Environmental Diversity of Bacteria and Archaea

EDWARD F. DELONG\textsuperscript{1} AND NORMAN R. PACE\textsuperscript{2}

\textsuperscript{1}Monterey Bay Aquarium Research Institute, 7700 Sandholdt Rd., Moss Landing, California 95039, USA; E-mail: delong@mbari.org

\textsuperscript{2}Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, Colorado 80309-0347, USA

Abstract.—The microbial way of life spans at least 3.8 billion years of evolution. Microbial organisms are pervasive, ubiquitous, and essential components of all ecosystems. The geochemical composition of Earth’s biosphere has been molded largely by microbial activities. Yet, despite the predominance of microbes during the course of life’s history, general principles and theory of microbial evolution and ecology are not well developed. Until recently, investigators had no idea how accurately cultivated microorganisms represented overall microbial diversity. The development of molecular phylogenetics has recently enabled characterization of naturally occurring microbial biota without cultivation. Free from the biases of culture-based studies, molecular phylogenetic surveys have revealed a vast array of new microbial groups. Many of these new microbes are widespread and abundant among contemporary microbiota and fall within novel divisions that branch deep within the tree of life. The breadth and extent of extant microbial diversity has become much clearer. A remaining challenge for microbial biologists is to better characterize the biological properties of these newly described microbial taxa. This more comprehensive picture will provide much better perspective on the natural history, ecology, and evolution of extant microbial life. [Archaea; bacteria; biodiversity; evolution; microbial; phylogeny.]

Historically, microbial biology has developed along research lines largely independent of other biological disciplines, mainly for technical reasons. Simply put, the natural microbial world, unlike that of visible organisms, cannot be observed in great detail by direct methods. The independent development of microbial biology encouraged the birth of some fields, a prime example being molecular biology. Meanwhile, however, appreciation of ecology and evolution in microbial systems has lagged far behind parallel developments in mainstream biology. Microbial biologists simply lacked most of the classical tools, concepts, and theory available to paleontologists, systematists, ecologists, and evolutionary biologists. Advances in molecular phylogenetics, macromolecular sequencing techniques, and emerging genomic technologies, however, are changing the playing field dramatically for microbial biologists (Pace, 1997). Now, serious inroads are being made in microbial ecology and evolution, paved largely by the application of comparative molecular phylogenetic methods.

The Three Domains and Microbial Evolution

Until recently, higher-order evolutionary relationships among microorganisms were virtually unknown and undescribed. Few conspicuous morphological features can be used to systematically differentiate microorganisms or infer their evolutionary relationships at higher taxonomic levels. Additionally, the microbial fossil record is neither extensive nor informative enough to provide much insight into ancestral microbial life. Until the mid-1960s, microbial biologists had to be content with simply distinguishing “prokaryotes” (which do not possess membrane-bound nuclei) from eukaryotes (which have true nuclear organelles). This situation changed dramatically when Zuckerkandl and Pauling (1965) pointed out that molecules can serve as documents for evolutionary history. Evolutionary relationships could now be deduced from sequence differences observed between homologous macromolecules. For the first time, universal comparisons of homologous macromolecular features from virtually all (known) cellular lifeforms became a practical reality.

