

Molecular geomicrobiology: genes and geochemical cycling

Jennifer Macalady¹, Jillian F. Banfield*

*Department of Earth and Planetary Science and Department of Environmental Sciences, Policy, and Management,
University of California, Berkeley, CA 94720, USA*

Received 11 July 2002; received in revised form 30 August 2002; accepted 3 October 2002

Abstract

Core geosciences concepts are being fundamentally revised as the result of breakthroughs in geomicrobiology. Revolutionary discoveries have resulted from increased effort devoted to study of microorganisms in the context of their environments. Much recent progress has been made possible by genetic data, particularly those that allow the description of microbial populations in situ. New gene and genome sequences are elucidating previously unexpected or unexplained interactions between microorganisms and Earth materials, with implications for key geological phenomena such as the formation of ore deposits and the regulation of global climate and surface oxidation state. Genetic data have also led to extensive revision of our understanding of the pace and mechanisms by which evolution occurs. Yet, the field of molecular geomicrobiology remains in its infancy. In the foreseeable future, merging of modern biogeochemistry with molecularly resolved ecological studies will inspire the development of integrated models for the processes that shape the Earth.

© 2002 Elsevier Science B.V. All rights reserved.

Keywords: geomicrobiology; geochemistry; evolution; genes; genomes; ecology

1. Introduction

What does ‘molecular geomicrobiology’ mean? Although this could be interpreted as ‘molecular biology applied to geologic systems’ we use the term to describe the search for molecular-level understanding of coupled biological and geochemical processes. Geoscientists have moved from describing complex mineral surface reactions

using empirical rate laws to analyzing molecular interactions at a fundamental level to determine kinetic parameters with predictive value [1]. Similarly, it is necessary for geoscientists seeking to understand biogeochemical processes to go beyond modeling microorganisms as black boxes. Access to complete microbial genome sequences (Fig. 1) and methods to monitor gene activity within organisms, as well as improvements in techniques for measurement of elemental composition, isotopic ratios, and mineral surface chemistry, make this possible. Genomes simultaneously convey information about metabolic capabilities and gene regulation and enable us to probe life’s origins and to explore the molecular mechanisms by which evolution occurs.

* Corresponding author.

E-mail address: jill@seismo.berkeley.edu (J.F. Banfield).

¹ Present address: Department of Geology, Carleton College, Northfield, MN 55057, USA

Fig. 1. A simple representation of a microbial genome. Bars around the circle represent genes; the gene type is indicated by color. A small section of the genome is enlarged to show a sequence of genes that code for proteins (in this case enzymes involved in energy generation), as well as genes of unknown function. Genes are constructed from nucleotides (A,T,C,G), groups of which can code for specific amino acids that are assembled into the proteins. The genome shown contains 1950000 nucleotides (1.95 megabases, mb).

→

Rapid progress in genomics is due to simultaneous innovations in DNA sequencing capabilities, technologies to monitor gene activities, and statistical and mathematical approaches for analyzing genetic data. To date, the genomes of about 100 organisms have been sequenced. A project is underway to sequence the genomes of thousands of prokaryotic organisms maintained in microbial culture collections within the next five years.

From a geoscience perspective, the new genome sequences will represent a gold mine. The proteins encoded by the genes include all of the biochemical machinery that transforms molecules important in natural environments. These enzymes have mediated, on a molecular level, global elemental cycles for the past 3–4 Gyr. Gene sequence analysis has the potential to reveal the order in which biological processes evolved, and whether proteins with a given function have evolved only once or multiple times in Earth history. Genes that code for proteins involved in lipid biosynthesis and transformations involving elements such as C, N, S, Fe, and other metals have direct relevance for the interpretation of lipid biomarkers and isotopic signatures in both modern environments and in the geologic record. Given the predominance of microorganisms over most (if not all) of Earth history, and in view of the fact that they are estimated to comprise the majority of the Earth's biomass [2], genomics and related technologies will teach us a great deal about biology, geology, and the coupling between these.

2. Fossils, genes, genomes, and evolution

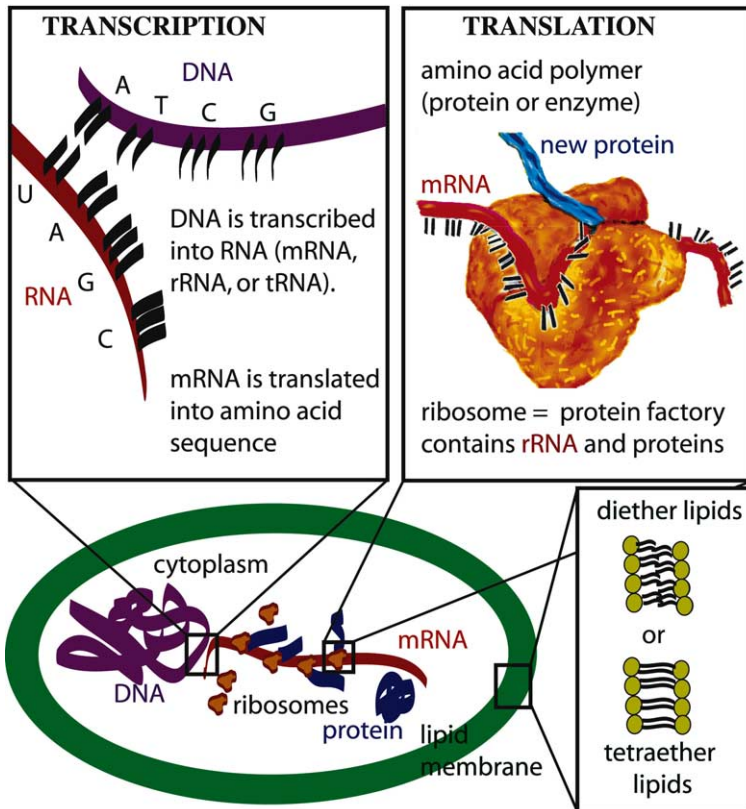
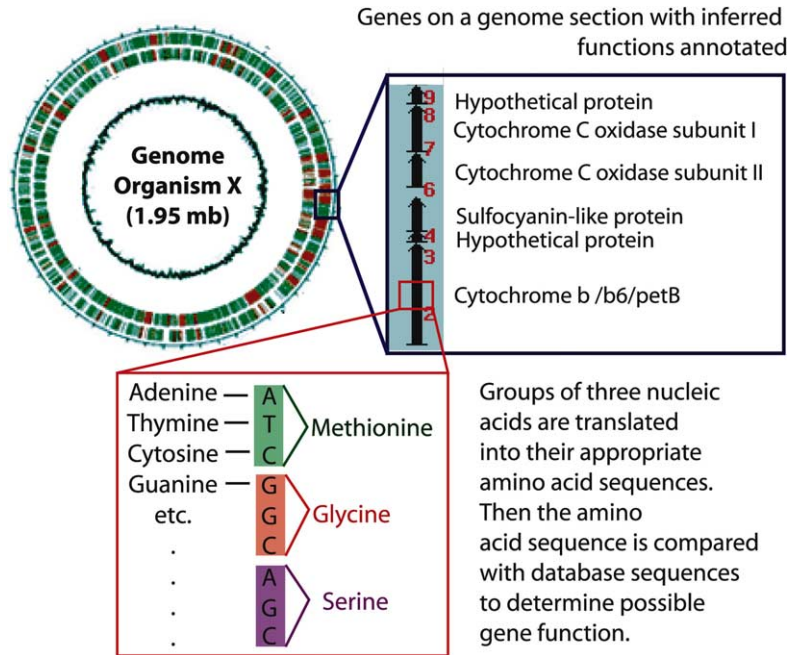
2.1. Putting microbes into the evolution picture

Geology has been connected to biology from the very beginning of these sciences. The fossil record was central to the development of the geologic timescale, and stratigraphic correlations based on fossil inventories have defined the scope and timing of major periods of biological innovation and extinction. Microorganisms are highly under-represented in this picture due to their poor fossil record. This may have led to considerable bias in how we think about natural selection and biological evolution. Evolution of macroorganisms by random mutation and natural selection is perturbed by major events such as meteorite impacts, climate change, volcanic eruptions, and alteration of ocean chemistry, and is closely followed by new species radiations. It is unclear, however, whether microbial evolution exhibited the same pattern of extinction and rapid speciation documented in the rock record for macroscopic eukaryotes.

