



Middlebury
College

Erin Eggleston, PhD

Assistant Professor of Biology

Department of Biology

eeggleston@middlebury.edu

276 Bicentennial Way, MBH 309A

Middlebury, VT 05753, USA

802-443-5779

July 2021

Dear Alumni College students,

Welcome to Microbial Ecology of Lake Champlain! I am excited to explore the amazing world of microbes with you. Our time together will focus on some of the core concepts pertaining to the microbes of Lake Champlain and freshwater ecosystems. You will become a microbial ecologist as we discuss a breadth of topics and gain hands-on experience with fieldwork and research tools.

We will investigate how tiny organisms impact the environment on local and global scales. We will work through the following central questions: How can organisms that are so small impact global nutrient cycling? What roles do microbes play in freshwater ecosystems? How do these microbial communities change in response to environmental factors, and what does this tell us about the health status of our local waterways?

Before you arrive, there are a few course materials with which I would like you to engage. These include a few readings, a short video, and some field observations. To provide context for our discussions of microbes in Lake Champlain we will ground ourselves with some background information about Lake Champlain ("Lake Champlain Basin Program: The Basin") and a video from Chief Don Stevens sharing the Abenaki *Bitawbágw* (Lake Champlain) creation story. Additionally, you should read two articles: Whitman *et al.* (1998) "Prokaryotes: The unseen majority," and Falkowski *et al.* (2008) "The Microbial Engines That Drive Earth's Biogeochemical Cycles." Finally, I ask that you complete a series of "Microbes in Motion" observations in which you sit, observe, and take notes about microbes in the environment around you. Additional information for each of these pre-course components and the course readings follow this letter.

Looking forward to meeting you all in August!

All the best,

Erin M Eggleston, PhD

Preparations for “Microbial Ecology of Lake Champlain”

The reading materials follow these guidelines. If you prefer that printed materials be mailed to you please contact the Alumni Office at alumni@middlebury.edu.

Introduction to Lake Champlain:

Read “Lake Champlain Basin Program: The Basin” and watch the ~10-minute film [Nebi](#). You can watch the video by clicking on this link: <https://www.lakechamplaincommittee.org/learn/news/item/nebi-abenaki-ways-of-knowing-water>

After engaging with these materials, reflect on the following questions and jot down some notes to bring with you to class:

- How might present-day variations in land use in Vermont and New York impact the nutrients in the five basins of Lake Champlain? What environmental factors (e.g., nutrients, temperature, etc.) might be most important in these waters?
- What did you learn from the *Nebi* video? How might the origin story of *Odzihozo* inform our thinking of long-term changes to the Lake Champlain region? What has changed since Lake Champlain formed, and how does this impact the kind of organisms that thrive in present-day Lake Champlain?

Microbes and Biogeochemistry:

We will spend a great deal of time thinking about microbes. When we discuss microbes we include bacteria, archaea, fungi, and viruses, most of which are too small to see without the aid of a microscope. Biogeochemistry, which links biology, geology, and chemistry, is also a key theme. There is a deep relationship between microbes, evolution, and these nutrient cycling processes. These two review articles are quite dense but provide rich context for the role of microbes in the environment. Take a stab at unfamiliar terms with quick Google searches, and bring any questions to class! Focal points are noted for each paper:

- Whitman *et al.* 1998 “Prokaryotes: The unseen majority.” You may read the whole article, but we will primarily discuss the “Aquatic Environments” subsection and Table 1. These numbers are hard to put into perspective. But I encourage you to grapple a bit with trying to contextualize these numbers and the implications of the magnitude of these microbes and their chemical composition in the environment. Be critical as you read. What assumptions are made with these calculations?
- Falkowski *et al.* 2008 “The Microbial Engines That Drive Earth’s Biogeochemical Cycles.” Our discussions will draw from the subsections “The Major Biogeochemical Fluxes Mediated by Life” and “Coevolution of the Metabolic Machines,” but feel free to read the full article. There is no expectation for you to understand all the details of the information presented here. Instead, make an

effort to identify the big-picture take-home message(s), and be prepared to engage in further discussion of these nutrient cycles.

Microbes in Motion:

Please aim to complete three of these observation sessions prior to your arrival. Feel free to do more! You will need ~20 minutes to complete each session. Bring some paper and your favorite writing implement. Find a comfortable place to sit in a natural setting (e.g., near a river or lake, in a park or forest, etc.). I recommend silencing any devices and trying to find a place away from others to deeply connect to the location. For each site:

- For a few minutes just sit, uninterrupted, and take in your location. After you feel grounded in your location, jot down some initial observations. Where are you? What did you notice? Are any “macrobes” passing by (e.g., birds, small mammals, etc.)?
- Next, think about the physical and chemical factors of this location. What is the temperature? When did it last rain here? Are there obvious nutrient sources? If not, from where might microbes gain their key nutrients in this environment (e.g., sugars, proteins, etc.)? Are these conditions constant in this location, or do they vary? What else do you notice that might allow or prohibit the growth of organisms, in particular microbes, in this location? What is the relative impact of humans at this location? Make some notes as you think through these physicochemical factors.
- Finally, turn your thoughts to the microbes in this landscape. Given the readings from Whitman *et al.* 1998 and Falkowski *et al.* 2008, in what ways might these invisibly small organisms be actively transforming this location? Is there any evidence of microbial activity here? Perhaps there are biofilms or symbiotic associations (e.g., lichens) that are conspicuous? Conclude with some final notes about your microbial ruminations.

