

MINIREVIEW

The Winds of (Evolutionary) Change: Breathing New Life into Microbiology

GARY J. OLSEN,^{1*} CARL R. WOESE,¹ AND ROSS OVERBEEK²

Department of Microbiology, University of Illinois, Urbana, Illinois 61801, and Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Illinois 60439-4801²

Of all the spectacular changes that have transformed biology over the last several decades, the least touted, but ultimately one of the most profound is that now under way in microbiology. It is not simply that we are coming to see microbiology *per se* in a new light but that we are coming to appreciate the central roles microorganisms play in shaping the past and present environments of Earth and the nature of all life on this planet. Because each organism is the product of its history, a knowledge of phylogenetic relationships—of common evolutionary histories—is essential to understanding the nature of any organism. Thus, it is unavoidable that evolution (a field much neglected in this molecular era) must be the conceptual heart of biology; all extant life traces back to common ancestors, and the earliest ancestors were microorganisms.

Despite considerable effort, microbiologists were never able to determine the phylogenetic relationships among prokaryotes (29). Not only did they ultimately give up on this problem, but some went so far as to declare it unsolvable (27, 28). By the 1960s most microbiologists concerned themselves little and cared even less about relationships among microorganisms. As a consequence, microbiology was not structured about a “natural,” phylogenetically valid system of classification. Lacking this essential evolutionary touchstone, microbiology developed in an incomplete, if not distorted, way.

Having no natural center, microbiology fissioned into separate disciplines pursuing (seemingly) divergent goals. The traditional pursuits of organism isolation and determinative classification (i.e., naming) continued, as did the applied side of microbiology (medical, industrial, etc.). However, the disciplines that came to dominate and define the field reflected the new molecular outlook: i.e., the detailed genetics of particular organisms, the comprehensive molecular biology of a “representative” prokaryote, and the detailed biochemistry of this or that metabolic pathway. Microbial ecology was a weak, in a sense immature, discipline, stymied not only by the lack of a natural system but also by the requirement that a species be cultured and characterized before its role in a microbial community could be explored.

The most profound symptom of microbiology’s unfortunate condition was its reliance on the prokaryote-eukaryote dichotomy as a phylogenetic crutch, something that replaced any useful understanding of microbial relationships: it represented microbiology’s “. . . only hope of . . . formulating a ‘concept of

a bacterium” (28). The problem with and the pernicious nature of this dichotomy lay in the fact that the prokaryote was initially defined negatively, in cytological terms. In other words, prokaryotes lacked this or that feature characteristic of the eukaryotic cell: even oil drops, or coacervates, could fit such a negative definition. Any virtue in the prokaryote-eukaryote dichotomy lay in what it could contribute to an understanding of the eukaryote, which might have evolved through “prokaryotic” stages. With repetition (as catechism) the prokaryote-eukaryote dichotomy served only to make microbiologists easily accept their near total ignorance of the relationships among prokaryotes; they were even dulled to the fact—one of the great challenges of today—that they did not in the slightest understand the relationship between the prokaryote and the eukaryote. The matter of relationships among bacteria had boiled down to “if it isn’t a eukaryote, it’s a prokaryote,” and to understand prokaryotes, we had only to determine how *Escherichia coli* differs from the eukaryotes. This was no invitation to creative thought, no unifying biological principle. This eukaryote-prokaryote dichotomy was a barrier that separated prokaryotic microbiology from eukaryotic microbiology.

This myopic view of microbiology failed to appreciate not only how important the problem of microbial relationships was but that an intractable problem today may not be so tomorrow. Molecular sequences had been used to determine evolutionary relationships since the 1950s, and Zuckerkandl and Pauling’s seminal article “Molecules as documents of evolutionary history” made the case most compellingly in 1965 (36). Yet the record shows that microbiology—the biological science most in need—was effectively blind to the significance and potential of these approaches.