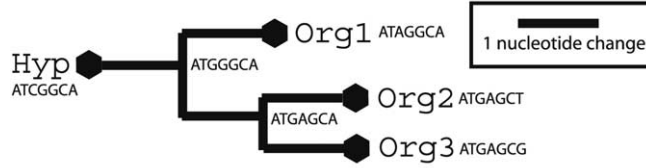
Genes have emerged as a key to unraveling the history of life. Analysis of genes in modern microorganisms partially circumvents their lack of a fossil record. Methods have been developed to analyze differences between DNA sequences (and their translated amino acid sequences; Fig. 2) so as to identify genes that were derived from a common ancestral sequence (homologs). For homologs, it is possible to use the pattern of differences to reconstruct the order of events

→

Fig. 2. Diagram showing key cell components and illustrating transcription (copying) and translation (conversion of the nucleotide sequence into proteins) by ribosomes. Two possible archaeal membrane lipid types are shown. The analogous bacterial lipids are ester linked and consist of fatty acid rather than isoprenoid chains.



Organism	DNA sequence	CHANGES PER SITE			
		Hyp	Org1	Org2	Org3
Hyp	...ATCGGCA	0/7			
Org1	...ATAGGCA	2/7	0/7		
Org2	...ATGAGCT	3/7	3/7	0/7	
Org3	...ATGAGCG	3/7	3/7	2/7*	0/7



* Org2 and Org3 are both one change different from their immediate ancestor, yet because the mutation occurred at the same site, Org2 and Org3 differ by only one base.

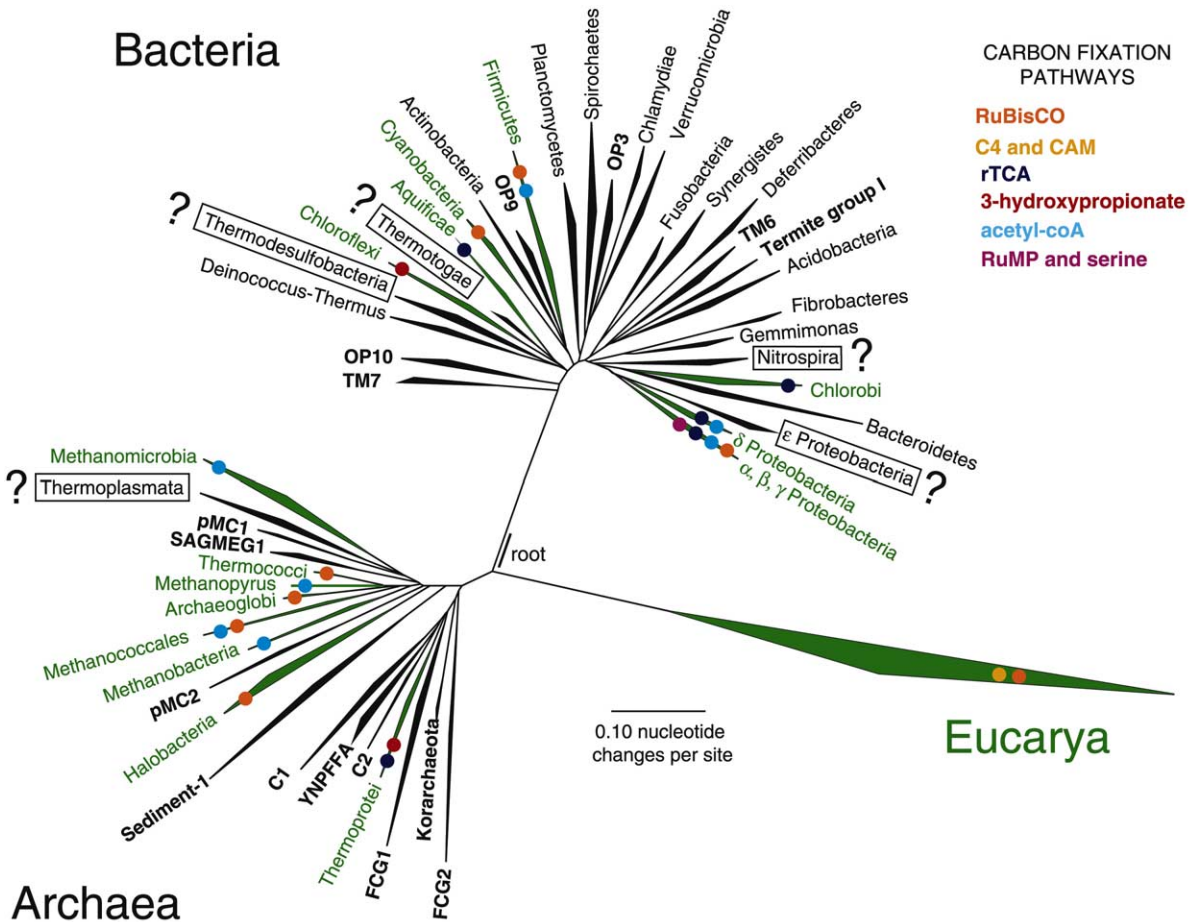


Fig. 3. An example of an alignment of DNA sequences from three organisms (Org1, Org2, and Org3) and the associated phylogenetic tree. Phylogenetic trees are interpreted based on two features. The horizontal branch length between two nodes of the tree reflects the number of nucleotide positions that differ between the two sequences, and is a measure of the ‘evolutionary distance’ between two genes. The branching order reflects the pattern of common ancestry and the sequence of evolutionary divergences that produced the genes at the fringe of the tree. For educational purposes we have included the sequences for the hypothetical ancestor (hyp) and for branch points in this tree.

←

that led to appearance of the modern genes (Fig. 3). This approach is referred to as molecular phylogeny. Molecular phylogeny is routinely practiced using gene sequences coding for ribosomal RNA, a functional RNA that is not translated into protein (Fig. 2). Because ribosomal RNA is a slowly evolving molecule that has the same function (Fig. 2) in all cells, ribosomal genes have been extensively used to construct representations of the tree of life (Fig. 4). As genomic data have become more readily available, it is now possible to view evolution through the phylogeny of genes coding for a wider diversity of functions.

2.2. The tree of life and where it all began

Some of the most intriguing discussion of evolution involves the concept that the tree has a single origin, i.e., a root. This idea has fueled speculation about the characteristics of this earliest organism, referred to as the last common ancestor (LCA). The early hope was that the habitat (temperature, pressure, pH, etc.) and metabolic attributes of lineages that branch near the base of the ribosomal gene tree would point toward the characteristics of the LCA, and even provide clues about the origin of life. For example, in ribosomal gene trees the organisms nearest the ‘root’ are

thermophiles, leading to the suggestion that life evolved at high temperature. This hypothesis is made more convincing by the observation that many of these thermophilic microorganisms are lithoautotrophs (Table 1). Phylogenetic trees based on large groups of genes [3] also support the idea that the first cells were heat-loving. Thermophiles near the root of the tree could be explained either by a thermophilic origin of life, or by the survival of thermophiles through a meteorite impact bottleneck. However, the simple picture constructed from early ribosomal gene datasets has become cloudier with the increase in number of organisms represented. Phylogenetic trees based on ribosomal genes now resemble shrubs with a poorly resolved branching order (Fig. 4), bringing into question correlations between physiological characteristics of the LCA and modern representatives of deep-branching lineages. Some methods for analyzing ribosomal genes do not support the placement of thermophilic lineages nearest the root [4]. Furthermore, if genes other than those involved in ribosomes are used, the tree topology sometimes differs from that of the ribosomal tree [5,6]. These discrepancies have highlighted the phenomenon of lateral gene transfer (LGT, the acquisition of genes from unrelated organisms), which may have been extremely prevalent early in life history.

←

Fig. 4. Phylogenetic tree based on 16S ribosomal RNA genes, showing the root and the separation of life into three domains: bacteria, archaea (prokaryotes), and eukarya (modified from [85]). Lineages with known autotrophs (Table 1) are shown in green, and the distributions of eight CO₂ fixation pathways are indicated with colored dots. The resulting phylogenetic pattern is surprisingly complex and does not follow the evolutionary pattern predicted by ribosomal genes. RuBisCO is the key enzyme in organisms using the Calvin–Bassham–Benson cycle, including cyanobacteria and other photosynthetic bacteria, some archaea, and photosynthetic eukaryotes. The serine and ribulose monophosphate (RuMP) pathways are unique to methylotrophic bacteria within the α,β,γ Proteobacteria lineage. The reverse TCA (rTCA) cycle is found in the photosynthetic bacterial genus *Chlorobium* and in certain other bacteria and archaea. The acetyl-coA (Ljungdahl–Wood) pathway is found in a variety of anaerobic, non-photosynthetic bacteria and archaea, especially methanogens. Anaerobic, photosynthetic bacteria such as *Chloroflexus* and some archaea in the group Sulfolobales (Thermoprotei lineage) use the 3-hydroxypropionate pathway. Not all species within lineages contain the indicated pathways. Lineages surrounded by boxes with question marks include autotrophs that use unknown CO₂ fixation pathways. Lineages in bold have no isolated (cultured) representatives available for study.

