time, not evolutionary "progression." One expects groups of anaerobic bacteria, such as the clostridia, to be older than aerobic groups like *Bacillus*, and therefore to exhibit a greater range of S_{AB} values. Indeed, the S_{AB} ranges allow one to distinguish the ancient phenotypes from the more modern ones. It is not germane to debate

whether taxonomic level for bacteria should be assigned on the basis of phylogenetic depth (a time measure) or—as is done in the case of the metazoans—extent of phenotypic variation (a measure of evolutionary "progression"). Both measures are significant evolutionarily and should be distinguished.

The phylogenetic patterns seen here are by no means an idiosyncracy of rRNA. To the extent to which they can be compared, very nearly the same relationships emerge from comparative analysis of protein sequences and cell wall analysis. Cytochrome c data from a variety of purple photosynthetic bacteria are in good agreement with the corresponding 16*S* rRNA data (*12, 34*). Likewise, results with ferredoxin (*35*), the 5*S* rRNA (*36*), and DNA-RNA hybridization (*37*) are in reasonable agreement; the observed differences are generally at the level of detail and do not for the most part affect major conclusions. Such congruence of phylogenies is of course satisfying, but more importantly it effectively rules out the possibility that the interspecific transfer of genes can obscure evolutionary relationships in the bacterial world. Bacterial phylogenies can be determined experimentally!

The only substantial disagreement among trees derived by molecular methods is between the present one and the attempt of Schwartz and Dayhoff (*38*) to construct a universal phylogeny by combining various approaches. Given the local agreement among methods, it appears that the composite tree generated by Schwartz and Dayhoff reflects unwarranted assumptions in the analysis of the data. A detailed critique of the Schwartz-Dayhoff tree has recently appeared (*39*).

Evolutionary Implications

The implications of the prokaryotic phylogenetic tree go far beyond the bounds of mere classification. The tree is essential to our understanding of the early events in cellular evolution. Previously we have had very few facts to guide us. Nevertheless, certain hypotheses and lesser prejudices have emerged as generally accepted truths concerning the primeval evolutionary course. The keystone of the conventional view is the notion—an integral part of the Oparin-Haldane view of the origin of life—that the first organisms were heterotrophic anaerobes, that is, organisms metabolically resembling today's clostridia (*40*). The ultimate goal of phylogenetic study is to construct a phylogenetic tree sufficiently detailed so that such suggestions can be treated as testable hypotheses.

Photosynthetic bacteria hold a prominent place in the hierarchy of the true bacteria. Four of the eight major groups are in fact defined by photosynthetic bacteria. Moreover, the traditional tendency to separate the photosynthetic from the nonphotosynthetic bacteria phylogenetically does not hold. Photosynthetic lines of descent tend to be intermixed with the lines of nonphotosyn-

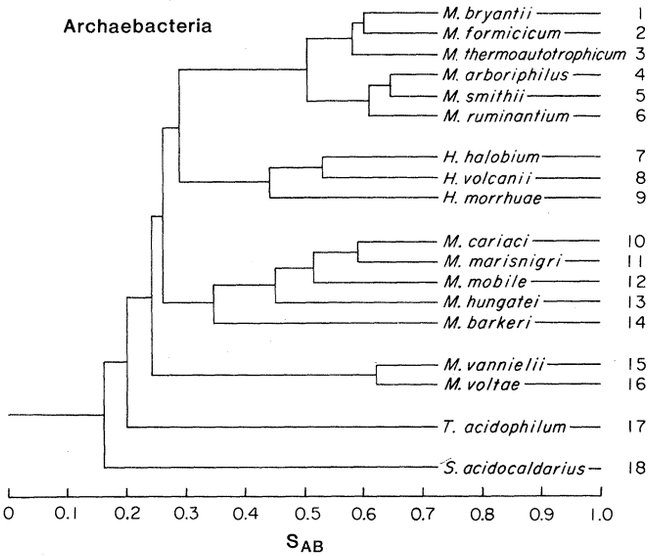


Fig. 2 (left). The phylogenetic relationships among archaebacteria. The abbreviated names stand for the following organisms (*11*): (1) *Methanobacterium bryantii*, (2) *Methanobacterium formicicum*, (3) *Methanobacterium thermoautotrophicum*, (4) *Methanobrevibacter arboriphilus*, (5) *Methanobrevibacter smithii*, (6) *Methanobrevibacter ruminantium*, (7) *Halobacterium halobium*, (8) *Halobacterium volcanii*, (9) *Halococcus morrhuae*, (10) *Methanogenium cariaci*, (11) *Methanogenium marisnigri*, (12) *Methanomicrobium mobile*, (13) *Methanospirillum hungatei*, (14) *Methanosarcina barkeri*, (15) *Methanococcus vannielii*, (16) *Methanococcus voltae*, (17) *Thermoplasma acidophilum*, (18) *Sulfolobus acidocaldarius*.

