

# The Molecular Tree of Life Changes How We See, Teach Microbial Diversity

In using terms such as procaryote in explaining evolutionary history or how life is organized, our textbooks and vocabularies are badly outdated

**Norman R. Pace**

**W**hile our understanding of the diversity and evolution of microbial life is advancing rapidly, our textbooks are not incorporating new information or concepts at nearly the same pace. One important consequence is that textbook explanations of life's organization and evolutionary history are badly outdated. In particular, textbooks fail to recognize that the procaryote concept is wrong in light of what recent biology teaches us and, thus, is conceptually misleading. Put simply, the concept of procaryote is obsolete.

## Toward a Tree of Life

The possibility of relating all organisms has long intrigued biologists. During the mid-19th century, for example, Charles Darwin, Ernst Hae-

ckel, and others compiled elaborate evolutionary trees based on life such as they knew it. However, because they depicted relationships among species based on morphological and developmental properties, they could not include microbial life in any meaningful way. With such criteria, how can a plant be meaningfully compared to an animal or a bacterium?

In general, microbes did not fit into early thought on biological diversity. Little was known about such organisms, and there was no objective way to relate them to one another, let alone to complex eucaryotes. The few physiological properties that have been and still are used to identify microbes provided little and often misleading information with which to deduce relationships.

Those problems were set aside during the 1960s and 1970s with the development of technology for determining gene sequences. In contrast to physiological properties, gene sequences provide an objective metric for evolutionary diversity: the extent of sequence difference between orthologous genes in different organisms is a measure of evolutionary distance. Consequently, comparisons of sequences can be used to measure evolutionary relatedness and to construct phylogenetic trees.

Carl Woese, in the early 1970s, presciently chose to focus on small-subunit rRNA sequences to compare and relate diverse organisms. Ribosomal RNA sequences are ubiquitous and highly conserved, and have become the standard for many comparative phylogenetic studies. Moreover, the technology available to Woese was amenable to comparing rRNA

### Summary

- Comparisons of gene sequences provide an objective view of evolutionary relationships and the course of evolution, in the context of a molecular tree of life.
- Culture-independent, sequence-based identifications of microbes in the environment are dramatically expanding our knowledge of microbial diversity.
- The results of environmental surveys affirm the three-domain model for phylogenetic organization and the course of evolution.
- Experimental results represented by the molecular tree render the concept of "procaryote" obsolete, making it a misleading term, particularly when used in teaching.

*Norman R. Pace is a professor in the Department of Molecular, Cellular and Developmental Biology, University of Colorado, Boulder.*



sequences by compiling catalogs of oligonucleotides that were released when nucleases were used to digest  $^{32}\text{P}$ -labeled rRNA molecules. Early results provided the first outlines of relationships among bacteria and led in 1977 to the discovery of “archaebacteria,” which in 1990 were renamed archaea to emphasize that they are fundamentally different from bacteria. Moreover, when analyzed in this way, organisms fall into three distinct relatedness groups: *Archaea*, *Bacteria*, and *Eucarya* (eucaryotes).

This molecular description of diversity is based on gene sequence comparisons, not properties of particular organisms. Not all gene sequences have tracked with the rRNA sequences in overall evolution. Lateral transfers of genes from one genetic environment to another—a significant driver of evolution—likely account for many of these irregularities. The rRNA genes, however, apparently did not undergo lateral transfers. Moreover, other widely conserved genes of the nucleic acids-based information processing system also track with the rRNA tree. Consequently, the rRNA-based tree represents the evolutionary flow of the genetic machinery, the essence of cells. Thus, large-scale molecular trees should not be viewed as organismic trees, but as the more abstract “lines of descent” of the genetic machinery. Although we may think of dinosaurs as extinct, the line of descent from which dinosaurs emerged is alive today in the form of birds.

### Into the Natural Microbial World

Well more than 99% of microbes from the environment do not thrive in culture and, historically, remained uncharacterized. However, by using molecular cloning and gene sequences to identify microbes, it is now possible to bypass traditional, but often impossible, requirements to study microorganisms in pure cultures. Microbiologists now are exploring the makeup of the natural microbial world, and their results continue to expand dramatically the known extent of microbial diversity.

