Microbial Diversity and Bioprospecting
Edited by Alan T. Bull

Chapter 11

Soils—the Metagenomics Approach

JO HANDELSMAN

THE HISTORY OF SOIL BIOLOGY

Soil is a living organ of the Earth. Until recent times, large and small cultures throughout the world recognized this truth, expressing it through a deep spiritual relationship with the soil. The ancient Greeks worshipped soil through the goddesses Gaia and Demeter, the Germans through the goddess Ertha, and the Native Americans through Mother Earth, and the Old Testament describes the soil as a source of healing: "The Lord hath created medicines out of the earth; and he that is wise will not abhor them." (*Ecclesiasticus,* XXXVIII, 4.)

Advances in science, medicine, and agriculture during the 20th century provided numerous reasons to continue to venerate the soil, well beyond the basic necessity of food production. Discoveries during this century about soil bacteria led to development of life-saving drugs, including antibiotics, antitumor agents, and immunosuppressants. And yet, just as these advances grew more impressive, urbanization and other global economic influences tended to disassociate many people from an intimate association with the environment. Gradually, we lost appreciation of our reliance on the soil, an ignoble decline typified by common use of the term "dirt." But that trend may be about to be altered by the renaissance in soil biology that presents the soil as one of the last great frontiers available for discovery and bioprospecting. What we are poised to discover, in the form of knowledge and medicinal and agricultural chemistry, may well restore the soil to a revered position in the now-global society. This renaissance is driven by the tools of molecular biology, which enable exploration of the life and chemistry of the soil in new ways. One emerging tool that draws together traditional soil biology and modern molecular biology is metagenomics, which entails analysis of the collective genomes of an assemblage of microorganisms.

The past 100 years of microbiological research has yielded experimental evidence verifying the historical belief that soil is the site of many life-sustaining processes. Microorganisms active in the soil are largely responsible for the biogeochemical cycles that support life on Earth (Paul and Clark, 1996). Microorganisms mediate the nitrogen cycle, providing the only biological route for atmospheric nitrogen to enter living systems through nitrogen fixation, oxidizing it further through nitrification, and recycling it into the atmosphere by denitrification. Similarly, microorganisms play a central role in the carbon and sulfur cycles and in the oxidative and reductive transformation of metals such as iron and mercury. Soil microorganisms also influence the health of plants and animals. Although they tend to receive more attention for the diseases they incite, microorganisms are likely more significant for their role in maintaining the health of plants and animals by protecting them from attack by other microorganisms, providing vitamins and other nutrients, and influencing developmental processes.

Although it is easy to celebrate these positive and essential activities, microorganisms exist within a carefully balanced natural system of checks and balances, and when disrupted through human agency can produce harmful environmental effects. One classic example is acid mine drainage, a microbially mediated process that generates a toxic waste typified by acid production, generating pH's below 1 or 0, and high soluble metal concentrations that can pollute waterways (Edwards et al., 1999; Schleper et al., 1995). The disruption of the Earth and exposure of coal or copper mines to oxygen sets the stage for the microbial activity that produces this highly damaging waste.

The biological processes in soil are executed in a complex physical and chemical environment (Fig. 1). The physical structure of soil is varied and dynamic. Soil particles vary in size from 1 to 1,000 μm, and the

Jo Handelsman • Department of Plant Pathology, University of Wisconsin, Madison, WI 53706.