Table 1
The energy and carbon sources define the microorganism type

Energy	Carbon	Terminology
Light	CO ₂	photoautotroph
Light	organic carbon	photoheterotroph
inorganic chemicals	CO ₂	lithoautotroph
inorganic chemicals	organic carbon	lithoheterotroph
organic carbon	organic carbon	heterotroph

Heterotrophs could alternatively be referred to as organo-heterotrophs.

The lack of resolution of branching order of major lineages and recognition of the possible importance of LGT led to the notion that the LCA was actually a community of pre-cellular gene complexes (last common community, LCC) that ‘crystallized’ into a diverse group of organisms that represented the last common ancestors of the modern lineages [7,8].

2.3. Lateral gene transfer in genome evolution

New whole genome sequences provide abundant evidence that microorganisms cannot be viewed purely as products of tree-like vertical descent. LGT is potentially of great ecological and evolutionary importance because it permits rapid acquisition of capabilities needed to colonize new habitats or respond to new challenges. Examples that illustrate the scope and range of LGT include acquisition of up to 24% of the genes in the genome of the thermophilic bacterium *Thermotoga maritima* from thermophilic archaea [9], and transfer of genes from plants to the radiation-resistant bacterium *Deinococcus radiodurans* [10]. Most conservative estimates of the fraction of ‘foreign’ (laterally transferred) DNA in microbial genomes range from 0 to 17% [11].

Despite the suggestion that LGT may be so pervasive as to obscure an organismal lineage [12], recent analyses lend support to the basic concept of evolutionary lines of descent. Genome trees produced using up to 37 genomes are broadly consistent with ribosomal gene trees [13,14,15,16]. This conclusion is strengthened when obviously foreign (laterally transferred) genes are removed from the analysis [17]. Similarity between trees based on ribosomal genes and

those constructed based on gene content in entire genomes (rather than genome sequences) also reinforce the evolutionary conclusions derived from ribosomal gene-based analyses [14]. Support for the division of all life forms into bacterial, archaeal, and eukaryal domains [18] is particularly strong.

3. Ancient environments

3.1. LGT and key biogeochemical cycles

Phylogenetic analysis of ancient genes can provide important clues about the roles played by heredity, gene invention, and LGT in evolution of pathways critical in geochemical cycling. Selection of genes for analysis can be guided by geochemical data from the geologic record. There is isotopic evidence for microbial sulfate reduction at 3.47 Ga [19], indicating that this trait evolved early in Earth’s history. The root of the phylogenetic tree for a key enzyme in sulfate reduction, dissimilatory sulfite reductase (*dsr*), is within the bacterial domain. Although most modern sulfate-reducing microbial groups are bacterial, archaeal *Archaeoglobus* species are exceptions. *Archaeoglobus* species contain *dsr* homologs that appear to have been acquired via LGT [20]. The observation that all modern *dsr* genes have a bacterial origin suggests that sulfate reduction was not a trait present in the LCA or LCC. Together, sulfur isotope data and *dsr* phylogeny suggest that the LCA pre-dated 3.47 Ga. Alternatively, there may be another explanation for the sulfur isotope signature.

In order to reveal the evolution of important biogeochemical functions, it is advantageous to focus directly on phylogeny of genes for the function of interest rather than on ribosomal genes. For example, a group of 31 genes involved in photosynthesis was recently analyzed to show that the earliest photosynthetic organisms diverged from a bacterial lineage called the Proteobacteria [21], a finding which has bearing on the evolution of photosynthetic pigments and other machinery for harvesting light energy. Since the evolutionary model for photosynthesis genes is at

odds with the model predicted by ribosomal genes, it is reasonable to conclude that LGT played an important role in the evolution of photosynthesis and photosynthetic organisms.

3.2. Acquisition of genes in response to geochemical change

Earth's evolution is intimately linked with the evolution of new microbial traits and capabilities. For example, volcanism could increase the temperature, selecting for heat-stable proteins; faulting could modify groundwater flow paths, selecting for modified electron transport chain components in response to alteration of the redox potential. The rise in oxygen in the Proterozoic created the need for mechanisms for iron acquisition, to deal with free radical toxicity, and to transform oxidized nitrogen and sulfur species. In modern environments, microbial populations have been shown to acquire genes that allow biodegradation of toxic, human-made organic compounds that they could never have previously encountered [22]. In fact, changes in microbial populations may be used as indicators for environmental perturbation. These cases illustrate the strong interplay of microbial activity, geochemical change, and evolution, and raise questions about how new biochemical capabilities arise and spread in response to environmental change.

Many genes providing resistances to toxic compounds such as metals and metalloids have homologs in both the archaeal and bacterial domains, suggesting that their ancestral sequence pre-dated the divergence of these domains. Because of their antiquity, these genes are good targets for studies of how pathways evolve in response to geochemical change. Arsenic resistance provides a particularly interesting example because the genes within this pathway are likely to have changed over time due to shifts in redox conditions. Arsenite (As^{3+}) resistance, conferred by an arsenite efflux pump encoded by the *arsB* gene, evolved early. Phylogenetic analysis indicates that another gene in the arsenic resistance pathway, *arsA*, was created by duplication of a gene of unknown function. Fusion of the two copies enabled the *arsA* product to bind arsenite and enhance the activity

of the arsenite efflux pump [23]. Arsenate (As^{5+}) toxicity would have become increasingly problematic following the rise of oxygen. Arsenate is often taken up by phosphate transporters but is exported from the cell as arsenite. It is likely that the necessary arsenate reductases evolved from reductases used for other functions. Co-option of the precursor molecule probably followed a gene duplication event that permitted one copy to evolve into an arsenate reductase while the other was available to carry out the original function. Phylogenetic analyses of *ars* genes also reveal the considerable importance of LGT via rapid exchange of plasmids (small, mobile DNA elements) that code for the entire arsenic reduction and efflux pathway, as well as plasmid incorporation into chromosomal DNA [24].

3.3. Interpreting the rock record

DNA degrades rapidly after organisms die, so it is impossible to use nucleic acid sequences from the rock record to date events reflected in phylogenetic trees. Much of what is known about the timing of major microbial innovations and allied changes in biogeochemical cycling comes from the geologic record of lipid biomarkers and isotope signatures. These anchor points have made it possible to begin to time-contour the tree of life [25]. At present, conclusions based on these biosignatures are limited by our incomplete knowledge of lipid biosynthetic and biodegradative pathways and the phylogenetic distributions of genes for isotope-fractionating enzymes.

3.3.1. Correlating lipid biosignatures with genes and organisms

Lipids make excellent molecular fossils because their hydrocarbon-rich chemical structures are resistant to biotic and abiotic degradation. Lipids preserve C and H isotope signatures, and some have structures unique to certain taxonomic groups. Lipid biomarkers extracted from rocks as old as 2.7 billion years include 2α -methylhopanes found in modern cyanobacteria [26] and steranes thought to be synthesized only in eukaryotes. The sterane biomarkers extend the fossil record of eukaryote-like organisms (or at least their

lipid synthesis pathways) by 500–1000 million years [27]. Biomarkers specific for dinoflagellates (single-celled eukaryotes) suggest that they had evolved by the early Cambrian, pushing the origin of this important planktonic lineage back by at least 250 million years relative to the dinoflagellate fossil record [28]. Archaeal-specific lipid biomarkers have been used to infer a mid-Cretaceous expansion of marine lithotrophic (Table 1) archaea [29].

A recent survey of Holocene sediments has shown that tetraether-linked lipids once thought to be unique to thermophilic archaea are ubiquitous in low-temperature environments [30]. We now know that low-temperature archaea commonly produce tetraether-linked lipids. However, we do not know whether non-archaeal groups could have generated these molecules. Until we determine the mechanisms by which diether-linked archaeal lipids are made and joined to form tetraether-linked lipids (Fig. 2) we can neither establish how widespread the ability to make them is nor infer the past distribution of this capability. Surprisingly, the Holocene sediments contained hybrid structures intermediate between bacterial/eukaryotic membrane lipids and archaeal membrane lipids [31]. A tantalizing possibility is that the hybrid lipids were derived from as yet unknown groups of organisms.

It is also possible that the lipid structures that cannot be assigned to known organisms are artifacts of partial degradation or diagenesis. The anaerobic biodegradation of *n*-alkane and isoprenederived hydrocarbons in the same chemical classes as many lipid biomarkers was unknown until very recently [32,33]. The anoxic degradation pathways for saturated isoprenoid compounds (e.g., archaeal membrane lipids) have yet to be elucidated. It remains to be seen whether anaerobic lipid biodegradation pathways could produce artifactual lipid structures that could be preserved in sediments.