With molecular technology, rRNA genes are cloned from environmental specimens containing DNA and then sequenced; information about those sequences is used to relate the organisms from particular environments to others whose sequences are known. Some properties of organisms identified solely through rRNA se-

quence analysis can be predicted, depending on how closely they are related to organisms with characterized properties. The presumption here is that the representatives of a phylogenetic group are expected to share many traits. Even if the environmental organisms are not closely related to known ones, however, the rRNA sequences are specific identifiers and provide a basis for the design of tools, such as PCR primers and hybridization probes, with which to study organisms of interest.

The numbers of environmental rRNA sequences in databases now substantially exceed those from cultured organisms, and encompass far more diversity. The basic foundation of the three-domain tree, comprised of *Archaea*, *Bacteria*, and *Eucarya*, has remained solid, while each domain continues to expand in phylogenetic diversity (Fig. 1).

### The Molecular Tree of Life Continues To Expand

A profound lesson from molecular phylogenetics is that all known cellular life is related and has a common origin. The location of the “root” of the universal tree of life, the last universal common ancestral state, cannot be inferred solely from rRNA data. However, other phylogenetic results, as well as biochemical correlates, put the root of the molecular tree on the bacterial line (“origin” in Figure 1). This bacterial origin means that the eucaryal and archaeal lines separated from one another after they separated from the bacterial line. Eucaryotes and archaea thus are more closely related to one another than either is to bacteria, and they are expected to share fundamental properties not found in bacteria.

Many biochemical correlates support this phylogenetic interpretation. For instance, whereas bacteria use sigma factors to control transcription initiation, both archaea and eucaryotes use a different mechanism, TATA-binding proteins. As another example, bacteria wrap their DNA in a variety of basic proteins, while eucaryotes and archaea both use histones.

The details of the major radiations in the domains cannot be resolved with current data and are interpreted in Fig. 1 as unresolved star radiations, or polytomies. Recent molecular analyses of environmental organisms greatly influence the extent of diversity now depicted in the universal tree. Among the bacteria, only 12

phyla (the main relatedness groups of bacteria, sometimes called divisions) were identified in 1987, and all had representative cultured organisms. Now more than 70 phylum-level bacterial lines are known, but only about 20 have representatives that have been cultured. Only seven of the cultured phyla, those that contain human pathogens, have significant representation through culture studies.

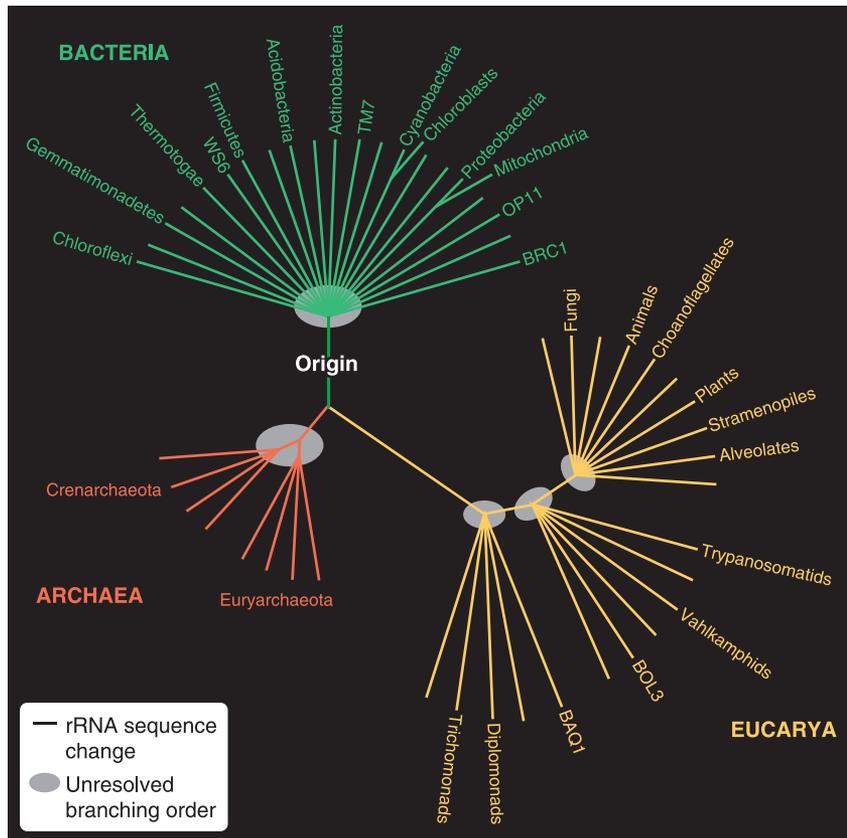
Archaeal phylogeny also proves much more complex than originally thought, as environmental sequences have swamped those from the few cultured versions of these organisms. The molecular observations contradict some previous notions. For instance, based on limited culture studies, the crenarchaeota (Fig. 1) commonly are portrayed as extreme thermophiles. Yet, in the light of environmental sequences, the cultured, thermophilic crenarchaeotes are only one line among many other low-temperature organisms that occur globally. Although these environmental crenarchaeota are among the most abundant kinds of organisms on Earth, their roles in the global ecosystem remain obscure.