At the end of the 1970s however, the situation changed dramatically. rRNA sequences had been shown to provide a key to prokaryote phylogeny (e.g., 8). No matter that on the cellular and physiological levels the prokaryotes did not provide characteristics that permitted their reliable phylogenetic ordering; their rRNAs were more than sufficient to do so. By the early 1980s, as the rRNA-based phylogeny of prokaryotes began to emerge, microbiologists began (albeit extremely slowly) reawakening to the importance of knowing microbial phylogenies.

The folly of having taken all prokaryotes to be of a kind was dramatically revealed by the totally unanticipated discovery of the *Archaea* (originally called archaeobacteria), a group of prokaryotes that, if anything, is more closely related to *Eucarya* (eukaryotes) than to the other prokaryotes, the (true) *Bacteria* (11, 13, 32, 34). Even then, the power of the eukaryote-

* Corresponding author. Mailing address: Department of Microbiology, University of Illinois, 131 Burrill Hall, 407 South Goodwin Ave., Urbana, IL 61801. Phone: (217) 244-0616. Fax: (217) 244-6697. Electronic mail address: gary@phylo.life.uiuc.edu.

prokaryote dichotomy as phylogenetic dogma was strikingly demonstrated by the fact that the majority of microbiologists (and biologists) were initially unable to accept that there could be two types of prokaryotes that are not specifically related to one another. (An amazing aspect of the history is that [at the time] a biologist would most likely have agreed that the nuclear genome came from a "prokaryote-like" ancestor, i.e., the most recent common ancestor of eukaryotes and bacteria. After all, this is inherent in the prokaryote-eukaryote dichotomy. Yet the idea that specific prokaryotic relatives of the eukaryotes have persisted to the present was anathema. As best as we can tell, these noneukaryote descendants do exist; we call them *Archaea*. However, the nature of biologist's objections was not monolithic. For example, Margulis and Schwartz [19] suggested that the origin of [most of] the nuclear genes was indeed a specific group of [true] bacteria, the "aphagrambacteria." This view had the merit of proposing that some prokaryotes are more closely related to eukaryotes than are others, but it ignored the available molecular evidence as to which prokaryotes these were. This confusion was furthered by a drawing of the proposed relationships in which present-day organisms are not clearly distinguished from their ancestors.)

To date, over 1,500 prokaryotes have been characterized by small-subunit rRNA sequencing. Figure 1 encapsulates prokaryotic phylogeny as it is now known. Given the purview of this journal, our discussion will be limited to the two prokaryotic domains. However, it is worthwhile to note in passing that molecular phylogeny has had an equally profound effect on our understanding of relationship among eukaryotes (eukaryotic microorganisms in particular) and that anyone exposed to the universal phylogenetic tree readily appreciates how artificial the strong distinction between eukaryotes and prokaryotes has become. In Fig. 1, the split between the *Archaea* and *Bacteria* is the primary phylogenetic division (in that the *Eucarya* have branched from the same side of the tree as the *Archaea* [11, 13]).

Both prokaryotic domains would seem to be of thermophilic origin—suggesting that life arose in a very warm environment (1, 2, 31). Among the *Archaea*, all of the *Crenarchaeota* cultured to date are thermophiles, and the deepest euryarchaeal branchings are represented exclusively by thermophiles. Among the *Bacteria*, the deepest known branchings are again represented exclusively by thermophiles, and thermophilia is widely scattered throughout the domain.

The *Archaea* comprise a small number of quite disparate phenotypes that grow in unusual (to us, inhospitable) niches. All are obligate or facultative anaerobes. As stated above, all (cultured) crenarchaeotes are thermophilic, some even growing optimally above the normal boiling temperature of water. The *Archaeoglobales* are sulfate reducers growing at high temperatures. The extreme halophiles grow only in highly saline environments (such as the Dead Sea and salt evaporation ponds). The methanogens, which are responsible for virtually all of the biologically produced methane on this planet, are confined to a variety of anaerobic niches, often thermophilic.

The *Bacteria*, on the other hand, are notable as being the source of life's photosynthetic capacity. Five kingdoms of

bacteria contain photosynthetic species, and each of the five manifests a distinct type of (chlorophyll-based) photosynthesis. The cyanobacteria have also given rise to the chloroplasts, which are the basis for all photosynthesis found in eukaryotes. It would appear that photosynthetic metabolism is the ultimate evolutionary source of much of the metabolic diversity found among the bacteria, which in turn is the source of key biochemistries (e.g., photosynthesis and respiration) among the eukaryotes.