3.3.2. Correlating isotope biosignatures with genes and organisms

Isotope signatures preserved in rocks billions of years old can reflect the pathways by which elements were cycled. In the case of biologically

mediated reactions, the isotopic signature also reflects the specific characteristics of enzymes in biochemical reaction networks [34].

In many isotopic studies, fractionation factors for specific processes have been based on measurements for model organisms. However, kinetic isotope-fractionation factors sensitively depend on the specific enzymes involved. Many recent studies highlight the fact that a greater than expected diversity of enzymes can carry out the same biochemical transformation. Genomic data will be essential for cataloging which organisms use which enzymatic pathways, and thus to link the activity of specific organisms with the isotopic record.

Aerobic methane oxidation, an important determinant of redox and greenhouse gas balances, is a good example of how the diversity of enzymes carrying out the same function can lead to large differences in isotope fractionations. Bacteria which oxidize methane at the expense of molecular oxygen use one of two enzymes. The first, called particulate methane monooxygenase (pMMO), is a membrane-bound enzyme with multiple copper atoms in the active site. The second, called soluble methane monooxygenase (sMMO), is found in the cell cytoplasm and contains iron–sulfur clusters and no copper in the active site. The two enzymes are chemically and structurally unrelated, and recent work suggests that pMMO fractionates methane C more strongly than sMMO ($\epsilon_{\text{pMMO}} = -24\%$ vs. $\epsilon_{\text{sMMO}} = -13\%$) [35]. The phylogenetic distribution of the two enzymes is somewhat uncertain, but it appears that genes coding for both enzymes are widespread in the environment [36–39]. Surveys of the MMO genes and genes coding for subsequent enzymatic steps in methane oxidation in natural environments could be used to sort out where and when each enzyme (i.e., which competitive strategy) would be expected to dominate.

Carbon fixation, the biological transformation of CO₂ into cell components (organic matter), is another example of an important biogeochemical transformation associated with diverse pathways (Fig. 4). Most of the carbon fixation by the present-day biosphere involves the enzyme ribulose 1,5-bisphosphate carboxylase/oxygenase (Ru-

BisCO). RuBisCO discriminates strongly against ^{13}C ($\epsilon \sim -25\%$). RuBisCO has two major forms in bacteria and eukaryotes (I and II), and a third type has recently been identified in anoxic archaea [40]. Differences in the RuBisCO enzyme may lead to differences in isotopic fractionations, although work on this effect is in the very early stages. Several RuBisCO forms are abundant in deep-sea vent and seep environments, where carbon-fixing microorganisms and their symbionts support a highly productive biosphere derived from chemical energy. Early evidence suggests that the isotopic compositions of organisms making up some vent communities fall into two clusters ($\delta^{13}\text{C} \sim -34\%$ and $\delta^{13}\text{C} \sim -12\%$), which may correspond to carbon fixation by different RuBisCOs [41].

In addition to RuBisCO, carbon fixation in plants may involve enzymes of the Hatch Slack (C4) pathway, or alternatively may proceed via Crassulacean acid metabolism (CAM). Both the C4 and CAM pathways result in less extreme fractionations than RuBisCO.

The few existing isotopic data on other carbon fixation pathways indicate that they are associated with isotopic fractionations ranging from small (reverse TCA cycle, 2–10%) to large (acetyl-coA pathway, 15–35%; reviewed in [42]). As in the case of RuBisCO, these carbon fixation enzymes have variants which may lead to differences in isotope fractionation. CO dehydrogenase, a key enzyme in the acetyl-coA pathway, comes in at least two varieties: an oxygen-tolerant form containing Mo and Cu, and an oxygen-sensitive form containing Fe and Ni [43]. Likewise, there is evidence that the 3-hydroxypropionate pathway in *Chloroflexus* differs significantly from the chemically analogous pathway in aerobic archaea such as *Metallosphaera*, *Sulfolobus*, and *Acidianus* [44].

It is important to note that observed isotope fractionations at the organismal or ecosystem level reflect the combined effects of isotope fractionations and differential flows within biochemical networks containing many enzymes and branch points. These networks are sufficiently complex that isotopic signatures preserved in the rock record are not simple recorders of particular enzymes

such as RuBisCO or pMMO. Nonetheless, new genomic data will facilitate a more complete analysis of the diversity and distribution of enzymes involved in methane oxidation and carbon fixation. Information about when and where multiple biochemical routes for these transformations evolved will also assist in interpretation of the isotopic record.

4. Modern environments

Probably <1% of all microorganisms have been cultivated. Previous estimates of biological diversity based on cultivation methods have been dwarfed by the number of organisms newly detected by extraction and amplification of ribosomal genes directly from natural environments [45–47]. Discovery of this enormous microbial diversity has motivated the effort to determine metabolic capabilities of the organisms and link their enzymes with specific functions.

4.1. Matching ribosomal gene signatures with geochemical reactions via cultivation-based approaches

Significantly for geoscientists, microorganisms detected in ribosomal gene surveys often mediate novel processes or link elemental cycles in unexpected ways. In some cases, it has been possible to isolate and thus characterize these organisms in terms of their optimal temperature, pH, and solution chemistry for growth and ability to utilize specific substrates. Examples include planctomycetes that carry out anaerobic ammonia oxidation [48], organisms capable of photosynthesis with Fe(II) as the electron donor [49,50], oxidation of sulfide at the expense of nitrate by *Thioploca* [51], iron oxidation in acid environments by archaea and new bacterial species instead of well-studied *Thiobacilli* [52], and detection of abundant lithoautotrophic iron-oxidizing bacteria associated with ocean crust and soils [53]. Many of these reports have been reviewed recently [25,54–56]. At present, most of the genes involved in these geochemical transformations, and their phylogenetic distributions, are unknown.

4.2. Matching uncultivated organisms with geochemical reactions

Many microbial organisms discovered through retrieval of their ribosomal genes have not been isolated, so that their basic biology remains obscure. In anaerobic environments, bacteria in the WS6 division (for which there are no isolated representatives), display high ribosomal gene sequence variability, suggesting a large reservoir of unknown metabolic diversity [57]. Pelagic archaea are numerically abundant in large regions of the ocean, yet we do not know how they fit into marine ecosystems [58]. Bacteria in the *Verrucomicrobia* [59,60] and *Acidobacteria* [61] lineages are numerous in a wide variety of habitats, including soils and sediments. These examples illustrate the importance of methods to connect the uncultivated microorganisms to their ecological roles.

Stable isotope measurements can be used to link microorganisms with processes in the environment, bypassing the need for isolation. In the 1950s Meselson and Stahl [86] showed that DNA molecules can be separated based on their isotopic content using ultracentrifugation. Because DNA contains C, N, O, and P, it is possible to use isotopically labeled compounds as specific tracers for microbial cells that transform the labeled compound in mixed cultures and in the environment. The potential of this approach has been demonstrated using $^{13}\text{C}\text{-CH}_4$ and $^{13}\text{C}\text{-acetate}$ to link $^{13}\text{C}\text{-DNA}$ with the relevant methane-oxidizing and sulfate-reducing bacteria [62].

A new approach for determination of the metabolic potential of uncultivable organisms is to correlate spatially resolved isotopic signatures indicative of specific geochemical transformations with DNA markers (oligonucleotide probes) that indicate groups or species. This method was used to identify consortia of methanogens and sulfate-reducing bacteria mediating anaerobic methane oxidation [63].

Another increasingly popular approach to linking uncultivated organisms and their physiological attributes is to recover large genome fragments from environmental samples and determine gene function via genetic methods [64,65]. The power of this approach is illustrated by discoveries of

protorhodopsin-mediated bacterial photosynthesis in the ocean [66], and aerobic, anoxygenic photosynthesis in ocean gyres [67,68]. Ongoing work in our group is adapting similar cultivation-free genomics methods to retrieve genes from organisms in simple microbial communities from extremely acidic mine-associated environments. The goal is to recover a sufficiently large fraction of the community gene inventory to permit detailed analysis of how the community functions. This approach will provide the data necessary to develop and test ecological models with molecular-scale resolution (Fig. 5).

4.3. Treating geochemical systems as ecosystems: a way forward

New discoveries about specific microorganisms change the way we think about certain aspects of elemental cycling. However, the ultimate challenge is to understand whole ecosystems. We need to analyze the architecture and rules for assembly of biogeochemical systems in order to build good models describing the flow of materials and energy. This requires information about the full range of metabolic processes and the ways in which they are regulated and interrelated. The need for this holistic approach is illustrated by a simple hypothetical example: the rate of sulfate reduction by a pure culture of sulfate-reducing bacteria is unlikely to be the same as that of a sulfate-reducing microbial community characterized by competition for shared resources, predation, symbioses, and consortia.