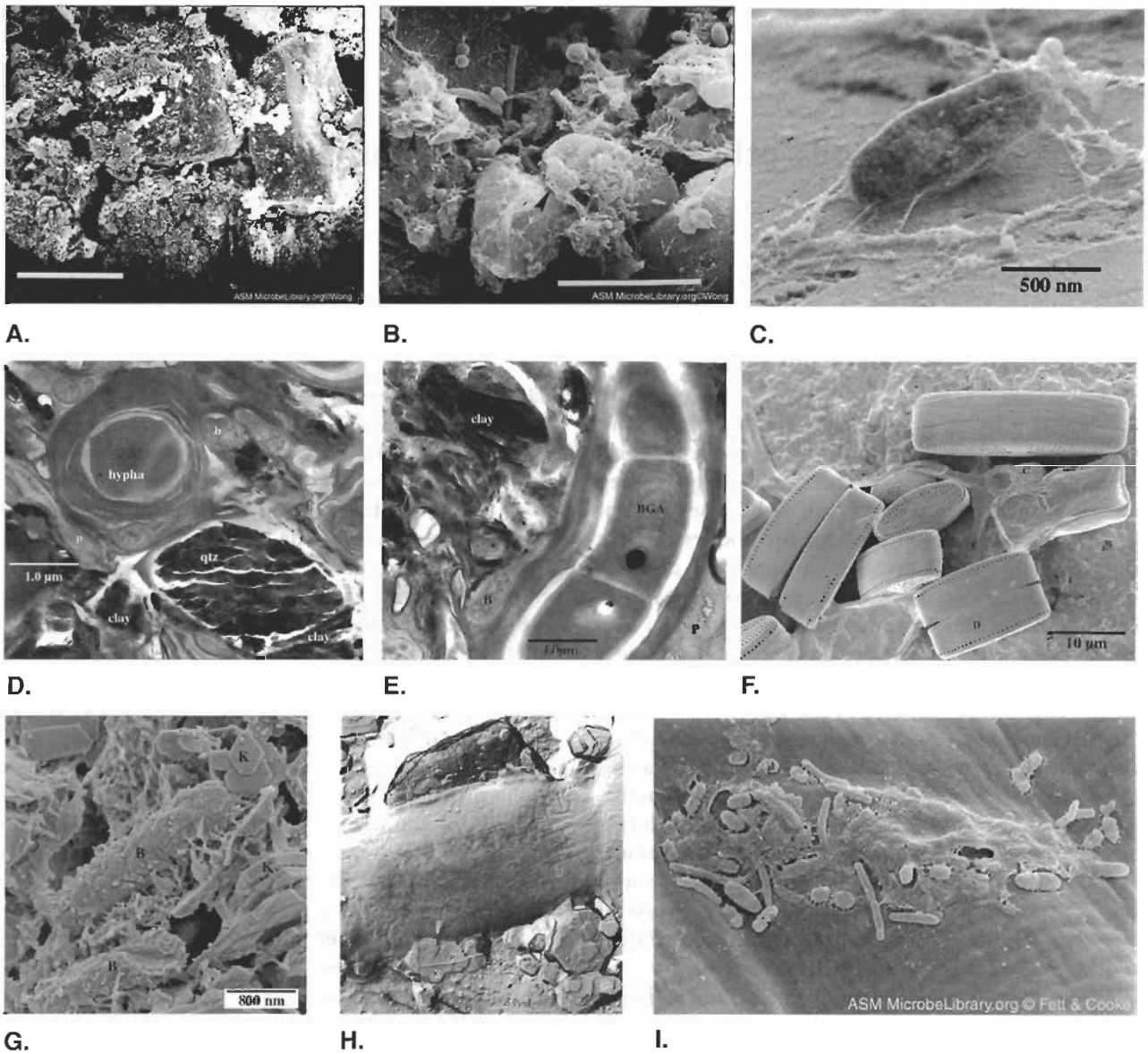

Figure 1. Microorganisms in soil, on minerals, and on plant surfaces.
(A) and (B) Scanning electron micrographs of the naturally occurring biofilm on sand grains in the clog mat of a septic system infiltration mound. The biofilm is composed of mineral particles, a variety of microorganisms, and a network of slime, or glycocalyx, that binds the microorganisms and particles together. Image (A) scale bar is 150 μm. Image (B) scale bar is 4.3 μm. Copyright Amy C. Lee Wong. Licensed for use, ASM MicrobeLibrary (linked to http://www.microbelibrary.org). (C) Unidentified bacterium attached to a feldspar surface by extracellular polymers. High-resolution, low-voltage cryoscanning electron micrograph of high-pressure frozen, freeze-fracture/sublimed culture sample. (D) Microbial soil assemblage consisting of extensive extracellular polymer networks (p), anhedral clay minerals (clay), quartz (qtz), bacteria (b), and fungi. Energy filtered transmission electron micrograph of ultrathin section of high-pressure frozen, freeze-substituted, undisturbed soil sample. (E) Microbial soil assemblage consisting of extensive extracellular polymer networks (P), anhedral clay minerals (clay), bacteria (B), and filamentous cyanobacteria (BGA). Energy-filtered transmission electron micrograph of ultrathin section of high-pressure frozen, freeze-substituted, undisturbed soil sample. (F) Diatoms (D), cyanobacteria (C), and fungal hyphae (F) inhabit the surface of a quartz grain from a spring seep sandstone outcrop near Mount Horeb, Wisconsin. High-resolution, low-voltage scanning electron micrograph. (G) *Pseudomonas fluorescens* (B) bound to euhedral kaolinite (K) by extracellular polysaccharides (arrows). High-resolution, low-voltage cryoscanning electron micrograph of high-pressure frozen, freeze-fracture/sublimed culture sample. (H) *Escherichia coli* (B) attached to euhedral kaolinite (k) by f-type sex pili (P). Transmission electron micrograph of a propane cryojet-frozen, freeze-etch Pt replica. (Micrographs in panels C to H courtesy of William W. Barker, The College of Letters and Science, University of Wisconsin-Madison.) (I) Micrograph of naturally occurring biofilm on a plant surface (alfalfa sprout hypocotyl). The size of the bacteria ranged from 0.4 μm in diameter (cocci) to 0.2 × 1.7 μm (rods). Copyright Peter Cooke and William Fett. Licensed for use, ASM Microbe Library (linked to http://www.microbelibrary.org).

moisture content and chemistry of the soil affect particle size. Particle size, in turn, affects water flow, gas exchange, and temperature gradients in soil. The chemical composition of soil is derived from a combination of its geologic and biologic origins. The bedrock contributes to the inorganic mineral fraction, and geologic events influence the ion exchange capacity, weathering rates, and particle size. The plants that grow in a soil and the microorganisms that live on them and decompose them create the organic fraction of soil, which is highly complex and changes rapidly. The soil, then, is a dynamic interplay of physical forces with chemical substrate and the life that grows in it. It changes over short and long time scales. The moisture content, structure, and nutrient solubility of a soil can change dramatically in the course of a day. Over years and centuries, the organic and mineral fractions can change substantially. Many of these changes are driven by microorganisms; all of them affect the structure and function of the microbial communities in soil (Brady and Weil, 2002).

It is this dynamic complexity of soil that has challenged and excited soil biologists for more than a century. The questions about soil are complex, the opportunities are profound, and the influence of new knowledge is expansive, constituting a broad frontier that intersects with many areas of science. Although we know vastly more about soil now than in the 19th century, the unknown is far greater than the known. The interplay between the known and the unknown and the integration of diverse sciences to dissect soil biology is reflected in the careers, discoveries, and beliefs of some of the most prominent microbiologists of the modern era. The impact of soil biology on human welfare has been molded by the integration of sciences by scientists of diverse origins and beliefs. Selman Waksman is remembered, for instance, for his discovery of streptomycin in the 1940s, hailed at the time as the ultimate weapon against the "White Plague," as tuberculosis was then known. Waksman was a biologist interested in the ecology of soil microorganisms, and as early as 1916 began publishing both a taxonomy of soil fungi and his findings regarding the effects of protozoa on populations of soil bacteria. He did not set out to find a medical miracle, but his explorations in basic soil biology led directly to the discovery of the actinomycete antibiotics that became that miracle (Waksman and Curtis, 1916; Waksman and Foster, 1937).

Waksman was a scientist who valued knowledge and appreciated how much was known about soil biology. His optimistic celebration of "the known" is sharply juxtaposed against the awe with which some of his contemporaries approached "the unknown." Waksman believed that the microbiology of the soil could be described in its entirety relatively quickly: "A large body of information has accumulated that enables us to construct a clear picture . . . of . . . the microscopic population of the soil . . ." (Waksman and Starkey, 1931).

