

The Microbial Engines That Drive Earth's Biogeochemical Cycles

Paul G. Falkowski,^{1*} Tom Fenchel,^{2*} Edward F. Delong^{3*}

Virtually all nonequilibrium electron transfers on Earth are driven by a set of nanobiological machines composed largely of multimeric protein complexes associated with a small number of prosthetic groups. These machines evolved exclusively in microbes early in our planet's history yet, despite their antiquity, are highly conserved. Hence, although there is enormous genetic diversity in nature, there remains a relatively stable set of core genes coding for the major redox reactions essential for life and biogeochemical cycles. These genes created and coevolved with biogeochemical cycles and were passed from microbe to microbe primarily by horizontal gene transfer. A major challenge in the coming decades is to understand how these machines evolved, how they work, and the processes that control their activity on both molecular and planetary scales.

Earth is ~4.5 billion years old, and during the first half of its evolutionary history, a set of metabolic processes that evolved exclusively in microbes would come to alter the chemical speciation of virtually all elements on the planetary surface. Consequently, our current environment reflects the historically integrated outcomes of microbial experimentation on a tectonically active planet endowed with a thin film of liquid water (1). The outcome of these experiments has allowed life to persist even though the planet has been subjected to extraordinary environmental changes, from bolide impacts and global glaciations to massive volcanic outgassing (2). Although such perturbations led to major extinctions of plants and animals (3), to the best of our knowledge, the core biological machines responsible for planetary biogeochemical cycles have survived intact.

The explosion of microbial genome sequence data and increasingly detailed analyses of the structures of key machines (4) has yielded insight into how microbes became the biogeochemical engineers of life on Earth. Nevertheless, a grand challenge in science is to decipher how the ensemble of the core microbially derived machines evolved and how they interact, and the mechanisms regulating their operation and maintenance of elemental cycling on Earth. Here we consider the core set of genes responsible for fluxes of key elements on Earth in the context of a global metabolic pathway.

Essential Geophysical Processes for Life

On Earth, tectonics and atmospheric photochemical processes continuously supply substrates and remove products, thereby creating geochemical cy-

cles (5, 6). These two geophysical processes allow elements and molecules to interact with each other, and chemical bonds to form and break in a cyclical manner. Indeed, unless the creation of bonds forms a cycle, planetary chemistry ultimately will come to thermodynamic equilibrium, which would lead inevitably to a slow depletion of substrates essential for life on the planetary surface. Most of the H₂ in Earth's mantle escaped to space early in Earth's history (7); consequently, the overwhelming majority of the abiotic geochemical reactions are based on acid/base chemistry, i.e., transfers of protons without electrons. The chemistry of life, however, is based on redox reactions, i.e., successive transfers of electrons and protons from a relatively limited set of chemical elements (6).

The Major Biogeochemical Fluxes Mediated by Life

Six major elements—H, C, N, O, S, and P—constitute the major building blocks for all biological macromolecules (8). The biological fluxes of the first five of these elements are driven largely by microbially catalyzed, thermodynamically constrained redox reactions (Fig. 1). These involve two coupled half-cells, leading to a linked system of elemental cycles (5). On geological time scales, resupply of C, S, and P is dependent on tectonics, especially volcanism and rock weathering (Fig. 1). Thus, biogeochemical cycles have evolved on a planetary scale to form a set of nested abiotically driven acid-base and biologically driven redox reactions that set lower limits on external energy required to sustain the cycles. These reactions fundamentally altered the surface redox state of the planet. Feedbacks between the evolution of microbial metabolic and geochemical processes create the average redox condition of the oceans and atmosphere. Hence, Earth's redox state is an emergent property of microbial life on a planetary scale. The biological oxidation of Earth is driven by photosynthesis, which is the only known energy transduction process that is not directly dependent on preformed bond energy (9).

The fluxes of electrons and protons can be combined with the six major elements to construct

a global metabolic map for Earth (Fig. 2). The genes encoding the machinery responsible for the redox chemistry of half-cells form the basis of the major energy-transducing metabolic pathways. The contemporary pathways invariably require multimeric protein complexes (i.e., the microbial "machines") that are often highly conserved at the level of primary or secondary structure. These complexes did not evolve instantaneously, yet the order of their appearance in metabolism and analysis of their evolutionary origins are obscured by lateral gene transfer and extensive selection. These processes make reconstruction of how electron transfer reactions came to be catalyzed extremely challenging (10).

In many cases, identical or near-identical pathways may be used for the forward and reverse reactions required to maintain cycles. For example, methane is formed by methanogenic Archaea from the reduction of CO₂ with H₂. If the hydrogen tension is sufficiently low, however, then the reverse process becomes thermodynamically favorable; methane is oxidized anaerobically by Archaea closely related to known, extant methanogens that apparently use co-opted methanogenic machinery in reverse. Low hydrogen tension occurs when there is close spatial association with hydrogen-consuming sulfate reducers (11–13); thus, this process requires the synergistic cooperation of multispecies assemblages, a phenomenon that is typical for most biogeochemical transformations. Similarly, the citric acid cycle oxidizes acetate stepwise into CO₂ with a net energy yield. In green sulfur bacteria, and in some Archaeobacteria, the same cycle is used to assimilate CO₂ into organic matter with net energy expenditure. Indeed, this may have been the original function of that cycle (14). Typically, in one direction, the pathway is oxidative, dissimilatory, and produces adenosine 5'-triphosphate, and in the opposite direction, the pathway is reductive, assimilatory, and energy consuming.