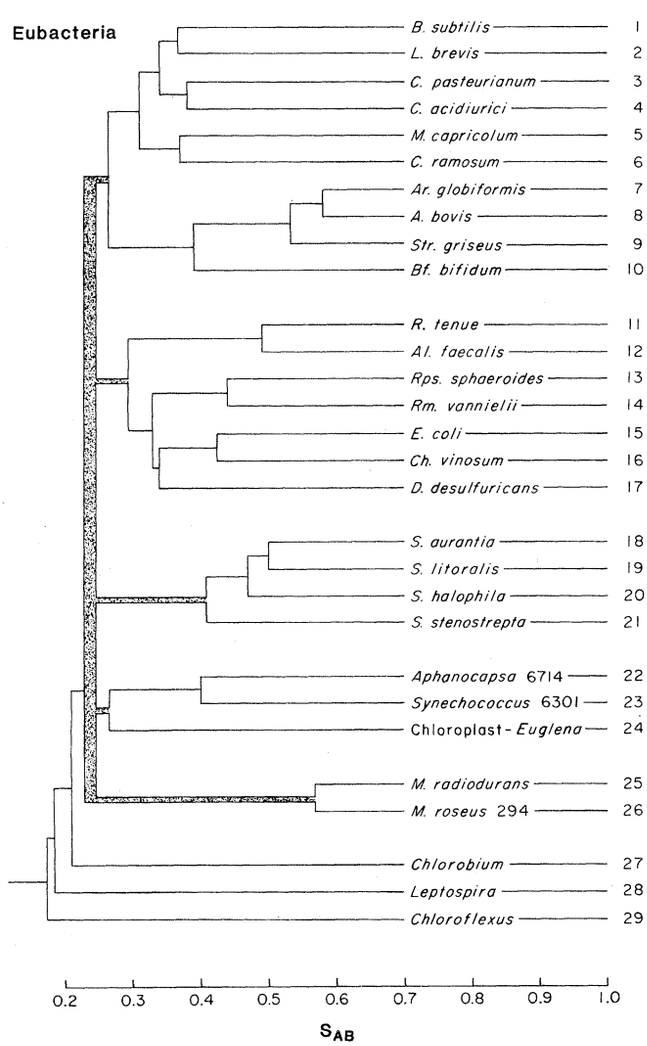


Fig. 3 (right). The major phylogenetic groups within the true bacteria. The abbreviated names stand for these organisms: (1) *Bacillus subtilis*, (2) *Lactobacillus brevis*, (3) *Clostridium pasteurianum*, (4) *C. acidiurici*, (5) *Mycoplasma capricolum*, (6) *C. ramosum*, (7) *Arthrobacter globiformis*, (8) *Actinomyces bovis*, (9) *Streptomyces griseus*, (10) *Bifidobacterium bifidum*, (11) *Rhodospirillum tenue*, (12) *Alcaligenes faecalis*, (13) *Rhodopseudomonas sphaeroides*, (14) *Rhodomicrobium vannielii*, (15) *Escherichia coli*, (16) *Chromatium vinosum*, (17) *Desulfovibrio desulfuricans*, (18 to 21) *Spirochaeta aurantia*, *S. litoralis*, *S. halophila*, and *S. stenostrepta*, (25 and 26) *Micrococcus radiodurans*, *M. roseus* UW0294, (27) *Chlorobium limicola*, (28) *Leptospira interrogans*, and (29) *Chloroflexus aurantiacus*.