Carl Woese was the first to fully exploit the power of molecular phylogenetics for inferring evolutionary relationships between “kingdoms” (deeply related groups), as he and his colleagues sought to create a unified picture of evolutionary relationship among prokaryotes in the 1970s (Woese and Fox, 1977). With remarkable insight, he
selected perhaps the single most optimal macromolecule for establishing deep relationships, 16S ribosomal RNA (rRNA). Woese realized that the optimal macromolecule for constructing global phylogenies should have a universal distribution, high conservation, some moderate variability, and minimal lateral genetic transfer. Initially, catalogues of rRNA oligonucleotides were used to infer relationships among disparate phylogenetic groups. Later, advanced nucleic acid sequencing techniques allowed direct acquisition and comparison of rRNA sequences. One of the first and most dramatic results of Woese’s studies was the discovery of a new microbial kingdom, the Archaea (then called archaebacteria; Woese et al., 1978). These anucleate microbes are as evolutionarily distant from common bacteria, as they are from eukaryotes. The representatives of archaea known in the early 1980s were a fairly bizarre collection of microbes: sulfur-respiring thermophiles, extreme halophiles, and obligately anaerobic methanogens. The evolutionary distinctiveness of the Archaea eventually showed that all known life can be organized into three major Domains: Eucarya (all eukaryotes), Archaea, and Bacteria (Woese et al., 1990). The fruit of Woese’s efforts, a universal tree that outlines the evolutionary relationships of all life, is shown in Figure 1. Analyses of most macromolecules involved in nucleic acid-based information processing (e.g., DNA replication, transcription, translation) yield this three-domain topology. Phylogenetic analyses of metabolic and regulatory genes, on the other hand, do not result in similar consistent topologies (Brown and Doolittle, 1997; Doolittle, 1999).

The three-domain tree is simple in form but profound in its implications. A casual glance at the universal rRNA tree (Fig. 1) shows that the lion’s share of phylogenetic diversity resides in the microbial world, whereas macroscopic organisms occupy small, terminal nodes on the tree of life. Considering the 3.8 billion years over which microbial life has evolved, this is perhaps not so surprising. Another brief look at the tree in Figure 1 reveals the endosymbiotic origins of several organelles—chloroplasts within the cyanobacterial cluster (near Synechococcus), and mitochondria within the alpha Proteobacteria lineage (near Agrobacterium).

A major point to recognize when considering molecular phylogenies is that the genealogy of the organisms is extrapolated from the evolutionary trajectory of single genes. Depending on the gene used, its particular history (including lateral transfer), and the magnitude of evolutionary distances considered, these extrapolations will vary in accuracy. Despite all of the uncertainties and artifacts associated with single-gene phylogenies, the major features of Woese’s rRNA tree—and its lessons—have overriding relevance for contemporary biology. Current genomic studies indicate that lateral gene transfer may have played a greater role in the evolution of major lineages than was previously appreciated (Doolittle, 1999). However, rRNA, and other core genes involved in information transfer, and therefore appear not to have been as extensively laterally exchanged, provide the most coherent framework for understanding and interrelating the main evolutionary branches on the extant tree of life (Ochman et al., 2000).

Cultivation-Independent Survey of Extant Microbial Diversity

Until recently, assessment of naturally occurring microbial diversity was an impossible undertaking, simply for lack of appropriate methods. The problem of accurately describing naturally occurring microbial assemblages has been extensively discussed and reviewed (Staley and Konopka, 1985; Pace et al., 1986; Amman, 1995; Pace, 1997) and is largely intertwined with the historical development of microbiology. The crux of the problem is this: Pure culture techniques, despite their tremendous utility, are inadequate for describing naturally occurring microbial assemblages. Microorganisms commonly recovered by standard microbiological procedures are not generally representative of the assemblages from which they originate. For a vast number of microbial species, appropriate media and conditions for growth are simply not well-developed, available, or practically feasible. In the past, microbial biologists depended nearly exclusively on the isolation and cultivation of individual microbial strains from the environment. Although useful, however, pure culture isolation obliterates the natural assemblage, and the predominant microbes are frequently, even generally, invisible to the approach.