However, reversible metabolic pathways in biogeochemical cycles are not necessarily directly related, and sometimes are catalyzed by diverse, multispecies microbial interactions. The various oxidation and reduction reactions that drive Earth's nitrogen cycle (which, before humans, was virtually entirely controlled by microbes) are a good example. N₂ is a highly inert gas, with an atmospheric residence time of ~1 billion years. The only biological process that makes N₂ accessible for the synthesis of proteins and nucleic acids is nitrogen fixation, a reductive process that transforms N₂ to NH₄⁺. This biologically irreversible reaction is catalyzed by an extremely conserved heterodimeric enzyme complex, nitrogenase, which is inhibited by oxygen (15). In the presence of oxygen, NH₄⁺ can be oxidized to nitrate in a two-stage pathway, initially requiring a specific group of Bacteria or Archaea that oxidize ammonia to NO₂⁻ (via hydroxylamine), which is subsequently oxidized to NO₃⁻ by a different suite of nitrifying bacteria (16). All of the nitrifiers use the small differences in redox potential in the oxidation reactions to reduce CO₂ to

¹Environmental Biophysics and Molecular Ecology Program, Institute of Marine and Coastal Sciences and Department of Earth and Planetary Sciences, Rutgers University, New Brunswick, NJ 08901, USA. ²Marine Biological Laboratory, University of Copenhagen, Strandpromenaden 5, DK-3000 Helsingør, Denmark. ³Department of Civil and Environmental Engineering and Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

*To whom correspondence should be addressed. E-mail: falko@marine.rutgers.edu (P.G.F.); tfenchel@bio.ku.dk (T.F.); delong@mit.edu (E.F.D.)

organic matter (i.e., they are chemoautotrophs). Finally, in the absence of oxygen, a third set of opportunistic microbes uses NO_2^- and NO_3^- as electron acceptors in the anaerobic oxidation of organic matter. This respiratory pathway ultimately forms N_2 , thereby closing the N cycle. Hence, this cycle of coupled oxidation/reduction reactions, driven by different microbes that are often spatially or temporally separated, forms an interdependent electron pool that is influenced by photosynthetic production of oxygen and the availability of organic matter (17).

Are the niches for all possible redox reactions occupied by microbial metabolism? Although

some metabolic transformations, and the microbes that enable them, have been predicted to exist solely on the basis of thermodynamics, and only later were shown to actually occur (18, 19), not all predicted pathways have been found. Some, such as the oxidation of N_2 to NO_3^- , may be too kinetically constrained for biological systems. Similarly, no known photosynthetic organism can photochemically oxidize NH_4^+ .

Coevolution of the Metabolic Machines

Due to physiological and biochemical convenience, elemental cycles generally have been

studied in isolation; however, the cycles have co-evolved and influence the outcomes of each other. Metabolic pathways evolved to utilize available substrates produced as end products of other types of microbial metabolism, either by modification of existing metabolic pathways or by using established ones in reverse (20). Photosynthesis is another example of the evolution of multiple metabolic pathways that lead to a cycle. Typically, reduction and oxidation reactions are segregated in different organisms. In photosynthesis, the energy of light oxidizes an electron donor, i.e., H_2O in oxygenic photosynthesis and HS^- , H_2 , or Fe^{2+} in anoxygenic