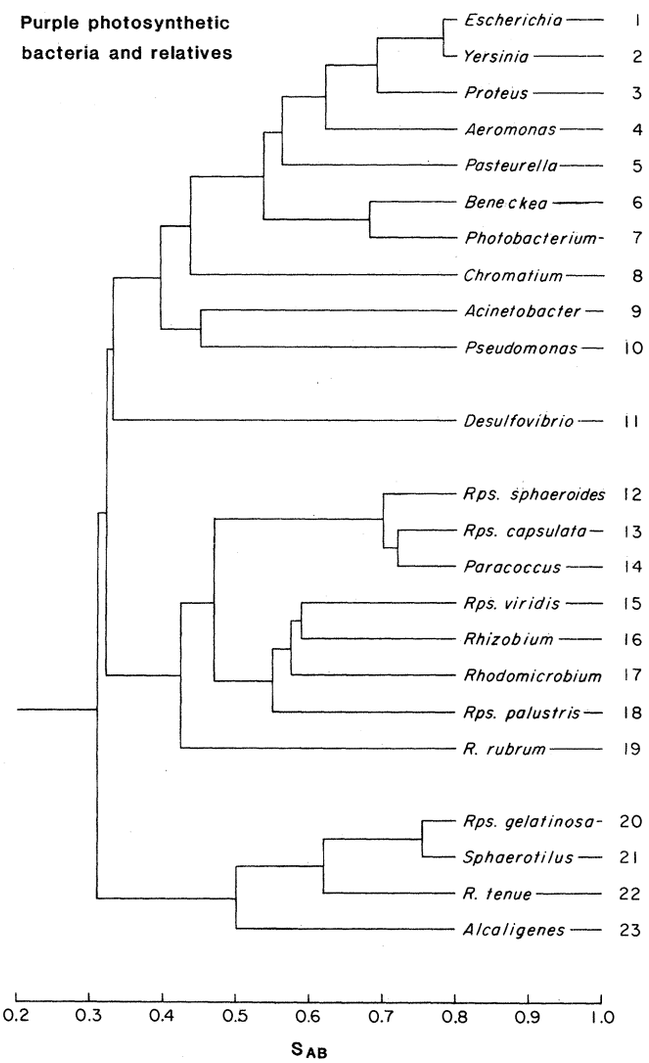
thetic species. The most thoroughly studied example up to now is that of the purple bacteria. This group of organisms comprises three distinct sublines, each containing photosynthetic representatives (Fig. 1): (i) the subline defined by the majority of *Rhodopseudomonas* species together with some *Rhodospirillum* and *Rhodomicrobium* species, (ii) the subline defined by *Rhodopseudomonas gelatinosa* and *Rhodospirillum tenue*, and (iii) the subline defined by the purple sulfur group, such as *Chromatium*. The first of these contains in addition *Paracoccus*, *Rhizobium*, and *Aquaspirillum*, while the second, in addition, contains *Alcaligenes* and *Sphaerotilus*; the third contains a large number of the classically recognized Gram-negative bacteria—enterics, vibrios, pseudomonads, and others (*12*). Thus many, if not most, of the Gram-negative nonphotosynthetic organisms appear to be subsumed by this group. Overall, the widest separations among the purple bacteria give S_{AB}'s of 0.31, which would suggest that this particular photosynthetic phenotype is as old as the clostridial one, whose widest separations are at $S_{AB} = 0.29$.

Realizing the quite incomplete nature of the eubacterial tree at its present stage, we can nevertheless draw a number of interesting if tentative conclusions concerning the evolutionary course. The phylogenetic arrangement of eubacterial phenotypes constitutes as good a proof as any that the earth passed from a globally anaerobic, highly reducing environment to an aerobic one. Those phenotypic groups with the greatest range of S_{AB} values, that is, the oldest groups, are all basically anaerobic—for example, the clostridia, the purple photosynthetic bacteria, and the cyanobacteria (if the various chloroplasts are included). Where a sufficient number of strains has been analyzed to give a reliable indication of diversity, aerobic phenotypes all form far shallower groupings. A good example is *Bacillus* and its aerobic relatives (minimal S_{AB} values of 0.50).

Given the speed of bacterial evolution, it is reasonable that a group such as *Bacillus* is about as old as the aerobic environment, so that S_{AB} values in the range of 0.5 should measure the age of the aerobic atmosphere. More important, it is clear that aerobic respiration has arisen many times; in the purple bacterial group in particular, there are three clear examples of it, and perhaps many more. Although the evolution of aerobic metabolism is not always associated with photosynthetic phenotypes initially, it often is.

It is a common belief, based on the interpretation of fossil evidence, that the cyanobacteria are a very ancient grouping. The data obtained by 16*S* rRNA cataloging of extant cyanobacteria have not detected enough diversity to warrant such a conclusion (*8*). Only when one includes the various examples of chloroplasts (*41*) is sufficient diversity found to suggest an age comparable to that of the clostridial or purple bacterial phenotypes. As has been pointed out many times, stromatolites (fossil algal mats) need not have been produced by cyanobacteria, however.