In the early 1980s, developments in comparative molecular phylogenetics were
FIGURE 1. Universal phylogenetic tree based on small-subunit rRNA sequences. “Universal” unrooted phylogenetic tree showing representative taxa was inferred by maximum likelihood analysis of 1,620 homologous positions of sequences of small subunit rRNAs from the indicated organisms or environmental clones. Numbers indicate percentage of bootstrap resamplings that support indicated branches in maximum likelihood (before slash) or maximum parsimony (after slash) analyses for only those groups that attained >60% support with at least one of the two methods. Analyses of duplicated protein genes have placed the root of the tree on the branch at the base of the bacterial line (Iwabe et al., 1989). Evolutionary distance (sequence changes) between the species shown is read along line segments. The scale bar corresponds to 0.1 changes per nucleotide. pSL50, pSL4, pSL22, pSL12, pJP27, pJP78, and marine SBRAR5 represent rRNA sequences obtained directly from environmental samples. [Figure from Barns et al., 1996; reproduced with permission.]
instrumental in removing the roadblocks that prevented accurate description of natural microbial diversity (Stahl et al., 1984; Pace et al., 1986). The cultivation-independent approach involves recovery of phylogenetically informative gene sequences from nucleic acids extracted directly from naturally occurring microbial biomass. Phylogenetically informative gene sequences extracted from mixed microbial populations are isolated as DNA clones that are then sorted and sequenced. Molecular phylogenetic comparisons provide phylogenetic identification of individual population constituents. Small subunit rRNA genes are the most commonly used phylogenetic markers to date, because of their ubiquity and conserved nature.

The above-described methods also provide phylogenetically based markers that can be used to tag specific phylogenetic types of organisms for identification. Small subunit rRNAs, the targets for these nucleic acid-based hybridization probes, have proven extraordinarily useful for such "molecular tagging" approaches. Relatively high amounts of intracellular rRNA provide abundant target for phylogenetic identification of individual cells by using fluor-labeled oligonucleotide probes (DeLong et al., 1989; Amann et al., 1995). Individual microbial cells can now be stained with color-coded probes and identified with fluorescence microscopic techniques. Nested suites of probes specific for different taxonomic levels (e.g., domains, genera, species) can now be designed, allowing hierarchical taxonomic dissection of complex, naturally occurring microbial assemblages (Amann et al., 1995).

Cultivation-independent molecular phylogenetic approaches, like any other methodologies, have their own specific pitfalls and biases (for a recent review and detailed discussion, see Wintzingerode et al., 1997). Potential biases include differential cell lysis, preferential DNA recovery from specific cell types, and several analytical complications introduced by using the polymerase chain reaction (Reysenbach et al., 1992; Suzuki and Giovannoni, 1996; Wang and Wang, 1996; Wintzingerode et al., 1997; Poltz and Cavanaugh, 1998). Nonetheless, results from culture-independent studies using different techniques, with each expected to produce different biases, tend to yield similar results (Schmidt et al., 1991; Mullins et al., 1995; Béjà et al., 2000). Additionally, the great concordance between many independent studies conducted in similar habitats tends to support the conclusions about microbial diversity and distribution provided by cultivation-independent approaches.

Cultivation-independent approaches such as those described above have invigorated the field of microbial ecology. Phylogenetic comparison of rRNA genes retrieved directly from the environment has fast become the standard for surveying natural microbial diversity (Amann et al., 1995; Pace, 1997). The approach has led to the discovery of many novel microbial taxa, ranging from new species, to new phyla, even new "kingdoms"! And these newly recognized microbes are not minor players in the environment: They often represent the major taxa present in both terrestrial and marine ecosystems. Cultivation-independent surveys have greatly expanded the known phylogenetic variety of known microbial species (Pace, 1997). These studies suggest that the phenotypic properties of many of the most abundant microbes inhabiting Earth remain to be determined.