Ecologists have thought in detail about how biological communities are shaped by interactions among organisms and their environments. These insights may prove useful for analysis of microbial systems. For example, artificial ecosystems created using culturable microorganisms have shown that higher species diversity in the communities results in greater temporal and spatial stability of overall system characteristics such as biomass [69] and rates of processes such as respiration [70]. Diversity is therefore considered to confer resilience, and thus represents a form of ‘biological insurance’ against catastrophic events (e.g., sudden geochemical changes). Artificial microbial

communities have also shown that competition and predation have direct consequences for productivity [71]. New methods for monitoring population size, structure, and organism activity make it possible to test these concepts in the environment.

The conceptual framework for ecological theories has been based largely on observations of macroscopic biological systems, despite the fact that these systems are underpinned by a diverse microbial biosphere. At present, we do not know if these ideas apply to the microbial component, or apply in microbially dominated extreme or ancient environments. Microorganisms can occupy far more diverse niches (pressure, temperature, pH, energy sources, etc.), potentially increasing efficiency of resource use and ecosystem resilience. Specific microorganisms within a large dormant pool can rapidly reactivate when conditions change, perhaps shortening timescales for succession. Microorganism–microorganism signaling, sharing of genetic information, and resource interdependence may be fundamentally different than analogous processes among macroorganisms. In addition, microorganisms can respond to environmental perturbation via genome changes so much faster than macroorganisms that microbial evolution becomes a relevant factor in ecosystem models. For example, in times of extreme stress, microorganisms may induce formation of large numbers of mutations within their genes, each of which has a low probability of successfully generating a new, advantageous molecule or pathway [72,73]. However, if only one cell in billions successfully generates a mutation that allows the organism to survive, rapid rates of cell division lead to re-establishment of the community. For the above reasons, microbial communities may not be well-described by existing ecological theories. Alternatively, microorganisms may fit ecological theories better than macroorganisms because the observed results sample billions of individuals rather than small populations and so are less subject to random effects.

Microbially dominated habitats from some extreme environments should be especially suitable for molecularly resolved ecological studies because they contain few species and it will be prac-

tical to monitor the activity of the genes of all community members via genome-enabled gene expression arrays (Fig. 5). The potential of this approach is yet to be realized. Given that many geochemical processes have been fundamentally impacted by microbial activity over most of Earth history, the results from such studies may have far-reaching implications for the study of ancient biogeochemical cycles.

5. Speculative questions to ponder

In this article we have reviewed the ways in which genomic data are reshaping our understanding of the roles of microorganisms in modern and ancient biogeochemical cycling. We have also discussed how genomic data implicate lateral gene transfer as a driving force for microbial evolution. However, many questions about the coupling between biogeochemical systems and microbial evolution remain.

How do environmental factors impact the rates of microorganism evolution via lateral gene transfer? We have raised the question of whether periods of rapid microbial evolution correlate with periods of rapid macroorganism evolution documented in the fossil record. Microorganisms may not undergo rapid speciation following many catastrophic events because of their widespread distribution, their ability to rapidly reassemble into communities to populate new habitats, and their ability to survive unfavorable conditions in a dormant state. There may be as yet poorly appreciated connections between environmental perturbations and microbial evolution by LGT. Some microorganisms possess efficient gene capture systems (e.g., integrons in *Vibrio cholerae* [74]), others are receptive to uptake of plasmids [75], and yet others are especially prone to take up or delete foreign DNA [76]. These characteristics contribute to very significant differences in background evolutionary rates within modern microbial lineages. These rates might be dramatically perturbed by environmental stimuli. For example, pollution by arsenic pesticides stimulates rapid exchange of arsenic resistance genes via plasmids, which may also carry genes for antibiotic resis-

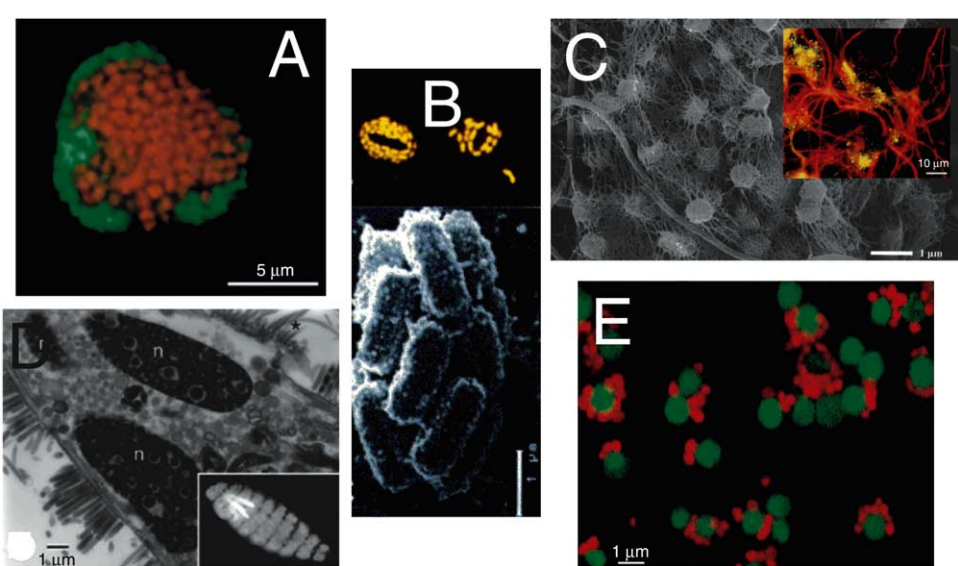
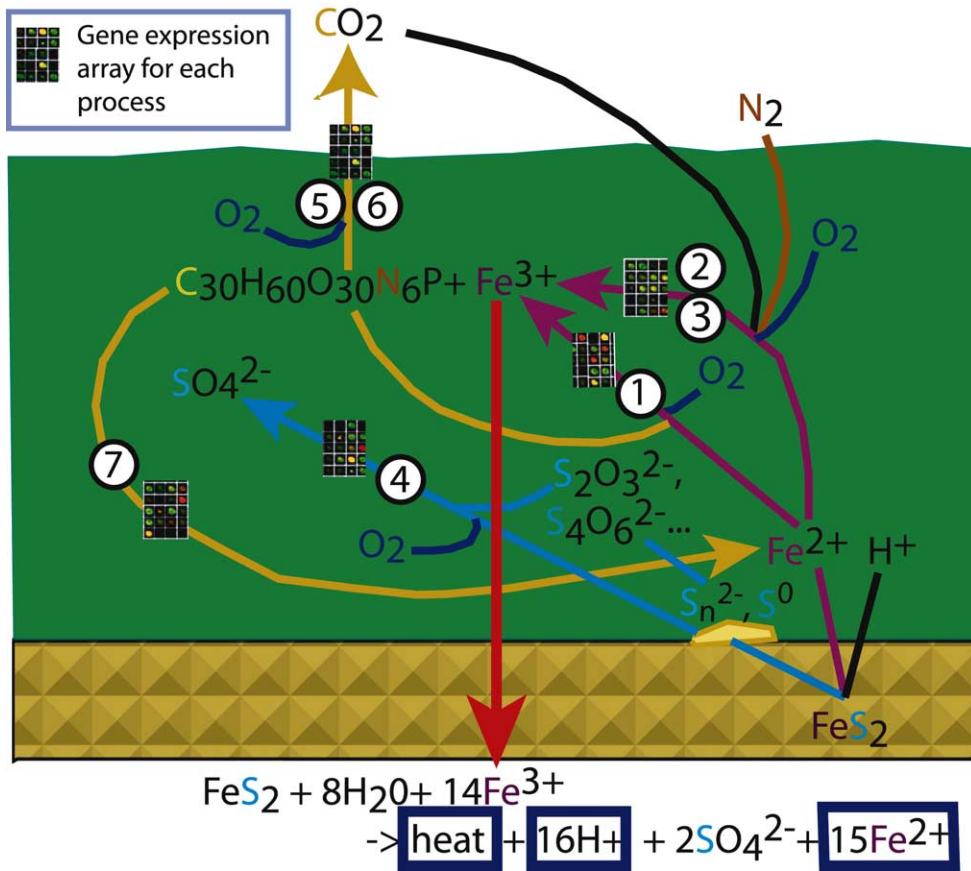