Given the widespread occurrence of



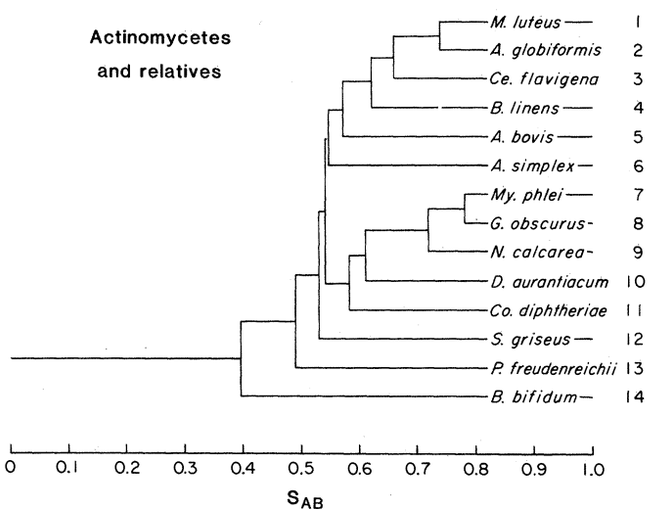


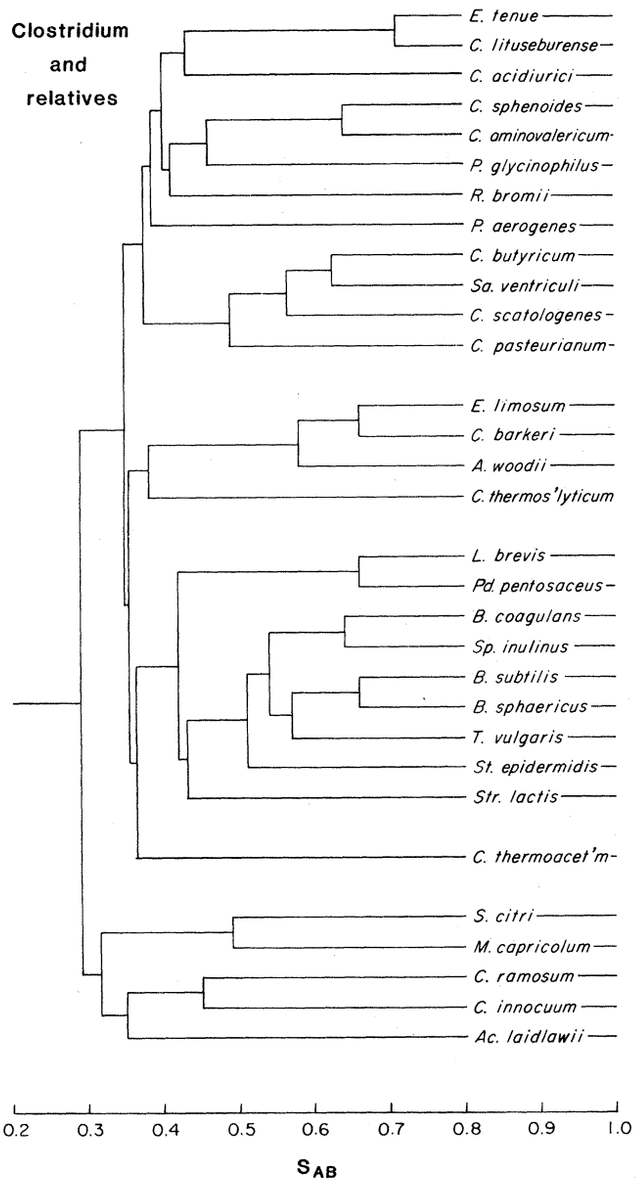
Fig. 4 (left). The phylogenetic relationships among the purple photosynthetic bacteria and their relatives. The organisms are as follows: (1) *Escherichia coli*, (2) *Yersinia pestis*, (3) *Proteus mirabilis*, (4) *Aeromonas hydrophila*, (5) *Pasteurella multocida*, (6) *Beneckea harveyi*, (7) *Photobacterium fischeri*, (8) *Chromatium vinosum*, (9) *Acinetobacter calcoaceticus*, (10) *Pseudomonas aeruginosa*, (11) *Desulfovibrio desulfuricans*, (12 and 13) *Rhodopseudomonas sphaeroides*, *Rps. capsulata*, (14) *Paracoccus denitrificans*, (15) *Rps. viridis*, (16) *Rhizobium leguminosarum*, (17) *Rhodomicrobium vannielii*, (18) *Rps. palustris*, (19) *Rhodospirillum (R.) rubrum*, (20) *Rps. gelatinosa*, (21) *Sphaerotilus natans*, (22) *R. tenue*, and (23) *Alcaligenes faecalis*. Fig. 5 (right). The phylogenetic relationships among the Gram-positive eubacteria whose DNA's have a high content of guanine and cytosine—that is, actinomycetes and their relatives. The organisms are as follows: (1) *Micrococcus luteus*, (2) *Arthrobacter (Ar.) globiformis*, (3) *Cellulomonas flavigena*, (4) *Brevibacterium linens*, (5) *Actinomyces bovis*, (6) *Ar. simplex*, (7) *Mycobacterium phlei*, (8) *Geodermatophilus obscurus*, (9) *Nocardia calcarea*, (10) *Dactylosporangium aurantiacum*, (11) *Corynebacterium diphtheriae*, (12) *Streptomyces griseus*, (13) *Propionibacterium freudenreichii*, and (14) *Bifidobacterium bifidum*.