**Less Known about the Eucaryal Branch of the Tree and about Broader Concepts**

The large-scale structure of the eucaryotic tree based on rRNA sequences is still unclear. The molecular results confirm that the major organelles, mitochondria and chloroplasts, are bacterial in origin. However, most eucaryal rRNA sequences available for comparison derive from a relatively limited diversity of organisms (metazoa, metaphyta, fungi), and the detail of phylogenetic trees can be distorted by uneven representation of sequences used in constructing such trees.

The topology for eucaryotes (Fig. 1) emerged from the first broad-range eucaryal rRNA sequence comparisons made by Mitchell Sogin of the Marine Biological Laboratory at Woods Hole, Mass., and his collaborators. Environmental sequences bolster that result through discovery of additional, deeply divergent, king-

FIGURE 1



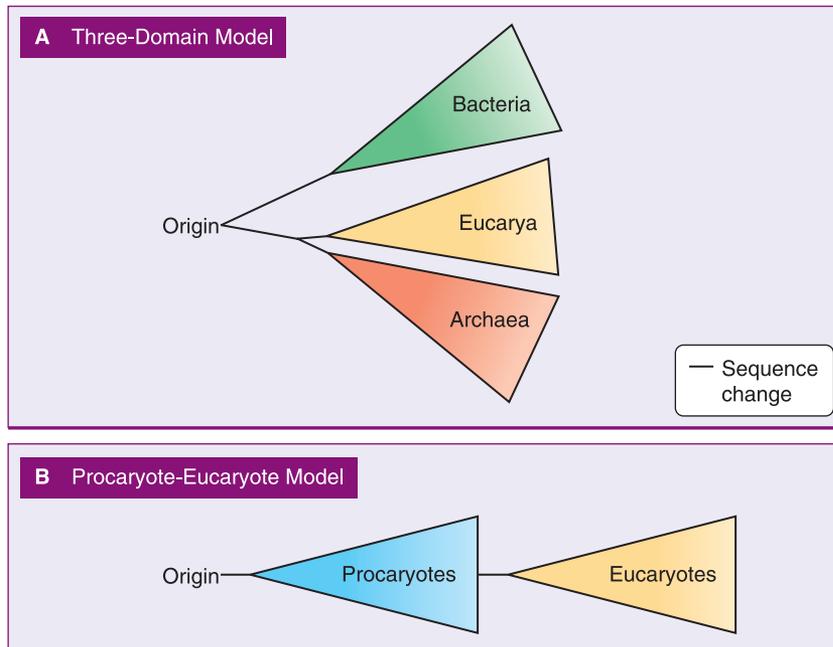
Molecular tree of life based on rRNA sequence comparisons. The diagram compiles the results of many rRNA sequence comparisons. Only a few known lines of descent are shown. The lineages TM7, OP11, WS6, BRC1, BOL3 and BAC1 (and many not shown) are known only from environmental sequences. Shaded zones represent areas in the diagram in which branching orders are not resolved.

dom-level groups, increasing the resolution of phylogenetic calculations.