The molecular revolution in microbiology means much more than just a new taxonomy. For example, microbial ecology is being reborn. The remarkable insight of Norman Pace that organisms can be (phylogenetically) identified directly in their niches—through a combination of rRNA gene cloning and sequencing, and the design and use of rRNA-directed "phylogenetic stains"—has sparked a revolution in that discipline (5, 24, 25). This simple realization removes the obstacle of having to culture all organisms in order to infer their characteristics, and it permits all manner of new detailed characterizations of the populations in any given niche to be made. Perhaps the most dramatic demonstration of these new approaches to microbial ecology is the discovery (through cloning and sequencing of genes directly from environmental samples) of two new groups of *Archaea*—both of which have to this point defied cultivation (4, 10). One is a new family or order within the euryarchaeotes, the other almost certainly a new taxon of even higher rank (see Pacific Ocean station 49 DNA clone NH49-9, Santa Barbara Channel DNA clone SBAR5, and Woods Hole DNA clone WHARQ in Fig. 1). For the first time, it is possible to count not just flowers and beetles, but also microorganisms, in taking a census of life on this planet.

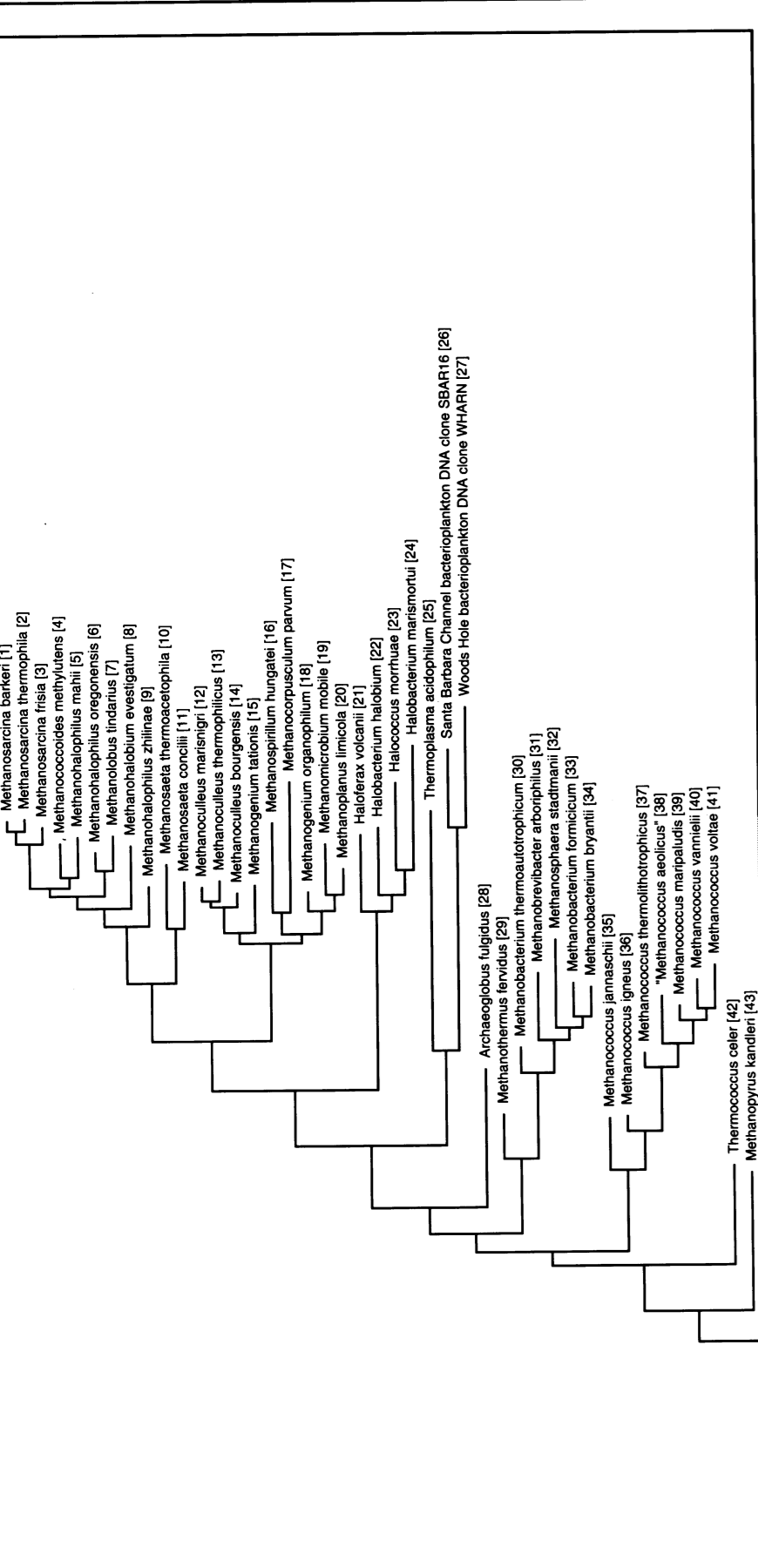
Another area in which evolutionary and comparative perspectives have been invaluable has been the inference of secondary and tertiary RNA structures. Comparisons of homologous RNA sequences from phylogenetically diverse organisms have provided our current understandings of the structures of rRNAs (9, 21, 22, 35), RNase P RNA (14), self-splicing introns (3, 20), signal recognition particle RNA (15, 18), and the small nuclear RNAs of the spliceosome (for a review, see reference 12). In several of these cases, it was rRNA-based phylogenetic information that guided the choice of organisms from which the corresponding molecules were identified and sequenced.