Waksman's contemporary, Francis Clark, by contrast, considered the true nature of soil to be unknowable because of the undiscovered extent and myriad functions of microorganisms. Clark, a medical microbiologist, reminded his readers of the salient observations of the 18th-century botanist Linnaeus, who suggested that the overwhelming complexity of the microbial world might be beyond the reach of tidy classification: "Linnaeus . . . recognized the existence of microscopic forms of life but skirted a taxonomic quagmire by simply placing all microbes in a group designated 'Chaos'" (Paul and Clark, 1989).

This contrast, between what is known or unknown, knowable or unknowable, introduces one of the most dramatic recent advances in soil science. For decades, soil microbiology has concentrated on cultured organisms. These studies, conducted through the essential but limiting screen of the petri dish, have produced significant advances. The advent of new molecular and microscopic methods has fostered a growing realization, however, that there is much to be known about soil microorganisms that cannot be elucidated through traditional approaches because the vast majority of soil microorganisms cannot be cultured. Thus, separating that unknown universe from Linnaeus' centuries-old description of "chaos" requires new methods.

SOIL BIOLOGY AND THE ORIGINS OF METAGENOMICS

Only a small minority of the microorganisms living in soil are readily culturable on standard media. Microscopic studies, analysis of DNA complexity, and studies of species richness with culture-independent methods indicate that between 0.1 and 1% of the viable prokaryotes in soil grow in a standard culturing experiment (Torsvik et al., 1990, 1996; Whitman et al., 1998). The most remarkable aspect of these studies is that the as-yet-uncultured bacteria are highly diverse. Reassociation of DNA from the bacterial component in soil indicates that the complexity of the bacterial community is comprised of at least 4,000 distinct genomes, which exceeds the complexity estimated by culturing by 200-fold (Torsvik et al., 1990). PCR amplification, cloning, and sequencing of rRNA genes from soil have led to a partial understanding of the species diversity (Pace, 1996; Hugenholtz et al., 1998) (Fig. 2). Phylogenetic analysis indicates that the

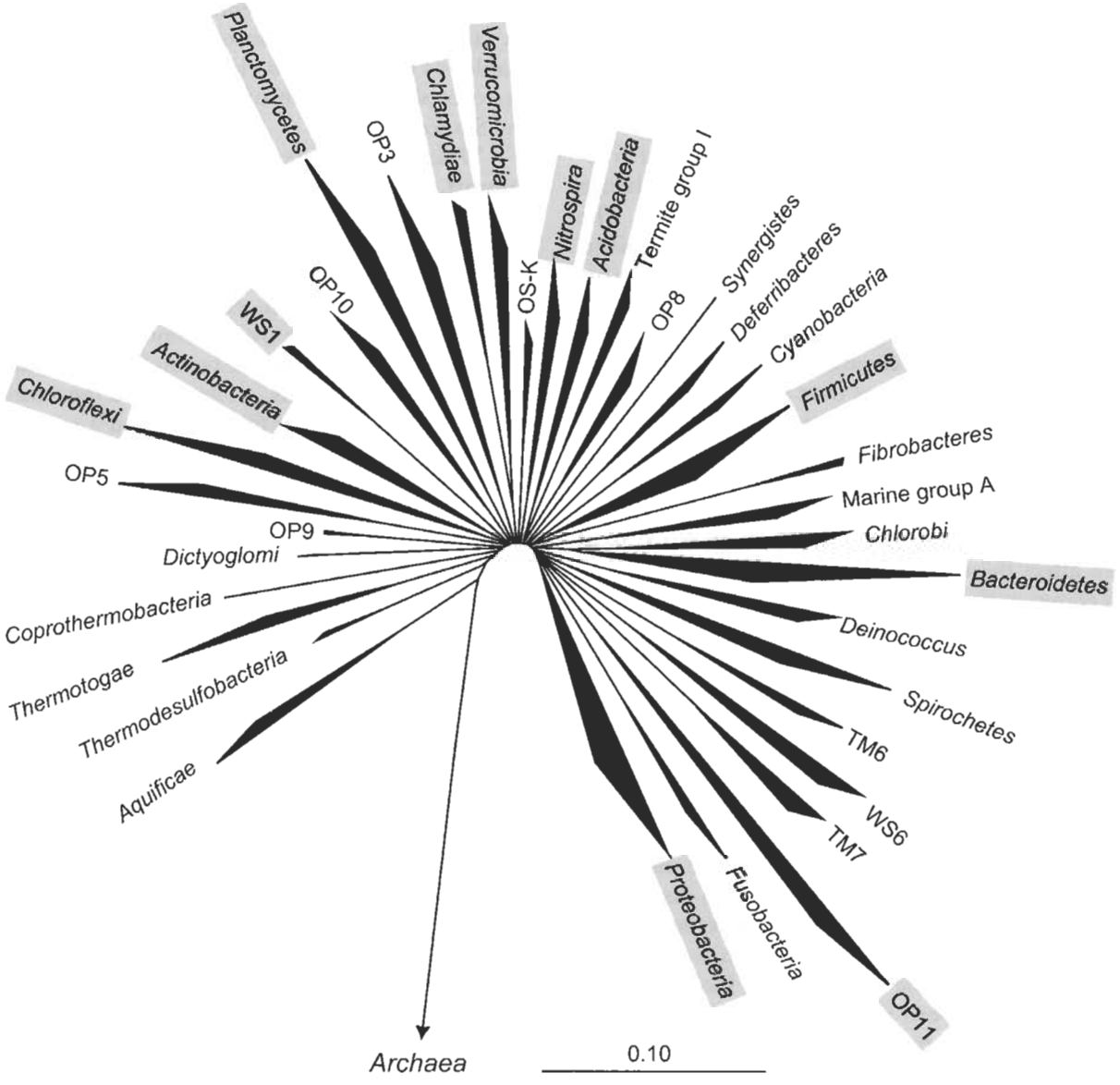

Figure 2. Phylogenetic diversity of microbial life. Phyla that have been detected in soil are indicated by shading.

approximately 2,400 sequences from soil contain few sequences that are identical to each other, span 13 phyla of *Bacteria*, contain clades of *Archaea* that, until recently, were unknown in terrestrial environments (Bintrim et al., 1997), and contain groups that diverge deeply from any cultured isolates (Borneman et al., 1996; Liles et al., 2003) (Fig. 2). The experimental evidence makes it abundantly clear that there are many thousands of species of prokaryotes yet to be discovered in soil. It is likely that these microorganisms mediate soil processes that have not been described, form novel associations with plant roots and soil animals, and produce chemicals that could be of benefit in medicine and agriculture.