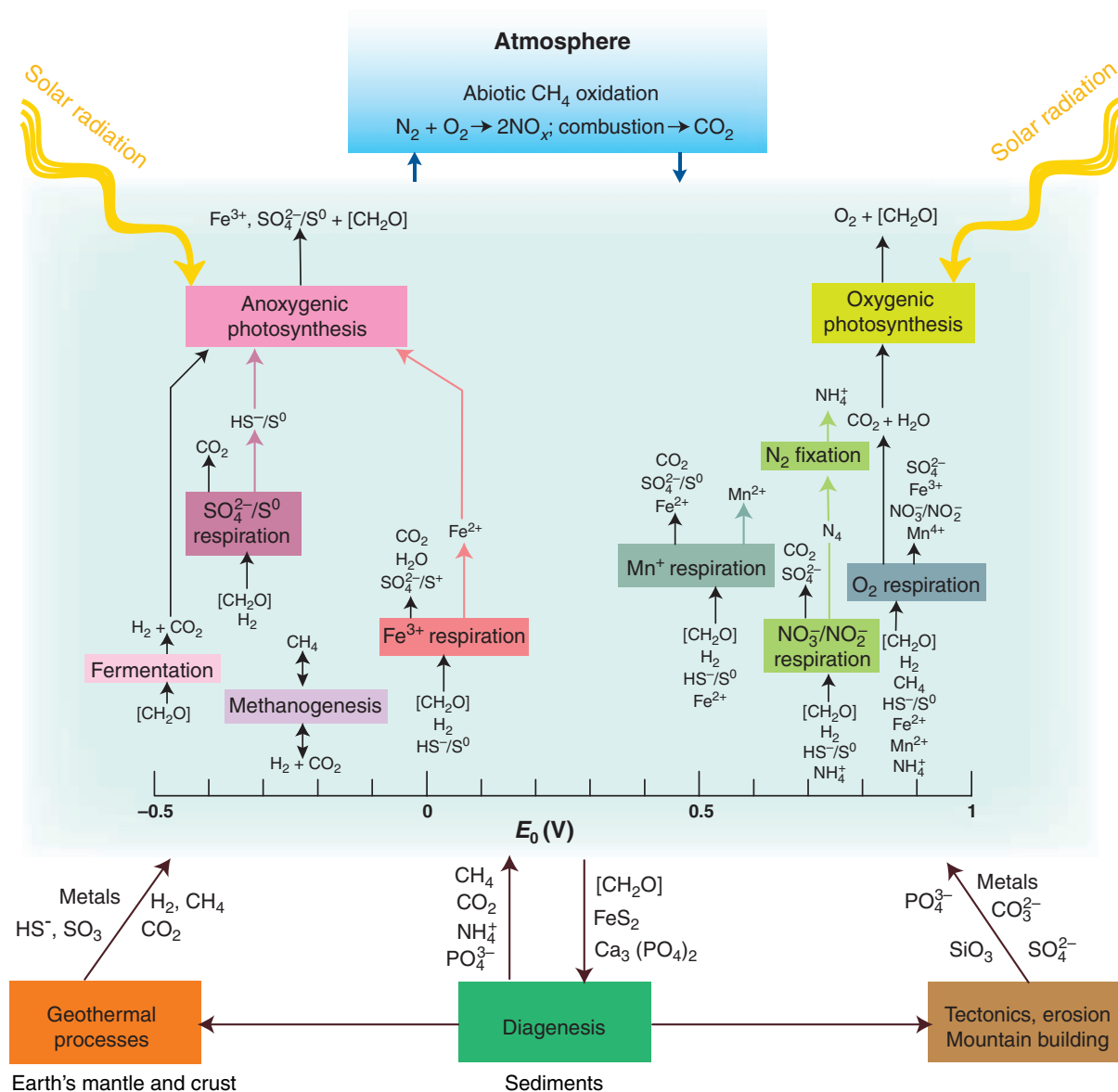


Fig. 1. A generalized biosphere model showing the basic inputs and outputs of energy and materials. Geochemical (abiotic) transformations are represented at the top (atmospheric) and bottom (tectonic and geothermal) compartments, while microbially driven biochemical processes are represented in the middle, biospheric compartment (in blue) and the sediments. Biological element cycling is not completely closed due to losses through sedimentation of organic carbon and nitrogen, carbonate, metal sulfides, sulfate, and phos-

phate, and losses to the atmosphere via denitrification. Regeneration of available forms of these elements is contingent on geological processes: erosion and geothermal activity. Electron acceptors (oxidants) in the respiratory processes have been arranged from left to right according to increasing capacity to accept electrons. The redox couples (at pH 7) for the reactions are approximate; the exact values depend upon how the individual reactions are coupled.