Fig. 5. An important objective for biogeochemical studies is the development of ecosystem models that explicitly include microorganisms. Ideally, these models should incorporate the rates of geochemical and biological processes and specify links between organisms, transformations, resources, and byproducts. In ‘molecularly resolved’ models, molecule-specific (geochemical species and substrates) fluxes are defined and pathways are resolved at the level of the enzymes that mediate them. This detailed information will be essential in order to fully understand the dynamic interactions that define higher levels of system organization. These include competition, predation, and symbiosis, and feedbacks between organism activity and rates of geochemical reactions. The environments where this approach is currently practical are likely to be microbially dominated, geochemically simple, and contain few organisms. Like many extreme environments, acid environments are excellent candidates. Studies of simple ecosystems may be applicable to more complex systems to determine how ecosystem perturbation impacts CO₂ sequestration, methane consumption, mineral weathering rates, etc. The figure shows an example of fluxes and feedbacks in an acid mine drainage ecosystem model. Organism 1 is a lithoheterotroph, 2, 3, and 4 are lithoautotrophs, 5 and 6 are heterotrophs (see Table 1). Note that 6 is an anaerobic heterotroph (using organic carbon as the electron donor and ferric iron as the acceptor), and would only be found in anoxic niches. The red arrow indicates one of the key feedbacks: increase in the rate of pyrite (FeS₂) dissolution due to re-supply of oxidant via biological-catalysis of iron oxidation. Each microscopic gene array (silicon ‘gene chip’) is used to monitor the activity of all the genes of all the organisms associated with the transformation of interest, e.g., iron oxidation, carbon fixation, nitrogen fixation, sulfur oxidation, etc. Colored spots signify the level of expression of a particular gene. The activity levels (organism-resolved rates) are key model inputs. Within the master equation, blue boxes signify major ecosystem parameters to be predicted.

tance and other functions. Geological phenomena may also increase rates of LGT. Uptake of DNA from solution may be accelerated by thermal or electrical shock or changes in concentrations of ions that inhibit DNA hydrolysis. Additional links between major geologic events and microbial evolution may arise if, as has been proposed, organisms respond to stress by self-mutation [73].

Is gene-swapping (LGT) an important mechanism for recovery following perturbations in biogeochemical systems? LGT may be particularly relevant in cases where re-colonization by organisms from adjacent habitats is restricted (e.g., by extreme conditions). Re-distribution of genes within a community can occur by uptake of naked DNA (transformation) and conjugation (cell-to-cell transfer of DNA via formation of a cytoplasmic bridge), both of which require relatively close cell proximity. Lateral transfer of

genes from organisms in distant environments can occur via viruses (transduction; [55]). The ease with which genes that confer new traits move within and between habitats may directly impact the ways in which biogeochemical systems respond to perturbation and evolve. The role of LGT in ecosystem recovery will be resolved by targeted studies of LGT in modern environments.

How much LGT is due to intimate physical associations between organisms? Symbioses are important in many microbially dominated biogeochemical systems, especially in biofilms and other layered microbial communities (Fig. 6). There are intriguing indications that endosymbiosis involving bacterial and eukaryal partners is widespread. Examples include the observation that certain protists (microbial eukaryotes) contain chloroplast-derived structures (apicoplasts) with relic genomes that enable the protists to live as parasites

Fig. 6. Notable examples of intimate physical associations that appear to have metabolic importance. (A) shows a consortium of sulfate-reducing bacteria (green) and methanogens (red) that carry out anaerobic methane oxidation [79]. (B) β -Proteobacterium (central cell) surrounded by green-sulfur bacteria [80] (reprinted from *FEMS: Microbiology Reviews*, Vol. 24, J. Overmann and H. van Gernerden, Microbial interactions involving sulfur bacteria: implications for the ecology and evolution of bacterial communities, pp. 591–599, 2000, with kind permission from the authors and Elsevier). The consortium carries out photosynthesis utilizing sulfide as the electron donor. Examples of consortia where the form of the interrelationship is currently unknown are (C) *Thiothrix* bacterial filaments (red) and unknown archaeal cells (green) [81,82] (reproduced from *Applied and Environmental Microbiology*, Vol. 67, p. 2336 and Vol. 68, p. 933, with kind permission of the authors and the American Society for Microbiology; photomicrographs provided by Robert Huber and Gerhard Wanner). (D) A protist and unknown surface-attached bacteria [83]. (E) A consortium of two archaea species: *Ignioccus* (green) and ‘*Nanoarchaeum*’ [84]. (Images in A, D and E reproduced from *Nature*, Vol. 407, pp. 623–626 (<http://www.nature.com>) with kind permission of the authors and publisher.)

[77], and evidence that some anaerobic protists contain hydrogenosomes (organelles for anaerobic energy generation via fermentation) that were acquired as the result of multiple endosymbiosis events in different lineages [78]. There is good evidence for transfers of genes between eukaryote nuclei and bacterially derived mitochondria or chloroplasts. At the moment, it is unclear how much LGT is taking place within these newly discovered endosymbioses. We will be able to evaluate the extent to which LGT accompanies close physical associations when we learn more about the genetics of microbial communities, including their microscopic eukaryote members.

6. Conclusions

Discoveries of the past decade reinforce the notion that microbial life has been able to colonize every habitat where biochemical molecules can function and to generate energy using virtually every thermodynamically favorable energy couple. This accomplishment must be attributed to fundamental connections between genome evolution and the geochemical environment. We have glimpsed the molecular basis for this coupling, and see profound implications for our understanding of biogeochemical cycling on local, global, and extraterrestrial scales.

Acknowledgements

We thank Philip Hugenholtz, Mary Power, Rachna Ram, Tom Hanson, and Clara Chan for their helpful comments on this manuscript. We also thank Roger Summons, Frank Spear, and Everett Shock for their insightful reviews. Preparation of this manuscript was partially supported by NSF, DOE, and NASA grants. [AH]

References

- [1] B.L. Phillips, W.H. Casey, M. Karlsson, Bonding and reactivity at oxide mineral surfaces from model aqueous complexes, *Nature* 404 (2000) 379–382.
- [2] W.B. Whitman, D.C. Coleman, W.J. Wiebe, Prokaryotes: the unseen majority, *Proc. Natl. Acad. Sci. USA* 95 (1998) 6578–6583.
- [3] J.R. Brown, C.J. Douady, M.J. Italia, W.E. Marshall, M.J. Stanhope, Universal trees based on large combined protein sequence data sets, *Nat. Genet.* 28 (2001) 281–285.
- [4] C. Brochier, H. Philippe, A non-hyperthermophilic ancestor for Bacteria, *Nature* 417 (2002) 244.
- [5] E. Pennisi, Genome data shake tree of life, *Science* 280 (1998) 672–674.
- [6] G. Deckert, P.V. Warren, T. Gaasterland, W.G. Young, A.L. Lenox, D.E. Graham, R. Overbeek, M.A. Snead, M. Keller, M. Aujay, R. Huber, R.A. Feldman, J.M. Short, G.J. Olsen, R.V. Swanson, The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*, *Nature* 392 (1998) 353–358.
- [7] C. Woese, The universal ancestor, *Proc. Natl. Acad. Sci. USA* 95 (1998) 6854–6859.
- [8] C.R. Woese, G.J. Olsen, M. Ibba, D. Soll, Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process, *Microbiol. Mol. Biol. Rev.* 64 (2000) 202–236.
- [9] K.E. Nelson, R.A. Clayton, S.R. Gill, M.L. Gwinn, R.J. Dodson, D.H. Haft, E.K. Hickey, J.D. Peterson, W.C. Nelson, K.A. Ketchum, L. McDonald, T.R. Utterback, J.A. Malek, K.D. Linher, M.M. Garrett, A.M. Stewart, M.D. Cotton, M.S. Pratt, C.A. Phillips, D. Richardson, J. Heidelberg, G.G. Sutton, R.D. Fleischmann, J.A. Eisen, O. White, S.L. Salzberg, H.O. Smith, J.C. Venter, C.M. Fraser, Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of *Thermotoga maritima*, *Nature* 399 (1999) 323–329.
- [10] K.S. Makarova, L. Aravind, Y.I. Wolf, R.L. Tatusov, K.W. Minton, E.V. Koonin, M.J. Daly, Genome of the extremely radiation-resistant bacterium *Deinococcus radiodurans* viewed from the perspective of comparative genomics, *Microbiol. Mol. Biol. Rev.* 65 (2001) 44–79.
- [11] H. Ochman, J.G. Lawrence, E.A. Groisman, Lateral gene transfer and the nature of bacterial innovation, *Nature* 405 (2000) 299–304.
- [12] W.F. Doolittle, Phylogenetic classification and the universal tree, *Science* 284 (1999) 2124–2128.
- [13] A.K. Bansal, T.E. Meyer, Evolutionary analysis by whole-genome comparisons, *J. Bacteriol.* 184 (2002) 2260–2272.
- [14] S.T. Fitz-Gibbon, C.H. House, Whole genome-based phylogenetic analysis of free-living microorganisms, *Nucleic Acids Res.* 27 (1999) 4218–4222.
- [15] B. Snel, P. Bork, M.A. Huynen, Genome phylogeny based on gene content, *Nat. Genet.* 21 (1999) 108–110.
- [16] F. Tekaia, A. Lazcano, B. Dujon, The genomic tree as revealed from whole proteome comparisons, *Genome Res.* 9 (1999) 550–557.
- [17] G.D.P. Clarke, R.G. Beiko, M.A. Ragan, R.L. Charlebois, Inferring genome trees by using a filter to eliminate phylogenetically discordant sequences and a distance ma-