Although our interests are primarily biological, to many people the justification of any pursuit lies in its practical consequences. The revolution in microbial ecology and the inference of molecular structures do, of course, have direct practical consequences (witness *Applied and Environmental Microbiology* and RNA-binding antibiotics), but it is the application in clinical microbiology that most people would think of first. Many clinically important microorganisms have now been subjected to molecular phylogenetic analysis (e.g., 6, 26). These data and analyses have been used to develop molecular diagnostics (e.g., 16) and to guide research into the nature (and control) of various pathogens.

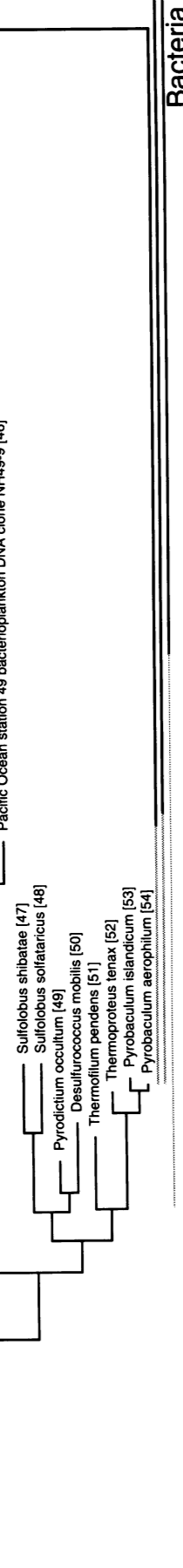
The universal phylogenetic tree brings us face to face with the great evolutionary questions and allows us to formulate them in molecular terms.

FIG. 1. Prokaryotic phylogenetic tree of 253 representative species derived by maximum likelihood analysis of small-subunit rRNA sequences. The tree is abstracted from that provided by the Ribosomal Database Project (17), which is a composite of several trees inferred by using maximum likelihood (7, 23). The suggested rooting is that inferred for the universal tree (11, 13). Organisms are consecutively numbered and are indexed in the accompanying alphabetic listing (Table 1). Names of the major groups (31, 34) are indicated. The distance scale indicates the expected number of changes per sequence position, for those positions changing at the median rate.