photosynthetic phenotypes among the eubacteria, the conventional notion that the eubacteria arose from a nonphotosynthetic, heterotrophic, anaerobic ancestry (*40*) must be reexamined. It seems just as reasonable that the ancestral phenotype for the eubacteria was a photosynthetic one, perhaps autotrophic as well.

The ancestral phenotype of the archaebacteria is less clear. In that three of the main archaebacterial sublines are methanogenic, methanogenesis may have been a part of the ancestral phenotype. The only light-utilizing archaebacteria known, the extreme halophiles, seem to have arisen considerably more recently, specifically from one of the methanogenic sublines. Indeed the minimal S_{AB} of 0.44 encountered among the halophiles suggests that they first appeared at about the same time as the major aerobic groups among the eubacteria. The origin of their photosynthetic capacity is for now purely a matter for speculation. In any event, it is clear that the diversity between the various methanogenic strains (that is, a minimal S_{AB} of 0.22) is sufficient to argue that the phenotype was present at a very early stage in evolution. It is interesting that the types of hydrocarbons found in ancient sediments are more closely approximated by the distribution of the lipid components found in methanogens than in any other bacteria (*42*).

The Eukaryotic Line of Descent and Its Origin

The biologist today is accustomed to viewing "eukaryote-prokaryote" as some fundamental phylogenetic dichotomy (*19–22*). However, this cannot be so, for two reasons (*5*): First, the results summarized herein show that there are at least three major lines of descent represented among extant living forms. Second, in that the eukaryotic cell is a phylogenetic chimera, it itself cannot be seen as a line of descent comparable to the two bacterial lines; only some of its component parts can be so viewed: The chloroplast has an origin in cyanobacteria-like entities (*20–21*); the (plant) mitochondrion apparently arose from the group of purple photosynthetic bacteria (*12, 43, 44*). We are less certain of the possible origins of other eukaryotic structures. It might of course be argued that that aspect of the chimeric eukaryote represented by its nucleus accounts for the bulk of the cell (informationally speaking), and therefore the corresponding line of descent, whatever it is, can be taken as the ancestral eukaryotic line. Such an argument makes the tacit assumption that the nuclear