- trix based on mean normalized BLASTP scores, *J. Bacteriol.* 184 (2002) 2072–2080.
- [18] C.R. Woese, O. Kandler, M.L. Wheelis, Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya, *Proc. Natl. Acad. Sci. USA* 87 (1990) 4576–4579.
- [19] Y. Shen, R. Buick, D.E. Canfield, Isotopic evidence for microbial sulphate reduction in the early Archaean era, *Nature* 410 (2001) 77–81.
- [20] M. Klein, M. Friedrich, A.J. Roger, P. Hugenholtz, S. Fishbain, H. Abicht, L.L. Blackall, D.A. Stahl, M. Wagner, Multiple lateral transfers of dissimilatory sulfite reductase genes between major lineages of sulfate-reducing prokaryotes, *J. Bacteriol.* 183 (2001) 6028–6035.
- [21] J. Xiong, W.M. Fischer, K. Inoue, M. Nakahara, C.E. Bauer, Molecular evidence for the early evolution of photosynthesis, *Science* 289 (2000) 1724–1730.
- [22] D.J. Lessner, G.R. Johnson, R.E. Parales, J.C. Spain, D.T. Gibson, Molecular characterization and substrate specificity of nitrobenzene dioxygenase from *Comamonas* sp. strain JS765, *Appl. Environ. Microbiol.* 68 (2002) 634–641.
- [23] B.P. Rosen, Families of arsenic transporters, *Trends Microbiol.* 7 (1999) 207–212.
- [24] T.M. Gihring, P.L. Bond, S. Peters, J.F. Banfield (2002) Arsenic resistance in the archaeon *Ferroplasma acidarmanus*: new insights into the structure and evolution of arsenic genes, *Extremophiles* (in press).
- [25] J.F. Banfield, C.R. Marshall, Earth science and evolution: genomics and the geosciences, *Science* 287 (2000) 605–606.
- [26] R.E. Summons, L.L. Jahnke, J.M. Hope, G.A. Logan, 2-Methylhopanoids as biomarkers for cyanobacterial oxygenic photosynthesis, *Nature* 400 (1999) 554–557.
- [27] J.J. Brocks, G.A. Logan, R. Buick, R.E. Summons, Archean molecular fossils and the early rise of eukaryotes, *Science* 285 (1999) 1033–1036.
- [28] J.M. Moldovan, N.M. Talyzina, Biogeochemical evidence for the dinoflagellate ancestors in the Early Cambrian, *Science* 281 (1998) 1168–1170.
- [29] M.M.M. Kuypers, P. Blokker, J. Erbacher, H. Kinkel, R.D. Pancost, S. Schouten, J.S. Sinningh-Damste, Massive expansion of marine Archaea during a mid-Cretaceous oceanic anoxic event, *Science* 293 (2001) 92–94.
- [30] S. Schouten, E.C. Hopmans, R.D. Pancost, J.S. Sinningh-Damste, Widespread occurrence of structurally diverse tetraether membrane lipids: evidence for the ubiquitous presence of low-temperature relatives of hyperthermophiles, *Proc. Natl. Acad. Sci. USA* 97 (2000) 14421–14426.
- [31] J.M. Hayes, Lipids as a common interest of microorganisms and geochemists, *Proc. Natl. Acad. Sci. USA* 97 (2000) 14033–14034.
- [32] F. Widdel, R. Rabus, Anaerobic biodegradation of saturated and aromatic hydrocarbons, *Curr. Opin. Biotechnol.* 12 (2001) 259–276.
- [33] T.P.-A. Bregnard, A. Haner, P. Hohener, J. Zeyer, Anaerobic degradation of pristane in nitrate-reducing microcosms and enrichment cultures, *Appl. Environ. Microbiol.* 63 (1997) 2077–2081.
- [34] J.M. Hayes, Factors controlling C-13 contents of sedimentary organic compounds: principles and evidence, *Mar. Geol.* 113 (1993) 111–125.
- [35] L.L. Jahnke, R.E. Summons, J.M. Hope, D.J. DesMarais, Carbon isotopic fractionation in lipids from methanotrophic bacteria. II. The effects of physiology and environmental parameters on the biosynthesis and isotopic signatures of biomarkers, *Geochim. Cosmochim. Acta* 63 (1999) 79–93.
- [36] T. Shigematsu, S. Hanada, M. Eguchi, Y. Kamagata, T. Kanagawa, R. Kurane, Soluble methane monooxygenase gene clusters from trichloroethylene-degrading *Methylomonas* sp. strains and detection of methanotrophs during in situ bioremediation, *Appl. Environ. Microbiol.* 65 (1999) 5198–5206.
- [37] S.C. Koh, J.P. Bowman, G.S. Sayler, Soluble methane monooxygenase production and trichloroethylene degradation by a Type I methanotroph *Methylomonas-methanica* 68-1, *Appl. Environ. Microbiol.* 59 (1993) 960–967.
- [38] T. Iwamoto, K. Tani, K. Nakamura, Y. Suzuki, M. Kitagawa, M. Eguchi, M. Nasu, Monitoring impact of in situ biostimulation treatment on groundwater bacterial community by DGGE, *FEMS Microbiol. Ecol.* 32 (2000) 129–141.
- [39] H. Fuse, M. Ohta, O. Takimura, K. Murakami, H. Inoue, Y. Yamaoka, J.M. Oclarit, T. Omori, Oxidation of trichloroethylene and dimethyl sulfide by a marine *Methylobacterium* strain containing soluble methane monooxygenase, *Biosci. Biotechnol. Biochem.* 62 (1998) 1925–1931.
- [40] G.M.F. Watson, J.-P. Yu, F.R. Tabita, Unusual ribulose 1,5-bisphosphate carboxylase/oxygenase of anoxic Archaea, *J. Bacteriol.* 181 (1999) 1569–1575.
- [41] K.M. Scott, J.J. Robinson, D.T. Nguyen, C.M. Cavanaugh, Form I and Form II Rubisco in chemoautotrophic bacteria: influence on stable carbon isotope values, *Abstr. Gen. Meet. Am. Soc. Microbiol.* 99 (1999) 383.
- [42] M.T. van der Meer, S. Schouten, W.I. Rijpstra, G. Fuchs, J.S. Sinningh-Damste, Stable carbon isotope fractionations of the hyperthermophilic crenarchaeon *Metallorubra sedula*, *FEMS Microbiol. Lett.* 196 (2001) 67–70.
- [43] R.K. Thauer, Enzymology nickel to the fore, *Science* 293 (2001) 1264–1265.
- [44] C. Menendez, Z. Bauer, H. Huber, N. Gad'on, K.-O. Stetter, G. Fuchs, Presence of acetyl coenzyme A (CoA) carboxylase and propionyl-CoA carboxylase in autotrophic Crenarchaeota and indication for operation of a 3-hydroxypropionate cycle in autotrophic carbon fixation, *J. Bacteriol.* 181 (1999) 1088–1098.
- [45] N.R. Pace, A molecular view of microbial diversity and the biosphere, *Science* 276 (1997) 734–740.
- [46] D. Moreira, P. Lopez-Garcia, The molecular ecology of microbial eukaryotes unveils a hidden world, *Trends Microbiol.* 10 (2002) 31–38.