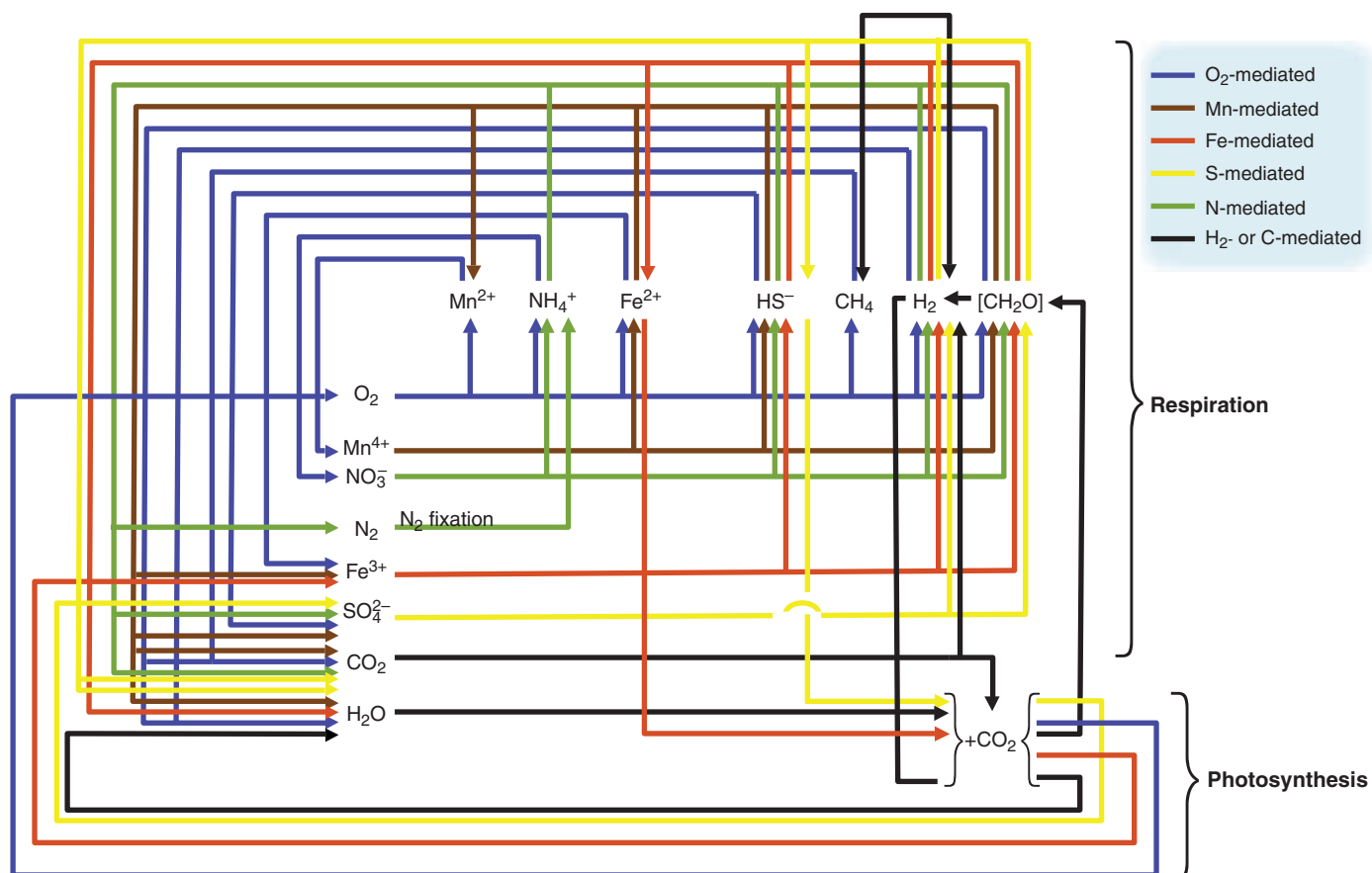


Fig. 2. A schematic diagram depicting a global, interconnected network of the biologically mediated cycles for hydrogen, carbon, nitrogen, oxygen, sulfur, and iron. A large portion of these microbially mediated processes are associated only with anaerobic habitats.

photosynthesis, and the electrons and protons generated in the process are used to reduce inorganic carbon to organic matter with the formation of higher-energy bonds. The resulting oxidized metabolites may in turn serve as electron acceptors in aerobic or anaerobic respiration for the photosynthetic organisms themselves or by other, nonphotosynthetic organisms that use these “waste products” as oxidants (21).

The outcome of the coupled metabolic pathways is that on geological time scales, the biosphere can rapidly approach relatively self-sustaining element cycling on time scales of centuries to millennia. On longer time scales, perpetuation of life remains contingent on geological processes and the constant flux of solar energy. Essential elements or compounds, such as phosphate, carbon (either as carbonate or organic matter), and metals, are continuously buried in sediments and are returned to the biosphere only through mountain building and subsequent erosion or geothermal activity (Fig. 1).

There is little understanding of how long it took for reaction cycles to develop from local events to global alteration of prevailing geologically produced redox set points. The last common ancestor of extant life presumably possessed genes for the adenosine triphosphatase complex required to maintain ion gradients generated by photochemical or respiratory processes. Regard-

less, one of the last metabolic pathways to emerge was oxygenic photosynthesis.

Oxygenic photosynthesis is the most complex energy transduction process in nature: More than 100 genes are involved in making several macromolecular complexes (22). Nevertheless, indirect evidence shows that this series of reactions had evolved by ~3 billion years ago (23), although the atmosphere and the upper ocean maintained a very low concentration of O_2 for the next ~0.5 billion years (24, 25). The production and respiration of nitrate must have evolved after the advent of oxygenic photosynthesis, as there can be no nitrate without oxygen (16). Although the succession of probable events that led to the global production of O_2 is becoming increasingly clear (26, 27), the evolutionary details delimiting important events for other redox cycles and elements are more ambiguous.

Attempts to reconstruct the evolution of major dissimilatory metabolic pathways are mainly based on geological evidence for the availability of potential electron donors and oxidants during the early Precambrian (23). Although we can gain some idea of the relative quantitative importance of different types of energy metabolism, we do not know the order in which they evolved. Indeed, the origin of life and the first reactions in energy metabolism probably never will be known with cer-

tainty. These events took place before any geological evidence of life, and while phylogenetic trees and structural analyses provide clues regarding key motifs, so far they have not provided a blueprint for how life began. Stable-isotope fractionation has provided evidence for sulfate reduction and methanogenesis in 3.5-billion-year-old deposits (28), but these metabolic processes are presumably older.

Modes of Evolution

Molecular evidence, based on gene order and the distribution of metabolic processes, strongly suggests that early cellular evolution was probably communal, with promiscuous horizontal gene flow probably representing the principal mode of evolution (29). The distribution of genes responsible for the major extant catabolic and anabolic processes may have been distributed across a common global gene pool, before cellular differentiation and vertical genetic transmission evolved as we know it today. In the microbial world, not only individual genes but also entire metabolic pathways central to specific biogeochemical cycles appear to be frequently horizontally transferred; a contemporary analog is the rapid acquisition of antibiotic resistance in pathogenic bacteria (30). The dissimilatory sulfite reductases found in contemporary sulfate-reducing δ -proteobacteria, Gram-positive

bacteria, and Archaea are examples of horizontal gene transfer that reflect the lateral propagation of sulfate respiration among different microbial groups and environments (31). Indeed, with the exception of chlorophyll- or bacteriochlorophyll-based photosynthesis, which is restricted to Bacteria, and methanogenesis, which is restricted to representatives within the Archaea (32), individual bacterial and Archaeal lineages contain most major metabolic pathways. Even some of the molecular components of methanogens seem to have been laterally transferred to methane-oxidizing members of the domain Bacteria (33). Nitrogenases appear to have been transferred to oxygenic photosynthetic cyanobacteria late in their evolutionary history, probably from an Archaean source (34), and are widespread among diverse groups of Bacteria and Archaea (35). Ammonia monooxygenase genes that encode the key enzyme required for the oxidation of ammonia to hydroxylamine, a key step of the nitrogen cycle, are also widely distributed (36, 37). Evidence also exists for lateral exchange of large “superoperons” encoding the entire anoxygenic photosynthetic apparatus (38). Presumably, severe nutritional or bioenergetic selective pressures serve as major drivers for the retention of horizontally transferred genes, thereby facilitating the radiation of diverse biogeochemical reactions among different organisms and environmental contexts.