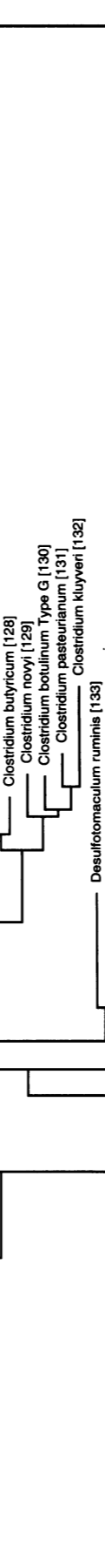
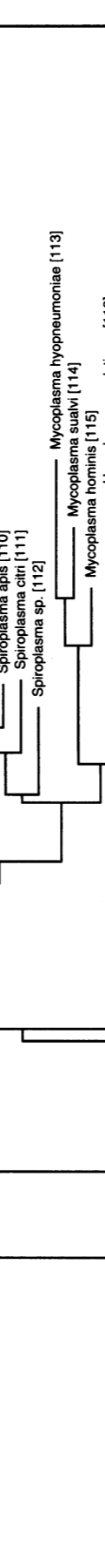
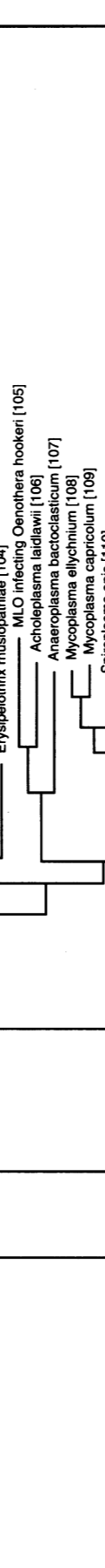
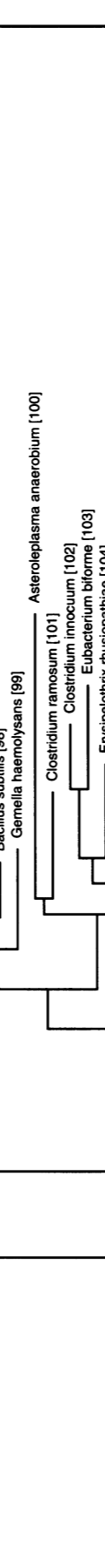
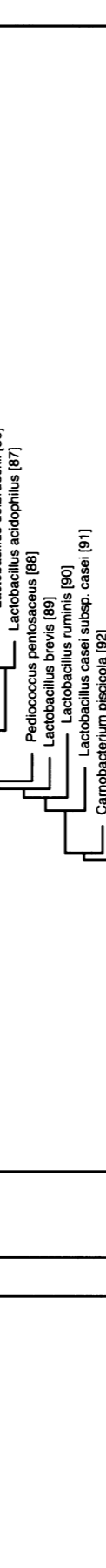
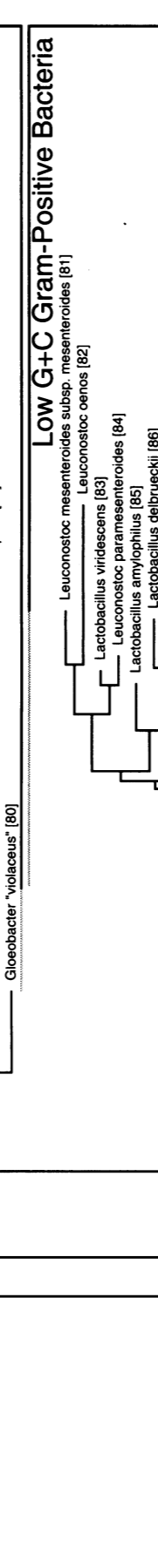
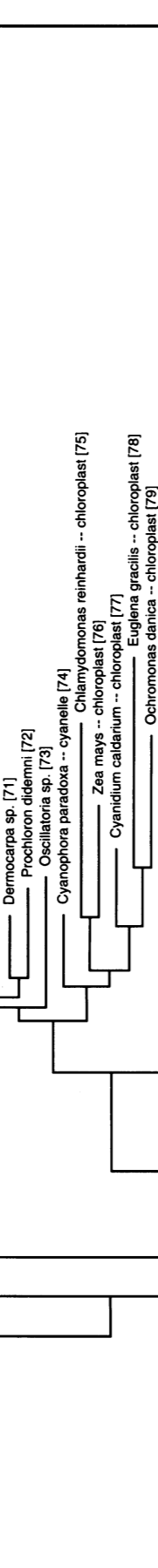
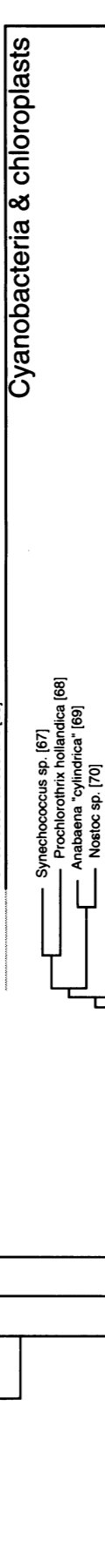
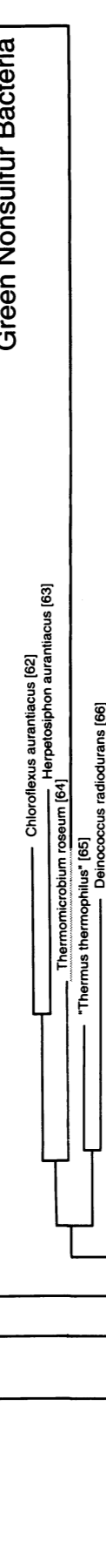
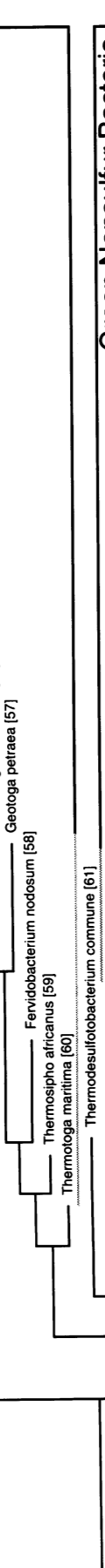
Archaea