**Cultivation-Independent Surveys of the Domain Bacteria**

In 1987, Carl Woese published a benchmark paper in microbial biology, the first comprehensive synthesis of bacterial evolution placed in the context of all lifeforms (Woese, 1987). In his treatise, Woese outlined the evolutionary relationships of the major bacterial phyla, as inferred from rRNA sequence data. The 12 major bacterial divisions identified still represent most of the taxa that can be readily cultivated and characterized by using cultivation methods. Nevertheless, recent cultivation-independent molecular surveys have revealed that the bacterial domain consists of many more divisions, having few or no cultured representatives. Figure 2 compares the view of bacterial phylogenetic representation in 1987 with the current view derived from cultivation-independent studies. The current tree of bacteria now contains >40 phylogenetically well-resolved bacterial divisions (Pace, 1997; Hugenholtz et al., 1998). The newly discovered groups within the domain Bacteria, most with no cultivated representatives, demonstrate that the microbial species in our
Figure 2. Diagrammatic representation of currently known major groupings in the bacterial domain, inferred from rRNA sequences. Division-level groupings of two or more rRNA sequences are depicted as wedges. Divisions that have cultivated representatives are shown in black; divisions represented only by rRNA sequences retrieved from environmental samples are shown in white. The small tree in the upper right is an outline of known bacterial divisions in 1987, as compiled by Woese (1987).

culture collections provide only a skewed and incomplete picture of extant microbial diversity.

Cultivation-Independent Surveys of the Domain Archaea

Archaea, although prokaryotic in cellular ultrastructure, are evolutionarily quite distant from their microbial cousins, Bacteria. At the time of their discovery, known and cultured Archaea (then archaebacteria) seemed to be an odd, perhaps rare, collection of organisms, united mainly by the evolutionary heritage indicated by rRNA genes and a few other molecular features. The known habitats of Archaea were hypersaline brines (extremely halophilic archaea), geothermal environments (hyperthermophilic archaea), and strictly anoxic habitats (methanogens). In common parlance, cultivated Archaea were, without exception, considered to be “extremophiles”. One of the two main lineages of Archaea, the Crenarchaeota, was especially notorious for growth at high temperatures. All cultivated Crenarchaeota originate from extremely hot environments, including hydrothermal vents and geothermal springs. Many are hyperthermophiles, requiring temperatures greater than 80°C for optimal growth.

It was surprising, then, when culture-independent surveys revealed an abundance of archaea in many diverse habitats. The discovery of widespread diversity of archaea in “nonextreme” habitats is one of the particularly striking findings of culture-independent surveys. Previously, there was no reason to believe that Archaea contributed significantly to the ecology of aerobic marine or terrestrial habitats. Yet now it is apparent that two major groups of Archaea are common and very abundant components of marine plankton and elsewhere.
Shallow and deep marine waters at polar, temperate, and tropical latitudes; the guts of abyssal sea cucumbers; and marine sediments all show evidence of these cold-adapted cousins of hyperthermophilic archaea (DeLong, 1998a, b; see Group 1a, Fig. 3). Archaea are now known to be commonly present in the marine environment at cell densities of $1 \times 10^5$/ml, at depths ranging from 100 m to the bottom of the abyss (DeLong et al., 1999a). At these cell densities, archaea represent ~20% or more of all microbial cells in the oceans, a habitat where 10 years ago they were not thought to exist!

Do these abundant archaeal cells, detected by rRNA gene surveys, actually represent active, authochthonous community members?

**Figure 3.** Diagrammatic representation of major groupings in the archaeal domain, inferred from rRNA sequences. Major archaeal divisions are depicted as wedges. Divisions that have cultivated representatives are shown in black; divisions represented only by rRNA sequences retrieved from environmental samples are shown in white. The predominant habitats of the uncultivated archaeal groups are indicated in italics.
Apparently yes, according to several lines of evidence. Fluorescence in situ hybridization microscopy of planktonic archaea shows intact (and sometimes dividing) cells that contain appreciable amounts of rRNA, cells that reach an abundance maximum (30% of the total planktonic microbial cell population) at specific depths in the water column (DeLong et al., 1999a); these data are suggestive of an active archael population. In Antarctica, surface-water archaea peak in relative abundance in late winter (up to 20% of the total planktonic microbial population), decrease to undetectable numbers during the austral summer (Murray et al., 1998), and again increase with the onset of the subsequent austral winter. These data are again suggestive of an active and dynamic population, one that in Antarctic waters has a distinct, seasonal cycle. Biomarkers of planktonic archael metabolism (specifically, archael-unique tetraether lipids) are also detectable and accumulate in planktonic and sedimentary marine habitats, a distinct biogeochemical influence of the cold-water archael on their surrounding environment (Hoefs et al., 1997; DeLong et al., 1998). Finally, a specific symbiotic association between a marine sponge and a strain of cold-water archael indicates that these microorganisms have radiated into many different habitats, including associations with metazoan hosts (Preston et al., 1996). All the above data strongly suggest that the newly detected marine archael are active, dynamic, and likely to have marked impacts and interactions with surrounding habitats and biota.