Although molecular methods of analysis have improved the quality of our image of the soil microbial community by unmasking the as-yet-uncultured members, the inability to decipher further the secrets of that community has both tantalized and frustrated soil microbiologists. The 16S rRNA gene analysis provides little functional information about the membership of uncultured communities, as phylogeny rarely predicts function in microbial taxa. Therefore, the contemporary quest for a fuller understanding of the entire biology of the soil hinges on approaches that can link phylogeny with function by integrating information about both cultured and uncultured microorganisms, as well as species abundance, nutrient cycling, signal molecules, antibiotics, and other small diffusible compounds.

Excitement about the potential for understanding the undiscovered forms of life indicated by 16S

rRNA gene sequence analysis, coupled with vexation at the lack of available technology to delve into the biology of the newly described taxa, propelled the development of the approach of metagenomics (also known as environmental or community genomics), which enables genomic analysis of uncultured organisms. The idea of directly cloning DNA from the environment was suggested by Pace et al. (1985) and Pace (1986) and was developed independently by microbiologists studying soil and seawater. It was first applied successfully to analysis of oceanic microbial communities (Schmidt et al., 1991; Stein et al., 1996) and subsequently to soil communities (Handelsman et al., 1998; Rondon et al., 2000).

METAGENOMICS AS AN EXPERIMENTAL STRATEGY

Application of genomics to uncultured communities in soil involves extracting DNA directly from the soil, cloning it into a suitable vector, and transforming it into a culturable host cell (Fig. 3). Some libraries of metagenomic DNA have been constructed in fosmids and bacterial artificial chromosomes to maintain large fragments of DNA, which are necessary to capture complete biosynthetic pathways or to study genome organization. Alternatively, some libraries have been constructed in expression vectors that maintain small fragments linked to a promoter known to express genes in the host organism, facilitating functional analysis by increasing the chance of obtaining expression of foreign genes. It is the functional expression studies that facilitate the link between phylogeny and function, lending unique power to metagenomics analysis.

FUNCTIONAL DIVERSITY IN SOIL

Little is known about the relationship between phylogenetic diversity and functional diversity. A premise of much of the metagenomic work conducted on soil microbial communities is that the diversity of species will be reflected in the diversity of physiology and biochemistry. The estimates of species diversity suggest that analysis of soil microbial communities will require libraries containing large amounts of DNA to access the diverse genomes. If the estimates of soil microbial diversity are correct, and if the genomes of the unknown species are similar to the sequenced genomes, an average of 5 Mb, then 1 million clones, each containing 50 kb of DNA would be required to provide single-fold coverage of the genomes of the estimated 10,000 different species in a gram of soil. To date, none of the libraries constructed approach this size, so substantial work remains to be done.

Access to minor species in the community will provide a challenge in the advanced stages of metagenomic analysis. If complete coverage of the soil "metagenome" is to be approached, then methods will need to be developed both to avoid repeatedly cloning DNA from the abundant organisms and to gain access to DNA from rare components of the community. There has been little redundancy detected in the libraries constructed thus far, suggesting that a complete census of the soil community is a distant prospect with current methods.

The relative abundance of culturable and as-yet-uncultured organisms has not been well described. Therefore, a future challenge is to determine the contribution of DNA from culturable and uncultured organisms to metagenomic libraries and elucidate the genetic and functional relationships among the organisms that can be dissected in the laboratory and those that cannot.

LINKING PHYLOGENY AND FUNCTION

Extracting the full value of metagenomic libraries requires systematically applying functional, sequence, and phylogenetic analyses. By identifying a phylogenetic marker on a clone that expresses a functional gene product or contains sequences with significant similarity to genes of known function, the linkage of phylogeny and the function of uncultured microorganisms can be accomplished (Stein et al., 1996). A striking example of this type of discovery is the identification of a clone derived from seawater that contained a 16S rRNA gene whose sequence indicated that it was of bacterial origin (Beja et al., 2000a, 2000b). Sequence analysis of the regions flanking the 16S rRNA gene revealed a new photorhodopsin gene, which was then expressed in *Escherichia coli* and shown to have light-harvesting capability. All microbial photorhodopsins discovered prior to this study were found in *Archaea,* and *Bacteria* were assumed not to contain this light-harvesting strategy. The bacteriorhodopsin discovery in a bacterial genome provides a new paradigm for linking phylogeny and function in uncultured microorganisms.

Although much can be inferred from sequence analysis of the genomes of uncultured microorganisms, sequence-driven study is limited to the functional knowledge about the sequences currently in the genetic databases. Functional analysis has the potential to expand the range of known functions and elucidate functions of genes with no known homologs.