After you have completed your third (or final) Microbes in Motion observation session, review your notes from different sites. What were the similarities and differences? How did these reflections shift your perspective of these locations, if at all? Feel free to take note of anything unexpected or exciting that occurred. If you are unable to visit sites in person, you are welcome to use any nature documentaries or other videos or photographs as inspiration for these observations. Bring your observation notes with you to class.

Optional resources:

Lake Champlain Natural History:

<https://www.lakechamplaincommittee.org/learn/natural-history-lake-champlain>

Lake Champlain Basin Atlas: <https://atlas.lcbp.org/>

LAKE CHAMPLAIN BASIN PROGRAM

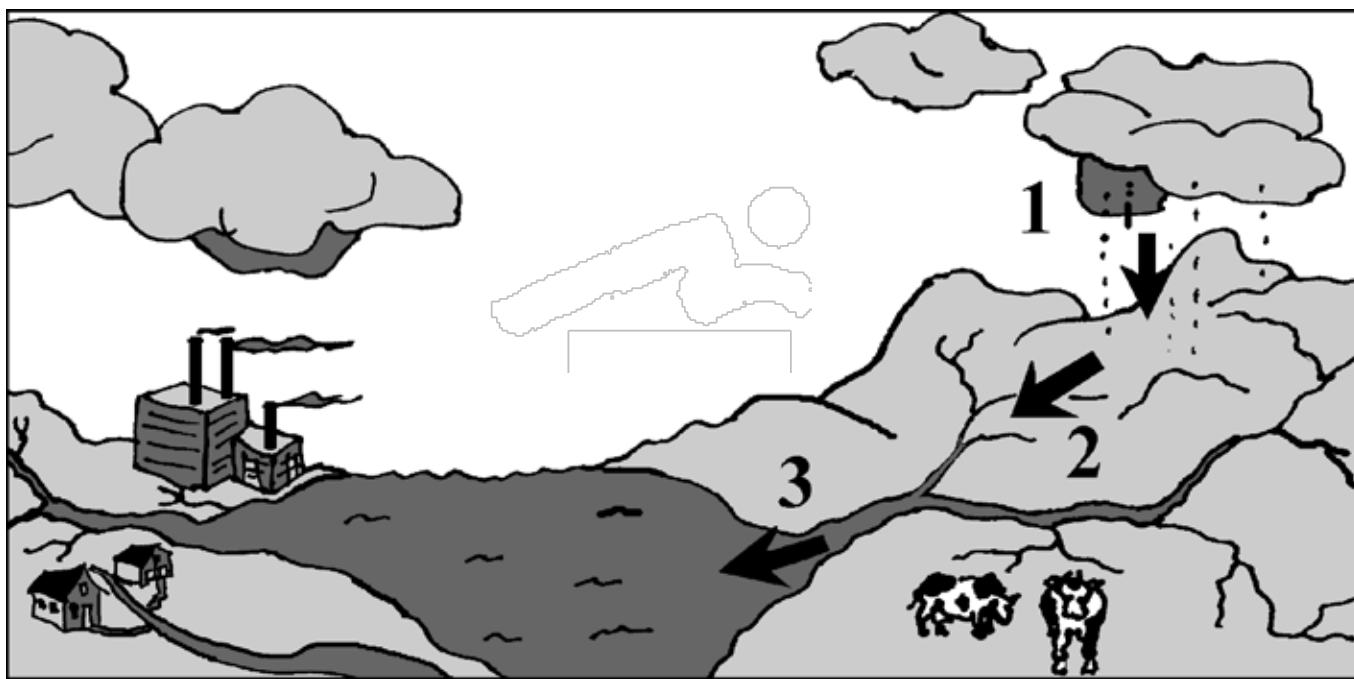
Fact Sheet Series

Number 3



T H E B A S I N

The Lake Champlain Basin is the entire watershed or drainage area for Lake Champlain. It spans the High Peaks of New York's Adirondack Mountains in the West, Vermont's Green Mountains in the East, and Quebec's St. Lawrence Valley in the North. Looking at the Lake from a watershed perspective is important since the water quality of the Lake is affected by land and water uses from the mountain tops to the Lake's shoreline.

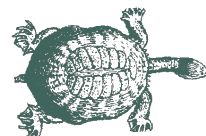


How Basins Work

Drainage basins or watersheds are like giant funnels. The land within Lake Champlain's basin collects precipitation in the form of rain, snow and sleet (1). This precipitation flows into wetlands, groundwater, tributaries, and streams (2), and eventually into the Lake (3). Unfortunately, this water also carries pollutants from the land. Some of the pollution sources include cities, farms, factories, houses, and roads. The Basin's boundary is determined by connecting the points of highest elevation around the Lake on a topographical map.