Regardless of detail, the rRNA perspective indicates that the nuclear line of descent is as old as the archaeal line, meaning that eucaryotes were around since the beginning. Whether the earliest representatives of the eucaryotic line had nuclear membranes becomes moot in the face of the genetic relationships established by DNA sequences. From the rRNA perspective, the general path of eucaryotic evolution seems to have been a basal radiation, one line of which resulted in subsequent radiation(s). One of those lines gave rise to the “crown radiation” of lineages that are most familiar and which include animals, plants, fungi, stamenopiles, and alveolates.



FIGURE 2



Two models for phylogenetic organization and the course of evolution. The wedges represent relatedness groups of organisms.

Researchers continue to identify thousands of new species of microbial life from all three domains, indicating that we are barely scratching the surface of an enormous extent of biological diversity throughout the environment. The emerging field of metagenomics, which involves determining and interpreting natural community DNA sequences, will play a central role as we learn more about how the microbial world influences the biosphere.

The molecular tree of life provides an objective measure of evolutionary relationships. While scientists continue to refine the molecular tree, the essence of the three-domain model for phylogenetic organization and evolution is established (Fig. 2A). Fundamentally, there are three phylogenetic kinds of organisms, representative of the three primary domains. Moreover, none of those primary domains is derived from another. These experimental observations represented by the molecular tree are at sharp odds with the notion of procaryote and eucaryote that currently dominates education in biology. The molecular tree shows that the procaryote-eucaryote model for biological orga-

nization and evolution is both wrong and misleading.

### Procaryote Is An Unfortunate Name and Concept

The term “procaryote” was popularized in the 1960s by Roger Stanier, Michael Douderoff, and Edward Adelberg, in the 2nd edition of their widely used text, *The Microbial World*. Coined for taxonomic purposes and not meant to carry evolutionary implications, this amalgam that combines “pro” (before) with “karyon” (Greek, for nucleus), inevitably invokes an evolutionary model. To many biologists of that period, “procaryote” was not much more than a name change from “monera” at the base of Haeckel’s 19th-century trees and a category for organisms that were little understood. Indeed, some texts freely interchange “monera” and “procaryotes.”

However, the textbook definition of “procaryote,” as an organism without a nucleus, more properly without a nuclear membrane, is negative, and there-

fore not a sound basis for classifying organisms. No one can say what truly is a “procaryote,” only what is not. Consequently, the term “procaryote” attracts different and confusing usages. In one context, it includes extant bacteria and archaea because they are organisms without nuclear membranes. In another context, “procaryote” refers to some long-gone ancestor of eucaryotes. Merging these two notions intellectually muddles the procaryote-eucaryote model for phylogenetic organization with evolutionary theory and thus misleads our students.

New editions of many textbooks incorporate the three-domain molecular tree to varying extents, although generally only in the context of comparing it to other systems for classifying organisms such as the “five-kingdom” or “six-kingdom systems.” Further, textbooks generally fail to convey a key lesson underlying the three-domain molecular tree—namely, that none of the three rRNA sequence-defined domains emerged from another. Instead, textbooks generally fall back on outmoded notions of cellular evolution, showing “procaryotes” giving rise to eucaryotes.

In other ways, biological textbooks are slippery in how they portray “procaryotes” in evolution, and the literature is rife with speculation about eucaryotes forming through various fusions of bacteria and archaea. Based on personal reviews of such textbooks and discussions with educators and students, I venture that a typical student of biology emerges from his or her university with the notions of biological relationships and evolution depicted in Fig. 2B. These include (i) all eucaryotes are specifically related; (ii) all procaryotes are specifically related; and (iii) eucaryotes evolved from procaryotes. These precepts now can and should be weighed in the light of the three-domain molecular tree (Fig. 2A), which emanates from experimental data and is not simply a free-standing hypothesis. Thus, the first of these precepts is correct in that all eucaryotes fall into the domain *Eucarya*, which is the relatedness group determined by sequence comparisons.

However, the second precept about relationships among procaryotes proves false. Indeed, the molecular tree shows there are two fundamentally distinct kinds of noneucaryotes, bacteria and archaea. Moreover, the group of archaea is more closely related to eucaryotes than is the other group, consisting of bacteria.

Finally, the third precept claiming that eucaryotes originate from procaryotes is also false. The molecular tree shows that the eukaryotic nuclear line did not derive from either bacteria or archaea (Fig. 2).