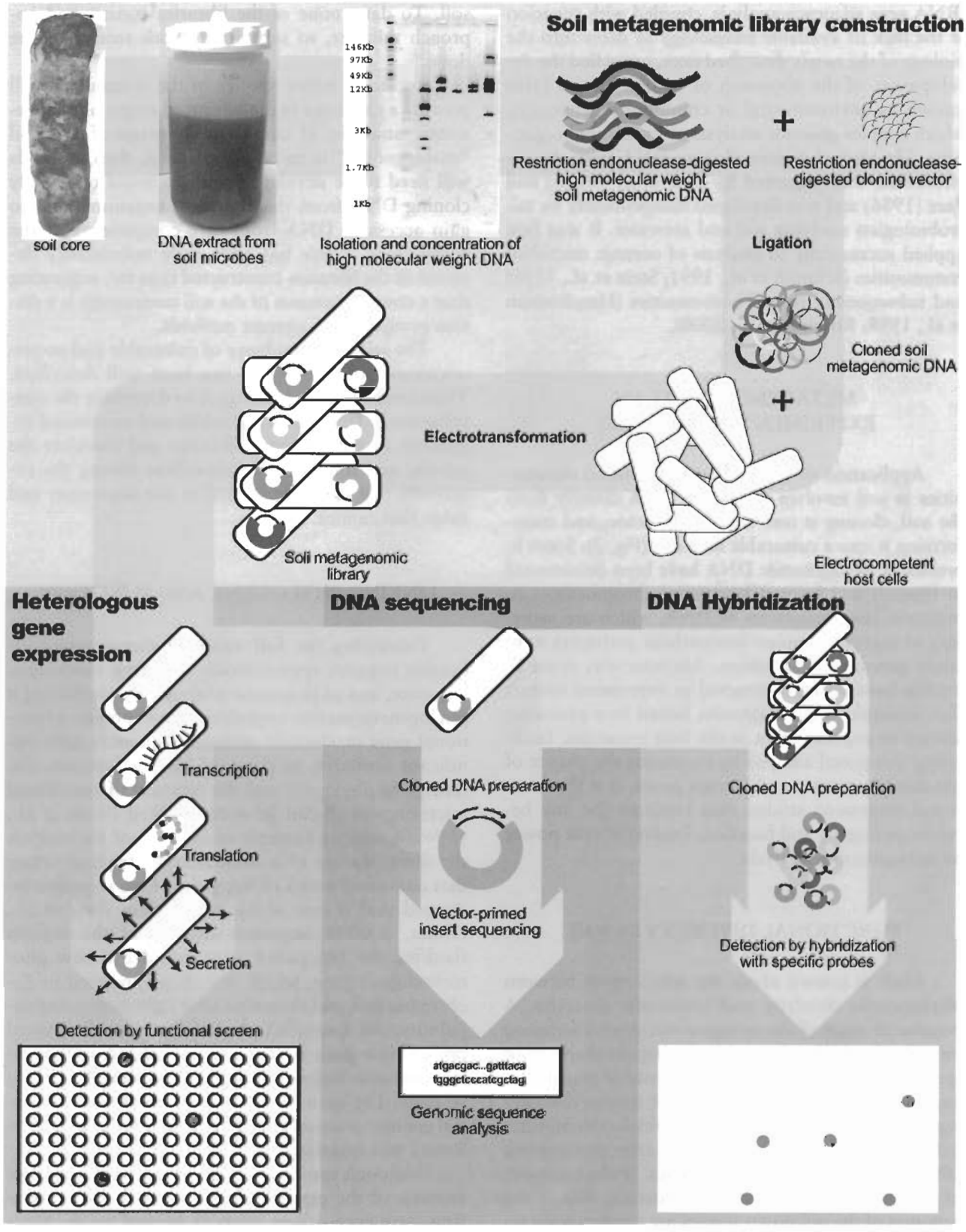

Figure 3. Strategy for metagenomic analysis.

Metagenomic libraries from soil have provided a rich source of new functions, which have enriched our knowledge of microbial biochemistry and contributed significantly to the genetic databases.

BIOLOGICAL INSIGHTS INTO THE SOIL FROM METAGENOMICS

Metagenomic analysis has added to our knowledge of proteins and small molecules encoded in the genomes of soil microorganisms. Soil metagenomic libraries produce diverse antibiotics. Biosynthetic pathways encoded by metagenomic DNA and expressed in *E. coli* direct the production of violacein, terragine, turbomycin, indirubin, fatty dienic alcohols, and the acyl tyrosines (Fig. 4). The synthesis of some of these antibiotics is encoded by a single gene whereas some require a complex operon for expression. Violacein, for example, requires coordinated expression of seven genes for its synthesis. The cloning of this pathway demonstrates the feasibility of heterologous expression of pathways from anonymous microorganisms from soil. Some of the antibiotics discovered are novel, such as the acyl tyrosines, terragine, and turbomycin, further illustrating the utility of metagenomics for biosprospecting. New insights into the evolution of antibiotic biosynthetic pathways may emerge as even the genes for synthesis of known antibiotics cloned in the metagenomic libraries diverged deeply from related genes from cultured organisms, suggesting that there is substantial genetic diversity to be discovered in known genes in uncultured organisms. In addition to new antibiotics, new enzymes, an antiporter, and antibiotic resistance determinants have been discovered (Table 1). A striking feature of the studies leading to these discoveries is the low frequency of clones expressing any given phenotype in metagenomic libraries. For example, one library consisting of 25,000 clones representing 1 Gb of DNA contained 3 that expressed detectable antimicrobial activity (Gillespie et al., 2002). Two screens for lipolytic activity detected 1 positive out of 730,000 clones and 3 positives out of 286,000 clones, respectively (Henne et al., 2000). The most feasible metagenomic studies thus involve selections or enrichments, rather than screens, for the phenotype of interest, negating the need to test thousands of clones individually. Robotics and high-throughput screening will accelerate the pace of discovery of genes for which there is no selection.

The emerging paradigm of metagenomic analysis is to screen libraries for 16S rRNA genes or other phylogenetic markers and sequence the flanking DNA to identify functions that can be recognized by sequence similarity, or to identify functionally active clones by selection or screen and then sequence the DNA flanking the function genes to identify phylogenetic markers (Sandler et al., 1999; Schleper et al., 1997). This approach will yield data linking phylogeny and function that will be used to design culturing strategies that are driven by the physiology of the organisms to be cultured.

CHALLENGES AND LIMITATIONS IN METAGENOMIC ANALYSIS

Although metagenomic libraries have yielded interesting information about biological systems, there are many barriers to extracting the full range of information that is locked in the genomes of the uncultured microbial world. First, some *Bacteria* and *Archaea* are recalcitrant to lysis and their DNA is not represented in the libraries. Some genes or their products will be unstable or toxic to the host cell and therefore will not be maintained in the libraries. Many genes will be derived from organisms that are phylogenetically distant from *E. coli* or another culturable species that is used as host to the library, and thus functional analysis will not detect these genes. The most interesting new genes are those that do not have high sequence similarity to those in the databases, and therefore expression and functional analysis are essential for discerning the nature of the gene product.