Since their initial detection in marine plankton, evidence for a widespread distribution of new types of archael has been further extended in forest and agricultural soils, deep subsurface environments, freshwater lake sediments, deep sea sediments, and in association with certain metazoan species (Pace, 1997; Delong, 1998a, b). Several of the new archael groups are now very commonly encountered in culture-independent ecological surveys (Fig. 3). Distant relatives of thermophilic Crenarchaela (Group C1a, C1b, C1c; Group C2, C3, C4) are found in many low-temperature terrestrial and marine habitats. Two recently discovered archael types fall within the other main branch of archael, the Euryarchaela. These uncultured groups (Group E2 and E3 in Fig. 3) are found predominantly in marine plankton and sediments, respectively. Many other yet uncultivated lineages continue to be discovered, as ongoing culture-independent surveys probe unexplored microbial habitats.

**CONCLUSION**

Despite the impact of culture-independent rRNA gene surveys on microbial ecology, phylogenetic information provided by a single gene (e.g., rRNA) is usually insufficient for inferring the physiology or ecological significance of organisms known only by rRNA sequence. To a large extent, the lack of alternatives to cultivation has severely limited the abilities of microbial biologists to characterize naturally occurring microbes. The unexpected phylogenetic breadth and diversity of many of the newly discovered microbial lineages imply novel evolutionary innovations, new phenotypic properties, and unanticipated ecological roles. New approaches for more comprehensive biological characterization of these novel microbial taxa are desperately needed to fully appreciate the significance of extant microbial life on Earth.

Recent developments in genome science now offer substantial promise for further characterizing uncultivated microbial species (Shizuya et al., 1992; Stein et al., 1996; Schleper et al., 1998; Vergin et al., 1998; DeLong et al., 1999b; Béjà et al., 2000; Rondon et al., 2000). Using bacterial artificial chromosomes (BACs), one can retrieve genome fragments larger than 100 kb from uncultivated microbial species. BAC clones prepared from the chromosomal DNA of uncultured microbes now provide the raw data necessary to dissect the genomes, and reconstruct the biochemical pathways, of naturally occurring microbes. Coupled with the high-throughput approaches developed for standard genome projects, large amounts of previously unavailable data are now becoming readily accessible. Soon, comparative information on the genome structure, content, and organization will become available from some of the most evolutionarily divergent lifeforms on our planet. This information promises to expand our understanding of the evolution and ecology of microbes remarkably. These empirical data should help catalyze new syntheses that include models and theory to help explain the observed ecological patterns and evolutionary processes of microbial life on Earth.
ACKNOWLEDGMENTS

We thank Carl Woese for raising general awareness in the area of microbial evolution, and numerous students and colleagues who have contributed greatly to the work and perspective described here. We also thank Matt Kane for inviting us to participate in this symposium and two anonymous reviewers for their comments.

REFERENCES


SUZUKI, M. T., AND S. J. JOVANNONI. 1996. Bias caused by template annealing in the amplification of mixtures


Received 9 June 2000; accepted 25 July 2000

Associate Editor: M. Kane