Sequence Space Available

Although the absolute number of genes and protein families currently in existence is unknown, several approaches have been used to evaluate the relative depth of protein “sequence space” currently sampled. Microbial community genome sequencing (i.e., metagenomics) provides a cultivation-independent, and hence potentially less biased, view of extant sequence space. The number of protein families within individual Bacterial and Archaeal genomes depends linearly on the number of genes per genome, and hence genome size (39). The higher levels of gene duplication found in nonmicrobial eukaryotic genomes potentially allows them to escape this constraint and has resulted in different evolutionary strategies and genome organization (39). Regardless, genome size appears to be correlated with evolutionary rate, but not with core metabolic processes (40). So, what does the apparent diversity in microbial genomes signify?

Genome Diversity in Nature

To date, the rate of discovery of unique protein families has been proportional to the sampling effort, with the number of new protein families increasing approximately linearly with the number of new genomes sequenced (41). The size of protein families (the number of nonredundant proteins found within a family) among fully sequenced genomes follows a power law, with the greatest number of protein families containing only a few members (39). These trends among

fully sequenced genomes are also mirrored in large-scale metagenomic shotgun sequencing efforts (42). Among the ~6 million newly predicted protein sequences from a recent ocean metagenomic survey, a total of 1700 new protein families were discovered with no homologs in established sequence databases. Even though this study increased the known number of protein sequences nearly threefold from just one specific habitat, the discovery rate for new protein families was still linear (Fig. 3). These data indicate that we have only just begun the journey of cataloguing extant protein sequence space.

The virtual explosion of genomic information has led to the hypothesis that there is limitless evolutionary diversity in nature. The vast majority of unexplored sequence space appears to encompass two categories of genes: a large and dynamic set of nonessential genes and pseudogenes, under neutral or slightly negative selective pressure (which we call “carry-on genes”), and a set of positively selected environment-specific gene suites, tuned to very particular habitats and organisms (which we call “boutique genes”). In contrast, the evolution of most of the essential multimeric microbial machines (including the basic energy transduction processes, nitrogen metabolic processes, ribosomes, nucleic acid replication en-

zymes, and other multienzyme complexes) is highly constrained by intra- and internucleic acid, RNA-protein, protein-protein, protein-lipid, and protein-prosthetic group interactions (22), to the extent that even when the machines function suboptimally, they are retained with very few changes. For example, the D1 protein in the reaction center of Photosystem II, a core protein in the water-splitting reaction center found in all oxygenic photosynthetic organisms, is derived from an anaerobic purple bacterial homolog. During oxygenic photosynthesis, this protein is degraded by photooxidative cleavage approximately every 30 min (43). Rather than reengineer the reaction center to develop a more robust protein in the machine, a complicated repair cycle has evolved that removes and replaces the protein. Consequently, photosynthetic efficiency, especially at high irradiance levels, is not as high as theoretically possible (44), yet the D1 is one of the most conserved proteins in oxygenic photosynthesis (22). Similarly, nitrogenase is irreversibly inhibited by molecular oxygen, yet this core machine is also very highly conserved even though many nitrogen-fixing organisms live in an aerobic environment. To compensate, nitrogen-fixing organisms have had to develop mechanisms for protecting this enzyme from oxygen by spatially or tempo-

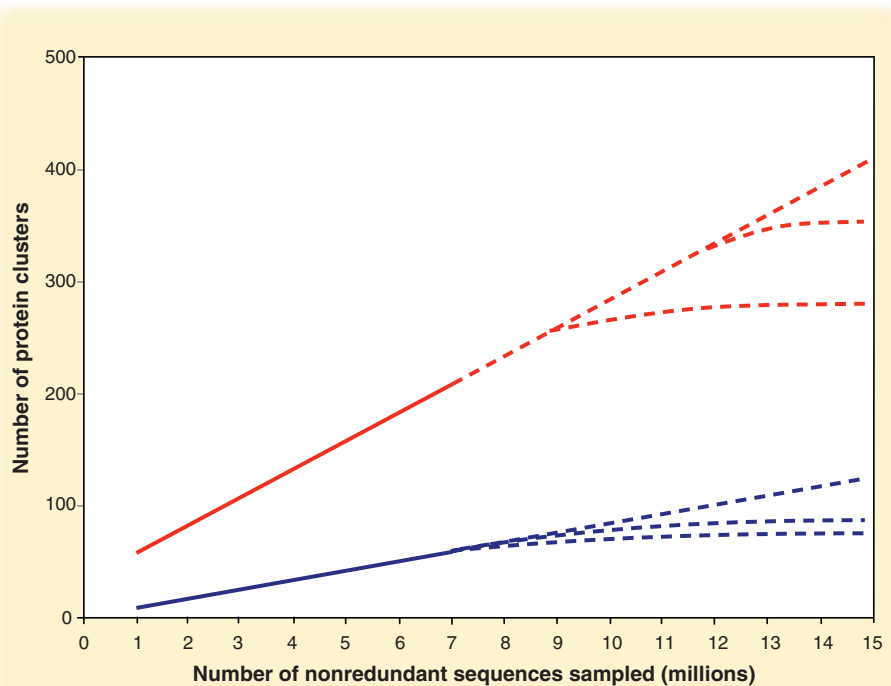


Fig. 3. Observed increases in new protein clusters with increasing sequence sampling [modified from Yooshef *et al.* (42)]. The number of new protein clusters discovered increases linearly with the number of nonredundant sequences sampled. We project hypothetical saturation profiles for the protein families. However, discovery of new protein families is much lower in protein clusters with greater membership. Seven different data sets of various sizes, including curated public databases and new data described in Yooshef *et al.* (42), were used to generate seven differently sized, nonredundant sequence data samples depicted. The red line shows protein clusters with ≥ 3 core sets of highly related sequences in a given cluster. The blue line shows protein clusters with ≥ 10 core sets of highly related sequences in a given cluster.

rally segregating nitrogen fixation from aerobic environments (45–47). In the contemporary ocean, ~30% of nitrogenase is nonfunctional at any moment in time, forcing overproduction of the protein complex to facilitate nitrogen fixation.