One goal of metagenomic analysis is to reconstruct genomes of uncultured organisms. This will be particularly challenging in soil communities because they are so complex. In addition to the large number of species, currently estimated at 10^4 or more per g of soil, there is likely microheterogeneity among genomes within a species. Recent work on *Cenarchaeum symbiosum,* a species that forms an intimate associate with a sponge, showed substantial variation among genomes within a local population of the bacterium (Schleper et al., 1998). Therefore, layers of complexity may make it impossible to reconstruct a genome, as is possible in simpler communities, but association of genes, their functions, and genomic organization with particular taxa will be a fruitful area of endeavor in soil metagenomics. Once the task of reassembling 5 to 10 genomes of uncultured organisms simultaneously has been accomplished in a simpler habitat, it will be easier to predict whether overcoming the technical and computational challenges of reconstructing genomes from soil is possible with today's expertise.

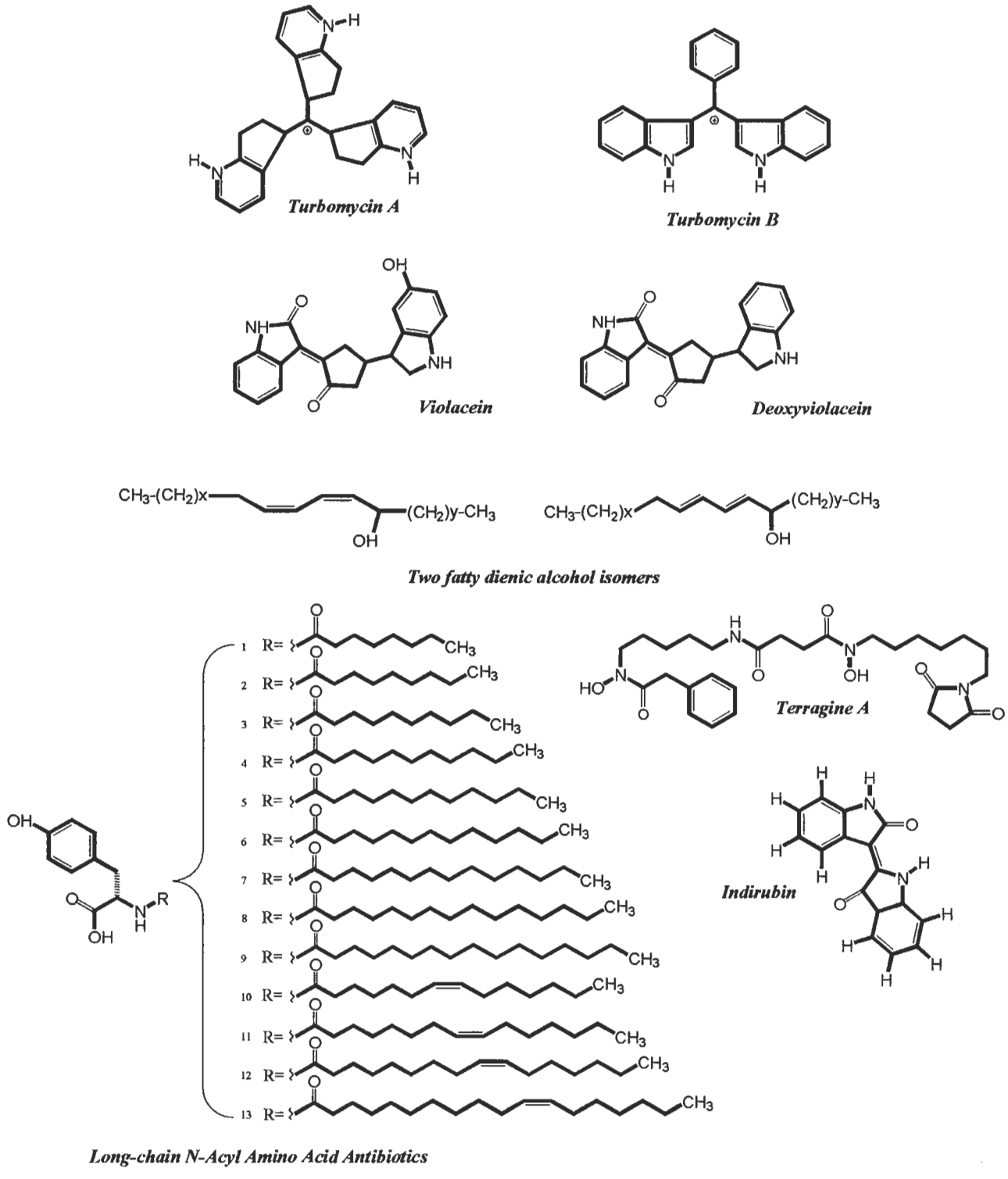

Figure 4. Small molecules discovered in metagenomic libraries.

Table 1. Genes and gene products discovered in metagenomic libraries constructed with DNA extracted directly from soil

Gene and/or function	Relevant characteristic	Reference(s)
N-Acyl amino acid compounds	Novel antibiotics	Brady and Clardy (2000)
Violacein	Antibiotic known in cultured organisms	Brady et al. (2001)
Turbomycin A and B	Tri-aryl cation antibiotics	Gillespie et al. (2002)
Indirubin and related compounds	Known compound; some compounds have antimicrobial activity	MacNeil et al. (2001)
Terragine A and related small molecules	Novel antibiotics	Wang et al. (2000)
Kanamycin resistance gene	New acetyltransferase homolog	Courtois et al. (2003)
Fatty dienic alcohol	Novel compounds	Courtois et al. (2003)
Na^+/H^+ antiporter	New homolog	Majernik et al. (2001)
4-Hydroxybutyrate utilization enzymes	New homologs	Henne et al. (1999)
Amylases	Diverse homologs	Rondon et al. (2000)
Lipolytic activity	New homologs	Henne et al. (2000), Rondon et al. (2000)
Hemolytic activity	Diverse homologs	Rondon et al. (2000)
DNase	New homolog	Rondon et al. (2000)
16S rRNA	Diverse sequences	Rondon et al. (2000)

FUTURE DIRECTIONS IN METAGENOMICS

Metagenomic analysis depends on technical innovations that will provide improved access to DNA, heterologous gene expression, and DNA sequence analysis. To access DNA from a wider range of organisms, we need methods that facilitate aggressive bacterial cell lysis while protecting the integrity of the DNA. We need strategies to increase representation of rare members of the community by removing the DNA from the abundant organisms. This might involve antibodies that remove certain bacterial cell types prior to cell lysis or methods for removing the genomes after lysis, although the latter will be challenging because of the conservation among all bacterial genomes.