Crenarchaeota



Bacteria



On the nature and origin of the eukaryotic cell, we can now inquire about the origins of the various parts of the eukaryotic genome: for what fraction of eukaryotic genes or gene families is the most similar homolog an archaeal gene, for what fraction is it (eu)bacterial, and for what fraction is there no detectable homolog among either *Archaea* or *Bacteria*? Conversely, what fraction of archaeal or (eu)bacterial genes find no recognizable representation in the eukaryotic genome? What nuclear genes are unique to eukaryotes with endosymbiotic organelles and which of these trace back to the expected bacterial lineages? Are there other major, undiscovered, genetic contributions from ancient symbioses, transfers, or fusions?

Considering still earlier events, we can pose questions about the nature of the most recent common ancestor of present-day life: What was its nature? Was it prokaryotic, in the sense of having a genome as complex (and well-defined) as those of extant prokaryotes, or could it have been a progenote, a hypothetical more rudimentary entity in which the link between genotype and phenotype is not yet as precise (accurate) as that seen today (33)? What were the evolutionary dynamics by which this universal ancestor spawned the three primary lineages? Was this typical evolution, or, for example, did it involve so much genetic exchange, gene duplication and divergence, and/or loss of genes, that it is not possible to conceptualize (and analyze) the ancestral state and its subsequent evolution in terms of well-defined lineages (30)?

The origins of key cellular functions can also be addressed: from what genetic roots did photosynthesis (or aerobiosis, or various other forms of metabolism) spring? What genes in other pathways are related to them? To what extent can we trace genes back to a basic aboriginal genetic complement, and what was it?

Fortunately, these questions will be to a significant extent answerable now that biology is moving into the era of genome sequencing. The answers can either be found randomly and anecdotally, or they can be found more quickly if microbiologists learn the history that links all life and use it for guidance.