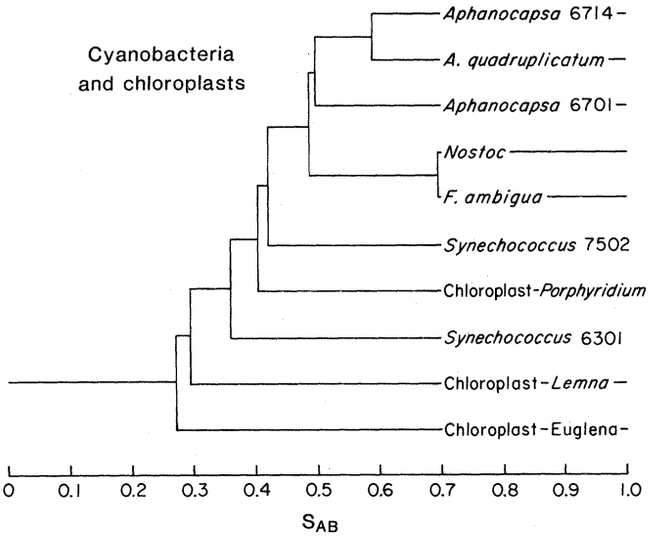


Fig. 6 (left). The phylogenetic relationships among the Gram-positive eubacteria whose DNA's have a low guanine and cytosine content—that is, the clostridia and their relatives. Generic abbreviations used are these: *A, Acetobacterium; Ac, Acholeplasma; B, Bacillus; C, Clostridium; E, Eubacterium; L, Lactobacillus; M, Mycoplasma; P. Peptococcus; Pd, Pediococcus; R, Ruminococcus; S, Spiroplasma; Sa, Sarcina; Sp, Sporolactobacillus; St, Staphylococcus; Str, Streptococcus;* and *T, Thermoactinomyces.* Fig. 7 (right). The phylogenetic relationships among the cyanobacteria and chloroplasts. The generic abbreviation *F.* stands for *Fischerella.* This tree differs slightly in topology from an earlier one (*8*) as a result of minor corrections and the use of a different method of accounting for unsequenced oligonucleotides in the calculation of the binary association coefficients.

component has a single origin—a point for which there is at present no evidence; the collection of genes constituting the nucleus could just as well have arisen through myriad gene or gene cluster captures from all manner of sources. Until this matter is settled—and molecular phylogenetic studies will ultimately do so—it is unwise to speak of an ancestral eukaryotic line of descent.

What we now know is that two primary lines of descent exist as free living forms, the true bacteria (eubacteria) and the archaebacteria, whereas a third line is represented in the eukaryotic 18*S* rRNA. The question concerning eukaryote evolution then becomes the manner in and the extent to which eubacterial and archaebacterial genes, as well as genes from other sources, are represented in the eukaryotic chimera. In addition to the eubacterial representation in terms of mitochondria and chloroplasts, it has recently been determined that at least one of the eukaryotic cytoplasmic ribosomal protein genes is almost certainly of archaebacterial origin (*45*).

The Universal Ancestor

A central issue remaining is the nature of the ancestor that gave rise to all extant life. Presumably features common to all three major lines of descent would be characteristic of their ancestor. There are surprisingly few such common features. Clearly all three major types of organisms have utilized the same basic processes—the same genetic code, similar biochemical pathways, and so on. Yet important details of just about every process appear to differ in one way or another in each of the kingdoms. Genetic control mechanisms seem to differ; RNA polymerase subunit structure differs; rRNA's and tRNA's differ in patterns of posttranscriptional modification; cell walls differ in composition, as do lipids, and so on. These findings suggest that the three lines of descent diverged before the level of complexity usually associated with the prokaryotic cell was reached, at a level of organization which we have called the "progenote" (*6*). Thus, those features of the cell that have to do with refining molecular functions, coping with a large genome, and so on, are evolved independently in the three primary lines of descent, as each reaches a more complex (that is, prokaryotic) level of organization.