Heterologous gene expression will likely be one of the areas of greatest improvement in metagenomic analysis, largely because of the sophisticated understanding of the elements in gene regulation that has emerged from the past few decades of study. Gene expression can be enhanced by transferring clones to a variety of organisms that represent a range of G+C content, preferred codon usage, and protein secretion strategies. This requires shuttle vectors that replicate in *E. coli* and a host cell with gene expression machinery that differs substantially from that of *E. coli,* such as *Bacillus* (a low G+C gram positive) or *Streptomyces* (a high G+C gram positive).

Alternatively, heterologous gene expression could be enhanced by using a host cell modified to express a broader range of genes. To enhance transcription, genes encoding sigma factors from organisms closely related to the uncultured clades of interest can be cloned into the host cell. To enhance translation, a broader range of tRNAs can be supplied by introducing a range of genes encoding tRNAs that may be rare in the host cell but are typical of other cells. *E coli* carrying genes for tRNAs typically required by organisms with G+C- or A+T-rich genomes are currently available and ready to be tested for their efficacy in enhancing functional expression in metagenomic libraries (Stratagene, BL21-Condon Plus Competent Cells, 2002, http://www.stratagene.com/).

Heterologous expression of many of the genes cloned in metagenomic libraries is likely to be poor even if measures are taken to enhance gene expression. Therefore, in order to capture most of the expressed genes in functional screens, highly sensitive screens are essential. Detection of gene products in assays that amplify an initial signal will increase the rate of discovery. Assays in which the detection system is inside the cell carrying the cloned DNA will enhance detection as local concentrations of gene products may be higher inside the cell than following secretion into the surrounding medium. Highly sensitive detectors and small-volume assays that maintain high local concentration of products from a clone will likely increase the rate of finding active clones.

CONCLUSIONS

Metagenomics offers a new look by accessing the as-yet-unculturable microorganisms that represent the majority of life in soil. Metagenomic analysis of soil is challenging because soil is such a complex environment, containing diverse organisms in a dynamic matrix. The biological, chemical, and physical properties of soil all contribute to the technical difficulties of the analysis, and thus numerous obstacles to cloning and analyzing the metagenome of soil remain to be overcome. The potential is vast. Discovery of metabolic networks and small molecules from the cultured microorganisms from soil has been unparalleled in any other environment, and therefore it is

likely that the potential for discovery from the uncultured community is similarly enormous. Because most of the uncultured life forms in soil appear to be new species and many represent new genera, there is much to learn about the fundamental functioning of soil microbial communities, and these communities have already yielded new enzymes and antibiotics. Technical advances in DNA recovery, gene expression, and functional analysis will enhance the rate of discovery and make possible productive prospecting of soil for the medicinal, agricultural, and industrial chemicals.

Acknowledgments. I thank Zakee Sabree and Tina Matta for preparing the figures and Ed DeLong for helpful discussions. The preparation of the manuscript was supported by the Howard Hughes Medical Institute, National Science Foundation grant MCB-0132085, and the University of Wisconsin-Madison.

REFERENCES

Beja, O., L. Aravind, E. V. Koonin, M. Suzuki, A. Hadd, L. P. Nguyen, S. B. Jovanovich, C. M. Gates, R. A. Feldman, J. L. Spudich, E. N. Spudich, and E. F. DeLong. 2000a. Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science* 289:1902–1906.

Beja, O., M. T. Suzuki, E. V. Koonin, L. Aravind, A. Hadd, L. P. Nguyen, R. Villacorta, M. Amjadi, C. Garrigues, S. B. Jovanovich, R. A. Feldman, and E. F. DeLong. 2000b. Construction and analysis of bacterial artificial chromosome libraries from a marine microbial assemblage. *Environ. Microbiol.* 2:516–529.

Bintrim, S. B., T. J. Donohue, J. Handelsman, G. P. Roberts, and R. M. Goodman. 1997. Molecular phylogeny of Archaea from soil. *Proc. Natl. Acad. Sci. USA* 94:277–282.

Borneman, J., P. W. Skroch, K. M. O'Sullivan, J. A. Palus, N. G. Rumjanek, J. L. Jansen, J. Nienhuis, and E. W. Triplett. 1996. Molecular microbial diversity of an agricultural soil in Wisconsin. *Appl. Environ. Microbiol.* 62:1935–1943.

Brady, N. C., and R. R. Weil. 2002. *The Nature and Properties of Soils*, 13th ed. Prentice-Hall, Upper Saddle River, N.J.

Brady, S. F., and J. Clardy. 2000. Long-chain *N*-acyl amino acid antibiotics isolated from heterologously expressed environmental DNA. *J. Am. Chem. Soc.* 122:12903–12904.

Brady, S. F., C. J. Chao, J. Handelsman, and J. Clardy. 2001. Cloning and heterologous expression of a natural product biosynthetic gene cluster from eDNA. *Org. Lett.* 3:1981–1984.

Courtois, S., C. M. Cappellano, M. Ball, E.-X. Francou, P. Normand, G. Helynck, A. Martinez, S. J. Kolvek, J. Hopke, M. S. Osburne, P. R. August, R. Nalin, M. Guérineau, P. Jeannin, P. Simonet, and J. L. Pernodet. 2003 Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products. *Appl. Environ. Microbiol.* 69:49–55.

Edwards, K. J., T. M. Gihring, and J. F. Banfield. 1999 Interdependence of microbial populations and environmental conditions at an extreme acid mine drainage environment. *Appl. Environ. Microbiol.* 65:3627–3632.