- [47] E.F. De Long, N.R. Pace, Environmental diversity of bacteria and archaea, *Syst. Biol.* 50 (2001) 470–478.
- [48] M. Strous, J.A. Fuerst, E.H.M. Kramer, S. Logemann, G. Muyzer, K.T. van de Pas-Schoonen, R. Webb, J.G. Kuenen, M.S.M. Jetten, Missing lithotroph identified as new planctomycete, *Nature* 400 (1999) 446–449.
- [49] F. Widdel, S. Schnell, S. Heising, A. Ehrenreich, B. Assmus, B. Schink, Ferrous iron oxidation by anoxygenic phototrophic bacteria, *Nature* 362 (1993) 834–836.
- [50] S. Heising, L. Richter, W. Ludwig, B. Schink, *Chlorobium ferrooxidans* sp. nov., a phototrophic green sulfur bacterium that oxidizes ferrous iron in coculture with a 'Geospirillum' sp. strain, *Arch. Microbiol.* 172 (1999) 116–124.
- [51] B.B. Jorgensen, V.A. Gallardo, *Thioploca* spp. Filamentous sulfur bacteria with nitrate vacuoles, *FEMS Microbiol. Ecol.* 28 (1999) 301–313.
- [52] K.J. Edwards, P.L. Bond, T.M. Gihring, J.F. Banfield, An archaeal iron-oxidizing extreme acidophile important in acid mine drainage, *Science* 287 (2000) 1796–1799.
- [53] D. Emerson, C. Moyer, Neutrophilic Fe-oxidizing bacteria are abundant at the Loihi Seamount hydrothermal vents and play a major role in Fe oxide deposition, *Appl. Environ. Microbiol.* 68 (2002) 3085–3093.
- [54] D.K. Newman, J.F. Banfield, Geomicrobiology: how molecular-scale interactions underpin biogeochemical systems, *Science* 296 (2002) 1071–1077.
- [55] E.G. Nisbet, N.H. Sleep, The habitat and nature of early life, *Nature* 409 (2001) 1083–1091.
- [56] A.-L. Reysenbach, E. Shock, Merging genomes with geochemistry in hydrothermal ecosystems, *Science* 296 (2002) 1077–1082.
- [57] M.A. Dojka, J.K. Harris, N.R. Pace, Expanding the known diversity and environmental distribution of an uncultured phylogenetic division of bacteria, *Appl. Environ. Microbiol.* 66 (2000) 1617–1621.
- [58] M.B. Karner, E.F. De Long, D.M. Karl, Archaeal dominance in the mesopelagic zone of the Pacific Ocean, *Nature* 409 (2001) 507–510.
- [59] A. Felske, A.D.L. Akkermans, Prominent occurrence of ribosomes from an uncultured bacterium of the Verrucomicrobiales cluster in grassland soils, *Letts. Appl. Microbiol.* 26 (1998) 219–223.
- [60] D.H. Buckley, T.M. Schmidt, Environmental factors influencing the distribution of rRNA from Verrucomicrobia in soil, *FEMS Microbiol. Ecol.* 35 (2001) 105–112.
- [61] S.M. Barns, S.L. Takala, C.R. Kuske, Wide distribution and diversity of members of the bacterial kingdom Acidobacterium in the environment, *Appl. Environ. Microbiol.* 65 (1999) 1731–1737.
- [62] S. Radajewski, P. Ineson, N.R. Parekh, J.C. Murrell, Stable-isotope probing as a tool in microbial ecology, *Nature* 403 (2000) 646–649.
- [63] V.J. Orphan, C.H. House, K.-U. Hinrichs, K.D. McKeegan, E.F. De Long, Methane-consuming archaea revealed by directly coupled isotopic and phylogenetic analysis, *Science* 293 (2001) 484–487.
- [64] O. Beja, E.V. Koonin, L. Aravind, L.T. Taylor, H. Seitz, J.L. Stein, D.C. Bensen, R.A. Feldman, R.V. Swanson, E.F. De Long, Comparative genomic analysis of archaeal genotypic variants in a single population and in two different oceanic provinces, *Appl. Environ. Microbiol.* 68 (2002) 335–345.
- [65] M.R. Rondon, P.R. August, A.D. Bettermann, S.F. Brady, T.H. Grossman, M.R. Liles, K.A. Loiacono, B.A. Lynch, I.A. MacNeil, C. Minor, C.L. Tiong, M. Gilman, M.S. Osburne, J. Clardy, J. Handelsman, R.M. Goodman, Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms, *Appl. Environ. Microbiol.* 66 (2000) 2541–2547.
- [66] O. Beja, L. Aravind, E.V. Koonin, M.T. Suzuki, A. Hadd, L.P. Nguyen, S.B. Jovanovich, C.M. Gates, R.A. Feldman, J.L. Spudich, E.N. Spudich, E.F. De Long, Bacterial rhodopsin: evidence for a new type of phototrophy in the sea, *Science* 289 (2000) 1902–1906.
- [67] O. Beja, M.T. Suzuki, J.F. Heidelberg, W.C. Nelson, C.M. Preston, T. Hamada, J.A. Eisen, C.M. Fraser, E.F. De Long, Unsuspected diversity among marine aerobic anoxygenic phototrophs, *Nature* 415 (2002) 630–633.
- [68] Z.S. Kolber, C.L. Van Dover, R.A. Niederman, P.G. Falkowski, Bacterial photosynthesis in surface waters of the open ocean, *Nature* 407 (2000) 177–179.
- [69] S. Naeem, S. Li, Biodiversity enhances ecosystem reliability, *Nature* 390 (1997) 507–509.
- [70] J. McGrady-Steed, P.M. Harris, P.J. Morin, Biodiversity regulates ecosystem predictability, *Nature* 390 (1997) 162–165.
- [71] B.J.M. Bohannan, R.E. Lenski, The relative importance of competition and predation varies with productivity in a model community, *Am. Nat.* 156 (2000) 329–340.
- [72] P. Capy, G. Gasperi, C. Biemont, C. Bazin, Stress and transposable elements: co-evolution or useful parasites?, *Heredity* 85 (2000) 101–106.
- [73] H. Ilves, R. Horak, M. Kivisaar, Involvement of sigmas in starvation-induced transposition of *Pseudomonas putida* transposon Tn4652, *J. Bacteriol.* 183 (2001) 5445–5448.
- [74] D. Mazel, B. Dychinco, V.A. Webb, J. Davies, A distinctive class of integron in the *Vibrio cholerae* genome, *Science* 280 (1998) 605–608.
- [75] A. Campbell, J. Mrazek, S. Karlin, Genome signature comparisons among prokaryote, plasmid, and mitochondrial DNA, *Proc. Natl. Acad. Sci. USA* 96 (1999) 9184–9189.
- [76] D. Dubnau, DNA uptake in bacteria, *Annu. Rev. Microbiol.* 53 (1999) 217–244.
- [77] C.Y. He, B. Striepen, C.H. Pletcher, J.M. Murray, D.S. Roos, Targeting and processing of nuclear-encoded apicoplast proteins in plastid segregation mutants of *Toxoplasma gondii*, *J. Biol. Chem.* 276 (2001) 28436–28442.
- [78] W. Martin, M. Mueller, The hydrogen hypothesis for the first eukaryote, *Nature* 392 (1998) 37–41.

- [79] A. Boetius, K. Ravenschlag, C.J. Schubert, D. Rickert, F. Widdel, A. Gieseke, R. Amann, B.B. Jorgensen, U. Witte, O. Pfannkuche, A marine microbial consortium apparently mediating anaerobic oxidation of methane, *Nature* 407 (2000) 623–626.
- [80] J. Overmann, H. van Gernerden, Microbial interactions involving sulfur bacteria implications for the ecology and evolution of bacterial communities, *FEMS Microbiol. Rev.* 24 (2000) 591–599.
- [81] C. Rudolph, G. Wanner, R. Huber, Natural communities of novel archaea and bacteria growing in cold sulfurous springs with a string-of-pearls-like morphology, *Appl. Environ. Microbiol.* 67 (2001) 2336–2344.
- [82] C. Moissl, C. Rudolph, R. Huber, Natural communities of novel archaea and bacteria with a string-of-pearls-like morphology: molecular analysis of the bacterial partners, *Appl. Environ. Microbiol.* 68 (2002) 933–937.
- [83] J.M. Bernhard, K.R. Buck, M.A. Farmer, S.S. Bowser, The Santa Barbara Basin is a symbiosis oasis, *Nature* 403 (2000) 77–80.
- [84] H. Huber, M.J. Hohn, R. Rachel, T. Fuchs, V.C. Wimmer, K.O. Stetter, A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont, *Nature* 417 (2002) 63–67.
- [85] P. Hugenholtz, Exploring prokaryotic diversity in the genomic era, *Genome Biology* 3 (2002) 3.1–3.8.
- [86] M. Meselson, F. Stahl, The replication of DNA in *Escherichia coli*, *Proc. Nat. Acad. Sci. USA* 44 (1958) 671–682.



Jillian F. Banfield is a Professor in the Departments of Environmental Science, Policy, and Management and Earth and Planetary Science at the University of California, Berkeley, and holds an appointment at Lawrence Berkeley National Laboratory. She received B.Sc. and M.Sc. degrees in Geology at the Australian National University and a Ph.D. in Earth and Planetary Sciences from Johns Hopkins University. Her current research focuses on the mineralogy and biomineralization of nanocrystals, the genomics and microbial ecology of extremely acidic environments, and microbial and geochemical controls on metals and mineral weathering in the environment.



Jennifer Macalady is an Assistant Professor in the Department of Geology at Carleton College in Northfield, Minnesota. She received a B.A. in Geology from Carleton College, and M.S. and Ph.D. degrees in Soil Science from the University of California at Davis. Her current research projects involve the geomicrobiology of soils and weathering profiles, stable isotopic tracer tools for microbial ecology, and the genetics and lipid chemistry of microbial membranes.