Is Everything Everywhere?

Abundant evidence exists for the rapid and efficient dispersal of viral particles and microbial cells, and for the genes they carry. At the same time, both microbial isolations and environmental genomic surveys indicate environmentally specific, quantitative distributional patterns of iron oxidation, methane metabolism, and photosynthesis (11, 48, 49). These distributions generally, but not always, reflect the environmental distributions of specific taxonomic groups. For example, the simplicity and modularity of rhodopsin-based photosynthesis appear to have led to the dispersal of this pathway into widely disparate taxonomic groups. The environmental distribution of these photoproteins therefore appears more reflective of habitat selective pressure than of any specific organismal or taxonomic distribution (50). Although the distributions of specific taxa may not vary greatly along a particular environmental gradient, in the absence of the relevant selection pressure, environmentally irrelevant genes may be lost rapidly (51).

The generalization that particular kinds of microbes always occur whenever their habitat requirements are realized is far from new (52). Although not necessarily metabolically active, viable bacteria of a particular functional type can be recovered from almost any environment, using appropriate types of enrichment cultures anywhere, even where that environment cannot support their growth. Hence, thermophilic bacteria can be grown from cold seawater (53), strict anaerobes from aerobic habitats (54), and microbial cells have been observed to accumulate in high numbers in surface snow at the South Pole (55). These observations can be explained by the sheer number of microbial cells occurring on Earth and consequent high efficiencies of dispersal and low probabilities of local extinction. Evidence for this also appears to be reflected in the vast number of very rare sequences revealed in rarefaction curves of deep microbial sampling surveys (56), which perhaps represents a sort of “biological detritus” from the very efficient microbial dispersal, coupled with extremely slow decay kinetics of individual microbial cells or spores in various resting states.

Very early in life’s history the atmosphere and oceans were anoxic and the distribution of the first aerobic respiring microbes was confined to the close vicinity of cyanobacteria. By contrast, in the extant surface biosphere, aerobic conditions are very widespread. During the late Proterozoic (between ~750 and 570 million years ago) glaciations, large parts of Earth’s surface may have been covered by ice, but even small remaining habitat patches will have assured the persistence and eventually dissemination of all types of prokary-

otes. By extension, it is unlikely that mass-extinction episodes in the Phanerozoic (the past 545 million years), which strongly influenced the evolution of animals and plants, did not fundamentally influence the core metabolic machines. How then has the ancient core planetary metabolic gene set been maintained over the vast span of evolutionary time?

Microbes as Guardians of Metabolism

Dispersal of the core planetary gene set, whether by vertical or horizontal gene transfer, has allowed a wide variety of organisms to simultaneously, but temporarily, become guardians of metabolism. In that role, environmental selection on the microbial phenotype leads to evolution of the boutique genes that ultimately protect the metabolic pathway. If the pathway in a specific operational taxonomic unit does not survive an environmental perturbation, the unit will go extinct, but the metabolic pathway has a strong chance of survival in other units. Hence, the same selective pressures enabling retention of fundamental redox processes have persisted throughout Earth’s history, sometimes globally, and at other times only in refugia, but able to emerge and exert ubiquitous selection pressure on ancillary genes. In essence, microbes can be viewed as vessels that ferry metabolic machines through strong environmental perturbations into vast stretches of relatively mundane geological landscapes. The individual taxonomic units evolve and go extinct, yet the core machines survive surprisingly unperturbed.

Humans may not yet be able to mimic the individual redox reactions that drive planetary processes; nevertheless, the interconnections between biogeochemical processes and the evolution of biologically catalyzed reactions are becoming more tractable for measurement and modeling. It is likely that the individual reactions that make life possible on Earth will be reasonably well described within the next few decades. Delineating how these machines coevolved and operate together to create the electron flows that predominate today on Earth’s surface remains a grand challenge. Understanding biogeochemical coevolution is critical to the survival of humans as we continue to influence the fluxes of matter and energy on a global scale. Microbial life can easily live without us; we, however, cannot survive without the global catalysis and environmental transformations it provides.

References and Notes

1. V. Vernadsky, *Geochemistry and the Biosphere* (Synergetic Press, Santa Fe, 2007).
2. A. H. Knoll, *Life on a Young Planet: The First Three Billion Years of Evolution on Earth* (Princeton Univ. Press, Princeton NJ, 2003).
3. A. Hallam, P. Wignall, *Mass Extinctions and Their Aftermath* (Oxford Univ. Press, Oxford, 1997).
4. www.pdb.org/
5. A. Kluver, H. Donker, *Chem. Zelle Gewebe* **13**, 134 (1926).
6. R. Williams, J. Frausto da Silva, *The Natural Selection of the Chemical Elements* (Clarendon, Oxford, 1996).