ACKNOWLEDGMENTS

This work was supported by NSF grant DIR 89-57026 to G.J.O., grants from NASA and NSF (Systematics) to C.R.W., and by the Office of Scientific Computing, U.S. Department of Energy, under contract W-31-109-Eng-38 (R.O.). We are deeply grateful for the opportunity to use both the Touchstone Delta at the California Institute of Technology and the SP-1 at Argonne National Laboratory to run the computations required to develop the phylogenetic tree.

REFERENCES

- Achenbach-Richter, L., R. Gupta, K. O. Stetter, and C. R. Woese. 1987. Were the original eubacteria thermophiles? *Syst. Appl. Microbiol.* **9**:34-39.
- Achenbach-Richter, L., R. Gupta, W. Zillig, and C. R. Woese. 1988. Rooting the archaeobacterial tree: the pivotal role of *Thermococcus celer* in archaeobacterial evolution. *Syst. Appl. Microbiol.* **10**:231-240.
- Cech, T. R., N. K. Tanner, I. Tinoco, Jr., B. R. Weir, M. Zuker, and P. S. Perlman. 1983. Secondary structure of the *Tetrahymena* ribosomal RNA intervening sequence: structural homology with fungal mitochondrial intervening sequences. *Proc. Natl. Acad. Sci. USA* **80**:3903-3907.
- DeLong, E. F. 1992. Archaea in coastal marine environments. *Proc. Natl. Acad. Sci. USA* **89**:5685-5689.
- DeLong, E. F., G. S. Wickham, and N. R. Pace. 1989. Phylogenetic stains: ribosomal RNA-based probes for the identification of single cells. *Science* **243**:1360-1363.
- Edman, J. C., J. A. Kovacs, H. Masur, D. V. Santi, H. J. Elwood, and M. L. Sogin. 1988. Ribosomal RNA sequences show *Pneumocystis carinii* to be closely related to the yeasts. *Nature (London)* **334**:519-522.
- Felsenstein, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.* **17**:368-376.
- Fox, G. E., E. Stackebrandt, R. B. Hespell, J. Gibson, J. Maniloff, T. A. Dyer, R. S. Wolfe, W. E. Balch, R. Tanner, L. Magrum, L. B. Zablen, R. Blakemore, R. Gupta, L. Bonen, B. J. Lewis, D. A. Stahl, K. R. Luehrsen, K. N. Chen, and C. R. Woese. 1980. The phylogeny of prokaryotes. *Science* **209**:457-463.
- Fox, G. E., and C. R. Woese. 1975. 5S RNA secondary structure. *Nature (London)* **256**:505-507.
- Fuhrman, J. A., K. McCallum, and A. A. Davis. 1992. Novel major archaeobacterial group from marine plankton. *Nature (London)* **356**:148-149.
- Gogarten, J. P., H. Kibak, P. Ditttrich, L. Taiz, E. J. Bowman, B. J. Bowman, M. F. Manolson, R. J. Poole, T. Date, T. Oshima, J. Konishi, K. Denda, and M. Yoshida. 1989. Evolution of the vacuolar H⁺-ATPase: implications for the origin of eukaryotes. *Proc. Natl. Acad. Sci. USA* **86**:6661-6665.
- Guthrie, C., and B. Patterson. 1988. Spliceosomal snRNAs. *Annu. Rev. Genet.* **22**:387-420.
- Iwabe, N., K. Kuma, M. Hasegawa, S. Osawa, and T. Miyata. 1989. Evolutionary relationship of archaeobacteria, eubacteria and eukaryotes inferred from phylogenetic trees of duplicated genes. *Proc. Natl. Acad. Sci. USA* **86**:9355-9359.
- James, B. D., G. J. Olsen, J. Liu, and N. R. Pace. 1988. The secondary structure of ribonuclease P RNA, the catalytic element of a ribonucleoprotein enzyme. *Cell* **52**:19-26.
- Kaine, B. P. 1990. Structure of the archaeobacterial 7S RNA molecule. *Mol. Gen. Genet.* **221**:315-321.