References and Notes

1. W. M. Fitch and E. Margoliash, *Science* **155**, 279 (1967); *Biochem. Genet.* **4**, 579 (1970); W. M. Fitch, *J. Mol. Evol.* **8**, 13 (1976).
2. L. Zablen, L. Bonen, R. Meyer, C. R. Woese, *J. Mol. Evol.* **4**, 347 (1975); L. Zablen and C. R. Woese, *ibid.* **5**, 25 (1975); W. E. Balch, L. J. Magrum, G. E. Fox, R. S. Wolfe, C. R. Woese, *ibid.* **9**, 305 (1977); G. E. Fox, L. J. Magrum, W. E. Balch, R. S. Wolfe, C. R. Woese, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 4537 (1977); L. J. Magrum, K. R. Luehrsen, C. R. Woese, *J. Mol. Evol.* **11**, 1 (1978).
3. C. Woese, M. Sogin, D. Stahl, B. J. Lewis, L. Bone, *J. Mol. Evol.* **7**, 197 (1976).
4. G. E. Fox, K. J. Pechman, C. R. Woese, *Int. J. Syst. Bacteriol.* **27**, 44 (1977).
5. C. R. Woese and G. E. Fox, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5088 (1977).
6. _____, *J. Mol. Evol.* **10**, 1 (1977).
7. C. R. Woese *et al.*, unpublished 16*S* rRNA catalog of *Sulfolobus acidocaldarius*.
8. W. F. Doolittle, C. R. Woese, M. L. Sogin, L. Bonen, D. Stahl, *J. Mol. Evol.* **4**, 307 (1975); L. Bonen and W. F. Doolittle, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 2310 (1975); *Nature (London)* **261**, 669 (1976); *J. Mol. Evol.* **10**, 283 (1978); _____, G. E. Fox, *Can. J. Biochem.* **57**, 879 (1979).
9. T. Dyer and C. R. Woese, unpublished 16*S* rRNA catalog of the chloroplast of *Lemna minor*.
10. L. B. Zablen, M. S. Kissil, C. R. Woese, D. E. Buetow, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 2418 (1975).
11. W. E. Balch, G. E. Fox, L. J. Magrum, C. R. Woese, R. S. Wolfe, *Microbiol. Rev.* **43**, 260 (1979).
12. J. Gibson, E. Stackebrandt, L. Zablen, R. Gupta, C. R. Woese, *Curr. Microbiol.* **3**, 59 (1979).
13. K. J. Pechman, B. J. Lewis, C. R. Woese, *Int. J. Syst. Bacteriol.* **26**, 305 (1976).
14. E. Stackebrandt and C. R. Woese, *Curr. Microbiol.* **2**, 317 (1979).
15. C. R. Woese, J. Maniloff, L. B. Zablen, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 494 (1980).
16. In addition to the previous results (*2*-*15*), Woese *et al.* have additional work in various stages of completion. Data from these unpublished works are included in the figures of this article.
17. T. Uchida, L. Bonen, H. W. Schaup, B. J. Lewis, L. Zablen, C. Woese, *J. Mol. Evol.* **3**, 63 (1974); F. Sanger, G. G. Brownlee, B. G. Barrell, *J. Mol. Biol.* **13**, 373 (1965).
18. M. R. Anderberg, *Cluster Analysis for Applications* (Academic Press, New York, 1973).
19. E. Chatton, *Titres et travoux scientifiques* (Sottano, Séte, 1937); R. G. E. Murray, in *Bergey's Manual of Determinative Bacteriology*, R. E. Buchanan and N. E. Gibbons, Eds. (Williams & Wilkins, Baltimore, ed. 8, 1974); A. Allsopp, *New Phytol.* **68**, 591 (1969).
20. R. Y. Stanier, *Symp. Soc. Gen. Microbiol.* **20**, 1 (1970).
21. L. Margulis, *Origin of Eucaryotic Cells* (Yale Univ. Press, New Haven, Conn., 1970).
22. R. H. Whittaker, *Parasitic Protozoa* **1**, 1 (1977).
23. C. R. Woese, L. J. Magrum, G. E. Fox, *J. Mol. Evol.* **11**, 245 (1978); R. K. Thauer and G. Fuchs, *Naturwissenschaften* **66**, 89 (1979).
24. H. König and O. Kandler, *Arch. Microbiol.* **121**, 271 (1979); *ibid.* **123**, 276 and 295 (1979); O. Kandler, *Naturwissenschaften* **66**, 95 (1979); _____ and H. König, *Arch. Microbiol.* **118**, 141 (1978); *Hoppe-Seyler's Z. Physiol. Chem.* **359**, 282 (1978); O. Kandler and H. Hippe, *Arch. Microbiol.* **113**, 57 (1977); J. B. Jones, B. Bowers, T. C. Stadtman, *J. Bacteriol.* **130**, 1357 (1977); J. Steber and K. H. Schleifer, *Arch. Microbiol.* **105**, 173 (1975); T. D. Brock, D. M. Brock, R. T. Belly, R. L. Weiss, *ibid.* **84**, 54 (1972).