7. D. C. Catling, M. W. Claire, *Earth Planet. Sci. Lett.* **237**, 1 (2005).
8. W. H. Schlesinger, *Biogeochemistry: An Analysis of Global Change* (Academic Press, New York, ed. 2, 1997).
9. P. G. Falkowski, L. Godfrey, *Proc. R. Soc. London B Biol. Sci.*, 10.1098/rstb.2008.0054 (2008).
10. E. Milner-White, M. J. Russell, *Biol. Direct* **3**, (2008).
11. S. J. Hallam *et al.*, *Science* **305**, 1457 (2004).
12. V. J. Orphan, C. H. House, K. U. Hinrichs, K. D. McKeegan, E. F. DeLong, *Science* **293**, 484 (2001).
13. B. Schink, *Antonie Van Leeuwenhoek* **81**, 257 (2002).
14. G. Wächtershäuser, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 200 (1990).
15. J. R. Postgate, *Nitrogen Fixation* (Cambridge Univ. Press, Cambridge, ed. 3, 1998).
16. P. Falkowski, *Nature* **387**, 272 (1997).
17. K. Fennel, M. Follows, P. G. Falkowski, *Am. J. Sci.* **305**, 526 (2005).
18. E. Broda, *Z. Allg. Mikrobiol.* **17**, 491 (1977).
19. M. Jettten *et al.*, *FEMS Microbiol. Rev.* **22**, 421 (1998).
20. J. A. Gerlt, P. C. Babbitt, *Annu. Rev. Biochem.* **70**, 209 (2001).
21. R. E. Blankenship, M. T. Madigan, C. E. Bauer, Eds., *Anoxygenic Photosynthetic Bacteria* (Kluwer Scientific, Dordrecht, Netherlands, 1995).
22. T. Shi, T. S. Bibby, L. Jiang, A. J. Irwin, P. G. Falkowski, *Mol. Biol. Evol.* **22**, 2179 (2005).
23. D. E. Canfield, *Annu. Rev. Earth Planet. Sci.* **33**, 1 (2005).
24. J. Farquhar, H. M. Bao, M. Thiemens, *Science* **289**, 756 (2000).
25. A. Bekker *et al.*, *Nature* **427**, 117 (2004).
26. R. Blankenship, S. Sadekar, J. Raymond, in *Evolution of Primary Producers in the Sea*, P. G. Falkowski, A. H. Knoll, Eds. (Academic Press, San Diego, 2007), pp. 21–35.
27. J. Raymond, *Rev. Mineral. Geochem.* **59**, 211 (2005).
28. D. Canfield, R. Raiswell, *Am. J. Sci.* **299**, 697 (1999).
29. C. R. Woese, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 8742 (2002).
30. J. P. Gogarten, W. F. Doolittle, J. G. Lawrence, *Mol. Biol. Evol.* **19**, 2226 (2002).
31. M. Klein *et al.*, *J. Bacteriol.* **183**, 6028 (2001).
32. C. R. Woese, *Microbiol. Rev.* **51**, 221 (1987).
33. L. Chistoserdova, J. A. Vorholt, R. K. Thauer, M. E. Lidstrom, *Science* **281**, 99 (1998).
34. T. Shi, P. Falkowski, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 2510 (2008).
35. K. J. Kechris, J. C. Lin, P. J. Bickel, A. N. Glazer, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 9584 (2006).
36. S. J. Hallam *et al.*, *PLoS Biol.* **4**, e95 (2006).
37. M. Konneke *et al.*, *Nature* **437**, 543 (2005).
38. K. V. P. Nagashima, A. Hiraishi, K. Shimada, K. Matsuura, *J. Mol. Evol.* **45**, 131 (1997).
39. A. J. Enright, V. Kunin, C. A. Ouzounis, *Nucleic Acids Res.* **31**, 4632 (2003).
40. M. Oliver, D. Petrov, D. Ackerley, P. G. Falkowski, O. Schofield, *J. Genome Res.* (2007).
41. V. Kunin, I. Cases, A. J. Enright, V. de Lorenzo, C. A. Ouzounis, *Genome Biol.* **4**, 401 (2003).
42. S. Yooseph *et al.*, *PLoS Biol.* **5**, e16 (2007).
43. I. Ohad, D. J. Kyle, C. J. Arntzen, *J. Cell Biol.* **99**, 481 (1984).
44. S. P. Long, S. Humpries, P. G. Falkowski, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **45**, 633 (1994).
45. P. W. Ludden, G. P. Roberts, in *Anoxygenic Photosynthetic Bacteria*, R. E. Blankenship, M. T. Madigan, C. E. Bauer, Eds. (Kluwer, Dordrecht, Netherlands, 1995), pp. 929–947.
46. I. Berman-Frank *et al.*, *Science* **294**, 1534 (2001).
47. R. C. Burns, R. W. F. Hardy, *Nitrogen Fixation in Bacteria and Higher Plants* (Springer, New York, 1975).
48. E. F. DeLong *et al.*, *Science* **311**, 496 (2006).
49. G. W. Tyson *et al.*, *Nature* **428**, 37 (2004).
50. N. U. Frigaard, A. Martinez, T. J. Mincer, E. F. DeLong, *Nature* **439**, 847 (2006).
51. E. Lerat, V. Daubin, H. Ochman, N. A. Moran, *PLoS Biol.* **3**, e130 (2005).
52. L. G. M. Baas Becking, *Geobiologie of Inleiding tot de Milieukunde* (Van Stockum and Zoon, The Hague, Netherlands, 1934).
53. M. F. Isaksen, F. Bak, B. B. Jørgensen, *FEMS Microbiol. Ecol.* **14**, 1 (1994).
54. M. Bianchi, J.-L. Teysseie, B. B. Fowler, *Mar. Ecol. Prog. Ser.* **88**, 55 (1992).

55. E. J. Carpenter, S. Lin, D. G. Capone, *Appl. Environ. Microbiol.* **66**, 4514 (2000).
 56. M. L. Sogin *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 12115 (2006).

57. Supported by NASA and the Agouron Foundation (P.G.F.); the Danish Natural Science Research Council and The Carlsberg Foundation (T.F.); and the Gordon and Betty Moore Foundation, the

Agouron Foundation, NSF, and U.S. Department of Energy (E.F.D.).