- Lane, D. J., and M. L. Collins. 1991. Current methods for detection of DNA/ribosomal RNA hybrids, p. 54-75. *In* A. Vaheri, R. C. Tilton, and A. Balows (ed.), *Rapid methods and automation in microbiology and immunology*. Springer-Verlag, New York.
- Larsen, N., G. J. Olsen, B. L. Maidak, M. J. McCaughey, R. Overbeek, T. J. Macke, T. L. Marsh, and C. R. Woese. 1993. The ribosomal database project. *Nucleic Acids Res.* **21**:3021-3023.
- Larsen, N., and C. Zwieb. 1991. SRP-RNA sequence alignment and secondary structure. *Nucleic Acids Res.* **19**:209-215.
- Margulis, L., and K. V. Schwartz. 1988. *Five kingdoms: an illustrated guide to the phyla of life on earth*, 2nd ed. W. H. Freeman and Company, New York.
- Michel, F., and B. Dujon. 1983. Conservation of RNA secondary structures in two intron families including mitochondrial-, chloroplast-, and nuclear-encoded members. *EMBO J.* **2**:33-38.
- Nishikawa, K., and S. Takemura. 1974. Nucleotide sequence of 5 S RNA from *Torulopsis utilis*. *FEBS Lett.* **40**:106-109.
- Noller, H. F., J. Kop, V. Wheaton, J. Brosius, R. R. Gutell, A. M. Kopylov, F. Dohme, W. Herr, D. A. Stahl, R. Gupta, and C. R. Woese. 1981. Secondary structure model for 23S ribosomal RNA. *Nucleic Acids Res.* **9**:6167-6189.
- Olsen, G. J., H. Matsuda, R. Hagstrom, and R. Overbeek. fast-DNAml: a tool for construction of phylogenetic trees of DNA sequences using maximum likelihood. *Comput. Appl. Biosci.*, in press.
- Pace, N. R., D. A. Stahl, D. J. Lane, and G. J. Olsen. 1985. The analysis of natural microbial populations by ribosomal RNA sequences. *ASM News* **51**:4-12.
- Pace, N. R., D. A. Stahl, D. J. Lane, and G. J. Olsen. 1986. The analysis of natural microbial populations by ribosomal RNA sequences. *Adv. Microb. Ecol.* **9**:1-55.
- Stahl, D. A., and J. W. Urbance. 1990. The division between fast- and slow-growing species corresponds to natural relationships among the mycobacteria. *J. Bacteriol.* **172**:116-124.
- Stanier, R. Y., M. Doudoroff, and E. A. Adelberg. 1970. *The microbial world*, 3rd ed., p. 529. Prentice-Hall, Inc., Englewood Cliffs, N.J.
- Stanier, R. Y., and C. B. van Niel. 1962. The concept of a bacterium. *Arch. Mikrobiol.* **42**:17-35.
- van Niel, C. B. 1946. The classification and natural relationships of bacteria. *Cold Spring Harbor Symp. Quant. Biol.* **11**:285-301.
- Woese, C. R. 1982. Archaeobacteria and cellular origins: an over-

- view. Zentralbl. Bakteri. Hyg. Abt. 1. Orig. C 3:1-17.
31. **Woese, C. R.** 1987. Bacterial evolution. *Microbiol. Rev.* **51**:221-271.
 32. **Woese, C. R., and G. E. Fox.** 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl. Acad. Sci. USA* **74**:5088-5090.
 33. **Woese, C. R., and G. E. Fox.** 1977. Progenotes and the origin of the cytoplasm. *J. Mol. Evol.* **10**:1-6.
 34. **Woese, C. R., O. Kandler, and M. L. Wheelis.** 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. USA* **87**:4576-4579.
 35. **Woese, C. R., L. J. Magrum, R. Gupta, R. B. Siegel, D. A. Stahl, J. Kop, N. Crawford, J. Brosius, R. Gutell, J. J. Hogan, and H. F. Noller.** 1980. Secondary structure model for bacterial 16S ribosomal RNA: phylogenetic, enzymatic and chemical evidence. *Nucleic Acids Res.* **8**:2275-2293.
 36. **Zuckerlandl, E., and L. Pauling.** 1965. Molecules as documents of evolutionary history. *J. Theor. Biol.* **8**:357-366.