Although biologists earlier embraced the procaryote-eucaryote model as a large-scale means for managing phylogenetic organization and for depicting evolutionary relatedness, that model has proved to be incorrect. Yet, following a half-century of steady usage, the term “procaryote” permeates our journals, texts, and language at all levels. However, language choices for expressing science are critical because we conceptualize that science through language. We must teach—and understand—biology through use of accurate terminology. Continued use of the term “procaryote” perpetuates incorrect concepts about phylogenetic organization and evolution, the very foundations of biologi-

cal thought. “Procaryote” has no place in modern scientific discourse.

### What To Do about It?

Microbiologists need to take the lead in removing procaryote and similar terminology from textbooks and the lexicon of biology. Because microbiologists are closest to the problem, their leadership is needed to rectify misconceptions that this terminology helps to perpetuate. Dropping the term “procaryote” will be difficult for microbiologists because of long conditioning. However, those who are tempted to continue using it risk saddling their students with misconceptions and muddled thinking about important biological problems.

What other terms can be used? For microorganisms in general, I usually use “microbe,” which encompasses the poorly acknowledged microbial eucaryotes. Beyond that, it often is necessary to be more specific about the scientific issues being described. For example, it is far more illuminating to distinguish bacterial from archaeal transcription than it is to lump them into the single category of “prokaryotic transcription.”

How can we handle this issue when use of the term “procaryote” is pervasive in teaching and research? In fact, the discordance between emerging data and traditional thought on deep evolution and relationships among organisms is a wonderful example of how the science of biology is itself an ongoing process. Bringing the subject to the attention of students shows them how new ideas based on emerging experimental evidence can change the ways in which we understand natural processes.

Dealing with the procaryote-eucaryote issue provides them with an example of how scientists evaluate specific models and examine hypotheses in the face of experimental data. Phylogenetic trees and maps of evolutionary relationships are not so difficult to understand. They are abstract, to be sure, but also can be readily interpreted by students. The three-domain concept poses many questions, but it also provides a solid framework for progress toward answering such questions.

### ACKNOWLEDGMENTS

A number of colleagues contributed significantly to this text, particularly Jim Darnell, Corrie Detweiler, Mike Klymkowsky, Carl Woese, and members of the Pace lab. Thanks to Kirk Harris for help with artwork. This article is based on the 2007 Abbott-ASM Lifetime Achievement Award lecture presented at the 2007 Annual Meeting in Toronto, Canada, May 22, 2007.



## SUGGESTED READING

Baumgartner, L., and N. R. Pace. 2007. Current taxonomy in classroom instruction: How to teach the new understanding of higher-level taxonomy. *Science Teacher* 74:46–51.

Pace, N. R. 2006. Time for a change. *Nature* 441:289.

Rappé, M. S., and S. J. Giovannoni. 2003. The uncultured microbial majority. *Annu. Rev. Microbiol.* 57:369–394.

Sapp, J. 2005. The procaryote-eucaryote dichotomy: meanings and mythology. *Microbiol. Mol. Biol. Rev.* 69:292–305.

Sogin, M. L., J. H. Gunderson, H. J. Elwood, R. A. Alonso and D. A. Peattie. 1989. Phylogenetic meaning of the kingdom concept: an unusual ribosomal RNA from *Giardia lamblia*. *Science* 243:75–77.

Woese, C. R. 1994. There must be a procaryote somewhere: microbiology's search for itself. *Microbiol. Mol. Biol. Rev.* 58:1–9.

Woese, C. R. 2004. A new biology for a new century. *Microbiol. Mol. Biol. Rev.* 68:173–186.

COULD BE VALUABLE



AMERICAN  
SOCIETY FOR  
MICROBIOLOGY

## American Board of Medical Microbiology (ABMM)

American College of Microbiology

### Certification in Medical and Public Health Microbiology, Doctoral-Level

- Documents your value and credibility with physicians
- Highest credential available to practicing medical laboratory microbiologists
- Benchmark your skills
- Enhance your credibility in the medical arena
- Recognized under CLIA '88

For more information please visit:

[www.asm.org/college](http://www.asm.org/college)

## Certify your worth.

Don't make people guess how valuable you are.