10.1126/science.1153213

REVIEW

Microbial Biogeography: From Taxonomy to Traits

Jessica L. Green,^{1*} Brendan J. M. Bohannan,¹ Rachel J. Whitaker²

The biogeographic variation of life has predominantly been studied using taxonomy, but this focus is changing. There is a resurging interest in understanding patterns in the distribution not only of taxa but also of the traits those taxa possess. Patterns of trait variation shed light on fundamental questions in biology, including why organisms live where they do and how they will respond to environmental change. Technological advances such as environmental genomics place microbial ecology in a unique position to move trait-based biogeography forward. We anticipate that as trait-based biogeography continues to evolve, micro- and macroorganisms will be studied in concert, establishing a science that is informed by and relevant to all domains of life.

And so it was indeed: she was now only ten inches high, and her face brightened up at the thought that she was now the right size for going through the little door into that lovely garden.

Lewis Carroll (1865)

Imagine Carl Linnaeus in Alice's shoes, shrinking to only 10 micrometers high. Afforded the opportunity to investigate biological diversity at this spatial scale, would Linnaeus have remained committed to plant exploration, or would he have turned his attention to microbial life? It is not surprising that Linnaeus and his contemporaries founded biogeography—a science that aims to document and understand spatial patterns of biological diversity—by studying organisms visible to the naked eye. Recent advances in our ability to quantify and visualize microbial diversity in natural environments have prompted a new era of microbial exploration, one that builds upon plant and animal biogeography surveys initiated roughly 250 years ago (1). These new explorations have already radically changed thinking in ecology and evolution and upset the hierarchical taxonomic structure that Linnaeus proposed (2). As microbiologists increasingly focus on biogeographical questions, textbook placeholders such as “microbial biogeography is poorly known and rarely discussed” (3) will become obsolete.

A long-held concept in microbial ecology is that microorganisms are dispersed globally and able to proliferate in any habitat with suitable envi-

ronmental conditions. This concept was prompted by Martinus Willem Beijerinck and concisely summarized by Lourens Gerhard Marinus Baas Becking in the widely referenced quote, “everything is everywhere, but the environment selects” (4). Seminal notions of ubiquitous dispersal and environmental determinism are not unique to microbiology. Linnaeus, for example, wrote that “the great Artificer of Nature has provided that every seed shall find its proper soil, and be equally dispersed over the surface of the globe” (1). The development of molecular approaches has allowed a more comprehensive view of microbial diversity than can be developed even with the aided eye, showing that like plant and animal distributions, microbial distributions can be the result of both deterministic (environmental) and stochastic (dispersal) processes [reviewed in (5)].

As with macroorganism biogeography, microbial biogeography initially adopted a taxonomic approach, focusing on sequence signatures to identify groups of microorganisms. These studies revealed classic patterns such as the species-area relationship and isolation by distance (5). Interpreting taxonomic patterns in terms of how they affect the function of a population or community is especially difficult in microorganisms, where a broad range of functional variation may occur among similar organisms (e.g., organisms with the same 16S rRNA sequence).

For macro- and microorganisms alike, there is growing interest in the biogeography of functional traits, characteristics of an organism that are linked with its fitness or performance (6). The study of biogeography includes the study of patterns in space, in time, and along environmental gradients. Such patterns in the distribution of traits can be used to understand complex phenomena, includ-

ing why organisms live where they do, how many taxa can coexist in a place, and how they will respond to environmental change. Although plants have been the focal group in this emerging research area, recent advances in environmental molecular biology such as genomics, proteomics, transcriptomics, and metabolomics place microbial ecology in a unique position to move trait-based biogeography forward.

Trait-Based Biogeography: A Macroorganism Perspective

Trait-based approaches to biogeography have been used since the pioneering work of Andreas F. W. Schimper more than a century ago (7). Although tending to wax and wane in favor over time, there has been a resurgence of interest in trait-based methodologies since the mid-1980s (6, 8). Here, we discuss some examples of plant trait-based research, focusing on applications likely relevant to both plant and microbial ecology.

An emergent theme in trait-based research is the identification of ecological strategies, suites of covarying ecological traits. The study of ecological strategies has been fundamental to the development of plant and animal ecology, and there is growing interest among microbial ecologists as well (9). Examining the slope, intercept, and correlation strength of relationships among traits provides insight into the nature of ecological strategies, including the underlying costs and benefits of different trait combinations. Quantification of trait variation with site properties such as climate, for example, is central to understanding how vegetation properties shift along geographical gradients, and thus for predicting habitat boundaries under changing land-use and warming scenarios. A similar approach has been suggested for microorganisms (9) and could be useful for predicting how microbial properties respond to environmental change.

Figure 1A illustrates how the relationship between two ecologically important plant traits—leaf life span and leaf mass per area—shifts with climate (10). This trait relationship is part of a fundamental ecological strategy known as the “leaf economics spectrum” that ranges from organisms with cheaply constructed, thin, and short-lived leaves to those with costly, thick, and long-lived leaves. The relationship is modulated by site climate, such that organisms at drier sites typically achieve shorter leaf life spans at a given leaf mass per area.

Community “assembly rules” was a concept formulated to understand why organisms live where they do and in what combinations. It has been used to understand how plant and animal communities change through time and to predict

¹Center for Ecology and Evolutionary Biology, Department of Biology, University of Oregon, Eugene, Oregon 97403, USA. ²Department of Microbiology, University of Illinois Urbana-Champaign, Urbana, Illinois 61801, USA

*To whom correspondence should be addressed. E-mail: jlgreen@uoregon.edu