25. T. G. Tornabene and T. A. Langworthy, *Science* **203**, 51 (1979); _____, G. Holzer, J. Oro, *J. Mol. Evol.* **13**, 73 (1979); T. G. Tornabene, R. S. Wolfe, W. E. Balch, G. Holzer, G. E. Fox, J. Oro, *ibid.* **11**, 259 (1978); T. A. Langworthy, W. R. Mayberry, P. F. Smith, *J. Bacteriol.* **119**, 106 (1974); T. A. Langworthy, *Biochim. Biophys. Acta* **487**, 37 (1977); M. Kates, in *Ether Lipids, Chemistry, and Biology*, F. Snyder, Ed. (Academic Press, New York, 1972), p. 351.
26. A. N. Best, *J. Bacteriol.* **133**, 240 (1978); R. Gupta and C. R. Woese, unpublished observations.
27. W. Zillig and K. O. Stetter, paper presented, Symposium on Evolution of Molecular Mechanisms, Hokeido, Japan, September 1979.
28. W. Zillig, K. O. Stetter, D. Janekovic, *Eur. J. Biochem.* **96**, 597 (1979); W. Zillig, K. O. Stetter, M. Tobien, *ibid.* **91**, 193 (1978).
29. P. J. Weimer and J. G. Zeikus, *J. Bacteriol.* **137**, 332 (1979); L. D. Eirch, G. D. Vogels, R. S. Wolfe, *Biochemistry* **17**, 4583 (1978); C. D. Taylor and R. S. Wolfe, *J. Biol. Chem.* **249**, 4879 (1974).
30. K. Luehrsen, G. E. Fox, C. R. Woese, unpublished results on 5*S* rRNA sequences of archaebacteria.
31. J. G. Zeikus and D. L. Henning, *J. Microbiol. Serol.* **41**, 171 (1975); J. G. Joshi, W. R. Guild, P. Handler, *J. Mol. Biol.* **6**, 34 (1963); R. L. Moore and B. J. McCarthy, *J. Bacteriol.* **99**, 248 (1969).
32. E. Stackebrandt and R. G. E. Murray, personal communication.
33. A. C. Wilson, S. S. Carlson, T. J. White, *Annu. Rev. Biochem.* **46**, 573 (1977); D. G. Wallace, L. R. Maxson, A. C. Wilson, *Proc. Natl. Acad. Sci. U.S.A.* **68**, 3127 (1973).
34. C. R. Woese, J. Gibson, G. E. Fox, *Nature (London)* **283**, 212 (1980); R. E. Dickerson, *ibid.*, p. 210.
35. M. Tanaka *et al.*, *J. Biol. Chem.* **246**, 3953 (1971).
36. H. Hori and S. Osawa, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 381 (1979).
37. J. De Ley, P. Segers, M. Gillis, *Int. J. Syst. Bacteriol.* **28**, 154 (1978); J. De Smedt and J. De Ley, *ibid.* **27**, 226 (1977); J. L. Johnson and B. S. Francis, *J. Gen. Microbiol.* **88**, 229 (1975); B. Pace and L. L. Campbell, *J. Bacteriol.* **107**, 543 (1971).
38. R. M. Schwartz and M. O. Dayhoff, *Science* **199**, 395 (1978).
39. V. Demoulin, *ibid.* **205**, 1036 (1979).
40. C. B. van Niel, *Cold Spring Harbor Symp. Quant. Biol.* **11**, 285 (1946).
41. The dendrogram shown in Fig. 7 implies nothing concerning the evolutionary relationships among the cytoplasmic components of the eukaryotes whose chloroplasts have been examined. Hence it should not be interpreted as implying that the cytoplasmic aspects of *Porphyridium*, a red algae, are of more recent origin than those of the higher plant *Lemna*. Rather it suggests that the red algal chloroplast is descended from the cyanobacteria while the *Lemna* and *Euglena* chloroplasts are not. Such a view is consistent with pigment characterizations of these chloroplasts.
42. G. Holzer, J. Oro, T. G. Tornabene, *J. Chromatog.* **186**, 795 (1979).
43. R. S. Cunningham, L. Bonen, W. F. Doolittle, M. W. Gray, *FEBS Lett.* **69**, 116 (1976).
44. Although the standard S_{AB} analysis of plant mitochondrial 16*S* rRNA does not reveal a specific relationship to any of the eubacteria, it is possible to show such a relationship, to the purple photosynthetic bacterial group, using a more complex analysis [C. R. Woese and G. E. Fox, unpublished data].
45. A. T. Matheson, W. Moller, R. Amous, M. Yaguchi, in *Ribosomes: Structure, Function, and Genetics*, G. Chamblis, G. R. Craven, J. Davies, K. Davis, L. Kahan, M. Nomura Eds. (University Park Press, Baltimore, 1980), p. 297.

46. The investigations reported herein received granting support as follows: to C.R.W., NASA NSG 7044 and NSF DEB 7727492; to G.E.F., NASA-NSG-7440; to J.G., NASA NSG 7529; NSF PCM 76-02652 to R.S.W.

* Address reprint request to C. R. Woese.