Gillespie, D. E., S. F. Brady, A. D. Bettermann, N. P. Cianciotto, M. R. Liles, M. R. Rondon, J. Clardy, R. M. Goodman, and J. Handelsman. 2002. Isolation of antibiotics turbomycin A and B from a metagenomic library of soil microbial DNA. *Appl. Environ. Microbiol.* 68:4301–4306.

Handelsman, J., M. R. Rondon, S. Brady, J. Clardy, and R. M. Goodman. 1998. Molecular biology provides access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem. Biol.* 5:R245–R249.

Henne, A., R. Daniel, R. A. Schmitz, and G. Gottschalk. 1999. Construction of environmental DNA libraries in *Escherichia coli* and screening for the presence of genes conferring utilization of 4-hydroxybutyrate. *Appl. Environ. Microbiol.* 65: 3901–3907.

Henne, A., R. A. Schmitz, M. Bomeke, G. Gottschalk, and R. Daniel. 2000. Screening of environmental DNA libraries for the presence of genes conferring lipolytic activity on *Escherichia coli*. *Appl. Environ. Microbiol.* 66:3113–3116.

Hugenholtz, P., B. M. Goebel, and N. R. Pace. 1998 Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *J. Bacteriol.* 180:4765–4774.

Liles, M. R., B. F. Manske, S. B. Bintrim, J. Handelsman, and R. M. Goodman. 2003. A census of rRNA genes and linked genomic sequences within a soil metagenomic library. *Appl. Environ. Microbiol.* 69:2684–2691.

MacNeil, I. A., T. C. Minor, P. R. August, T. H. Grossman, K. A. Loiacono, et al. B. A. Lynch, T. Philips, S. Narula, R. Sundaramoorthi, A. Tyler, T. Aldredge, H. Long, M. Gilman, and M. S. Osburne. 2001. Expression and isolation of antimicrobial small molecules from soil DNA libraries. *J. Mol. Microbiol. Biotechnol.* 3:301–308.

Majernik, A., G. Gottschalk, and R. Daniel. 2001. Screening of environmental DNA libraries for the presence of genes conferring $Na^+(Li^+)/H^+$ antiporter activity on *Escherichia coli:* characterization of the recovered genes and the corresponding gene products. *J. Bacteriol.* 183:6645–6653.

Pace, N. R., D. A. Stahl, D. J. Lane, and G. J. Olsen. 1985. Analyzing natural microbial populations by rRNA sequences. *ASM News* 51:4–12.

Pace, N. R. 1996. New perspective on the natural microbial world: molecular microbial ecology. *ASM News* 62:463–470.

Paul, E. A., and F. E. Clark. 1989. *Soil Microbiology and Biochemistry*. Academic Press, New York, N.Y.

Paul, E. A., and F. E. Clark. 1996. *Soil Microbiology and Biochemistry*, 2nd ed. Academic Press, New York, N.Y.

Rondon, M. R., P. R. August, A. D. Bettermann, S. F. Brady, T. H. Grossman, M. R. Liles, K. A. Loiacono, B. A. Lynch, I. A. MacNeil, M. S. Osburne, J. Clardy, J. Handelsman, and R. M. Goodman. 2000. Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl. Environ. Microbiol.* 66:2541–2547.

Sandler, S. J., P. Hugenholtz, C. Schleper, E. F. DeLong, N. R. Pace, and A. J. Clark. 1999. Diversity of *radA* genes from cultured and uncultured *Archaea:* comparative analysis of putative RadA proteins and their use as a phylogenetic marker. *J. Bacteriol.* 181:907–915.

Schleper, C., G. Pühler, B. Kühlmorgen, and W. Zillig. 1995 Life at extremely low pH. *Nature* 375:741–742.

Schleper, C., R. V. Swanson, E. J. Mathur, and E. F. DeLong. 1997. Characterization of a DNA polymerase from the uncultivated psychrophilic archaeon *Cenarchaeum symbiosum*. *J. Bacteriol.* 179:7803–7811.

Schleper, C., E. F. DeLong, C. M. Preston, R. A. Feldman, K. Y. Wu, and R. V. Swanson. 1998. Genomic analysis reveals chromosomal variation in natural populations of the uncultured psychrophilic archaeon *Cenarchaeum symbiosum*. *J. Bacteriol.* 180: 5003–5009.

Schmidt, T. M., E. F. DeLong, and N. R. Pace. 1991. Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing. *J. Bacteriol.* 173:4371–4378.

Stein, J. L., T. L. Marsh, K. Y. Wu, H. Shizuya, and E. F. DeLong. 1996. Characterization of uncultivated prokaryotes: isolation and analysis of a 40-kilobase-pair genome fragment from a planktonic marine archaeon. *J. Bacteriol.* 178:591–599.

Torsvik, V., J. Goksøyr, and F. L. Daae. 1990. High diversity in DNA of soil bacteria. *Appl. Environ. Microbiol.* 56:782–787.

Torsvik, V., R. Sorheim, and J. Goksøyr. 1996. Total bacterial diversity in soil and sediment communities—a review. *J. Ind. Microbiol.* 17:170–178.

Whitman, W. B., D. C. Coleman, and W. J. Wiebe. 1998. Prokaryotes: the unseen majority. *Proc. Natl. Acad. Sci. USA* 95:6578–6583.

Waksman, S. A., and R. E. Curtis. 1916. The actinomyces of the soil. *Soil Sci.* 1:99–134.

Waksman, S. A., and J. W. Foster. 1937. Associative and antagonistic effects of microorganisms: II. Antagonistic effects of microorganisms grown on artificial substrates. *Soil Sci.* 43: 69–76.

Waksman, S. A., and R. L. Starkey. 1931. *The Soil and the Microbe.* John Wiley, New York, N.Y., p. viii.

Wang, G.-Y.-S., E. Graziani, B. Waters, W. Pan, X. Li, J. McDermott, G. Meurer, G. Saxona, J. Andersen, and J. Davies. 2000. Novel natural products from soil DNA libraries in a Streptomycete host. *Org. Lett.* 2:2401–2404.