Formation of the Basin

The Green Mountains formed about 450 million years ago when the North American and European tectonic plates collided. As great blocks of land between the Green Mountains and the ancient Adirondack Mountains dropped down, the Champlain Valley was formed. Over time, the shape of this valley changed as glaciers plowed over the land, resulting in the U-shaped valley characteristic of New England.

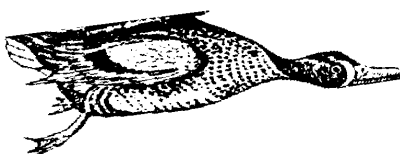


History and Development

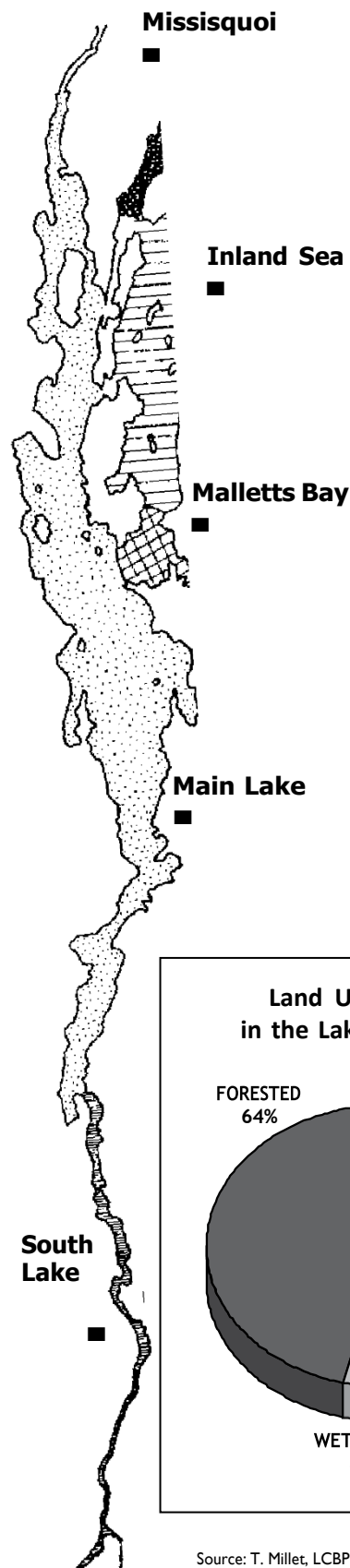
The first humans occupied the Lake Champlain Basin soon after the glaciers melted over 10,000 years ago. These early Native Americans hunted and fished and later became skilled at horticulture. They adapted to their environment without polluting, destroying, or depleting its natural resources. They left few traces behind making it difficult to precisely date their settlements.

Following Samuel de Champlain's exploration in 1609, which marked the beginning of European settlement in the Basin, the Lake became known as Lake Champlain. During this period, the Lake served as a boundary between the Abenakis along the Vermont shores and the Iroquois along the New York shoreline.

Following Champlain's arrival, a long history of military battles and power struggles began, including the French and Indian War, the American Revolution and the War of 1812. A rural economy, focusing largely on agriculture, was established in the Champlain Valley in the 1700s. The economy soon expanded to include natural resources such as timber, fish, ice, maple syrup, iron ore, and marble. Vacationing around Lake Champlain became very popular beginning soon after the Civil War. Throughout history, the demand for transportation over land, water, and ice increased, making the boat building industry and railroads very important to the Lake Champlain region during the 1800s and early 1900s.



The Water and The Land



Lake Champlain is made up of five distinct areas, each with different physical and chemical characteristics.

South Lake This segment is narrow and shallow, much like a river.

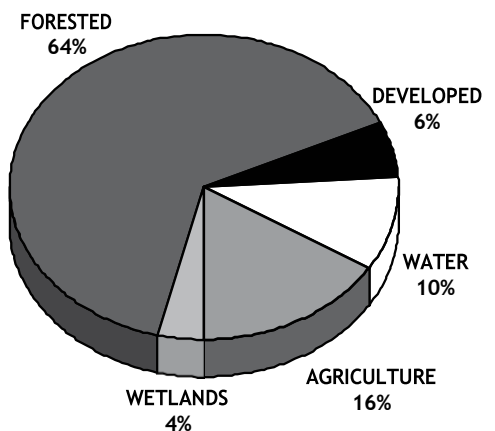
Main Lake This segment contains about 81% of the volume of the entire Lake, including the deepest, coldest water.

Malletts Bay This area lies between causeways built to the north and west. It has the most restricted circulation of any of the Lake's segments.

Inland Sea (Northeast Arm) The water here generally flows south from Missisquoi, north from Malletts Bay and passes through and around the Champlain Islands.

Missisquoi Bay Most of this segment lies within Canada. The Bay is very shallow and relatively warm. Water from the Bay flows into the Inland Sea.

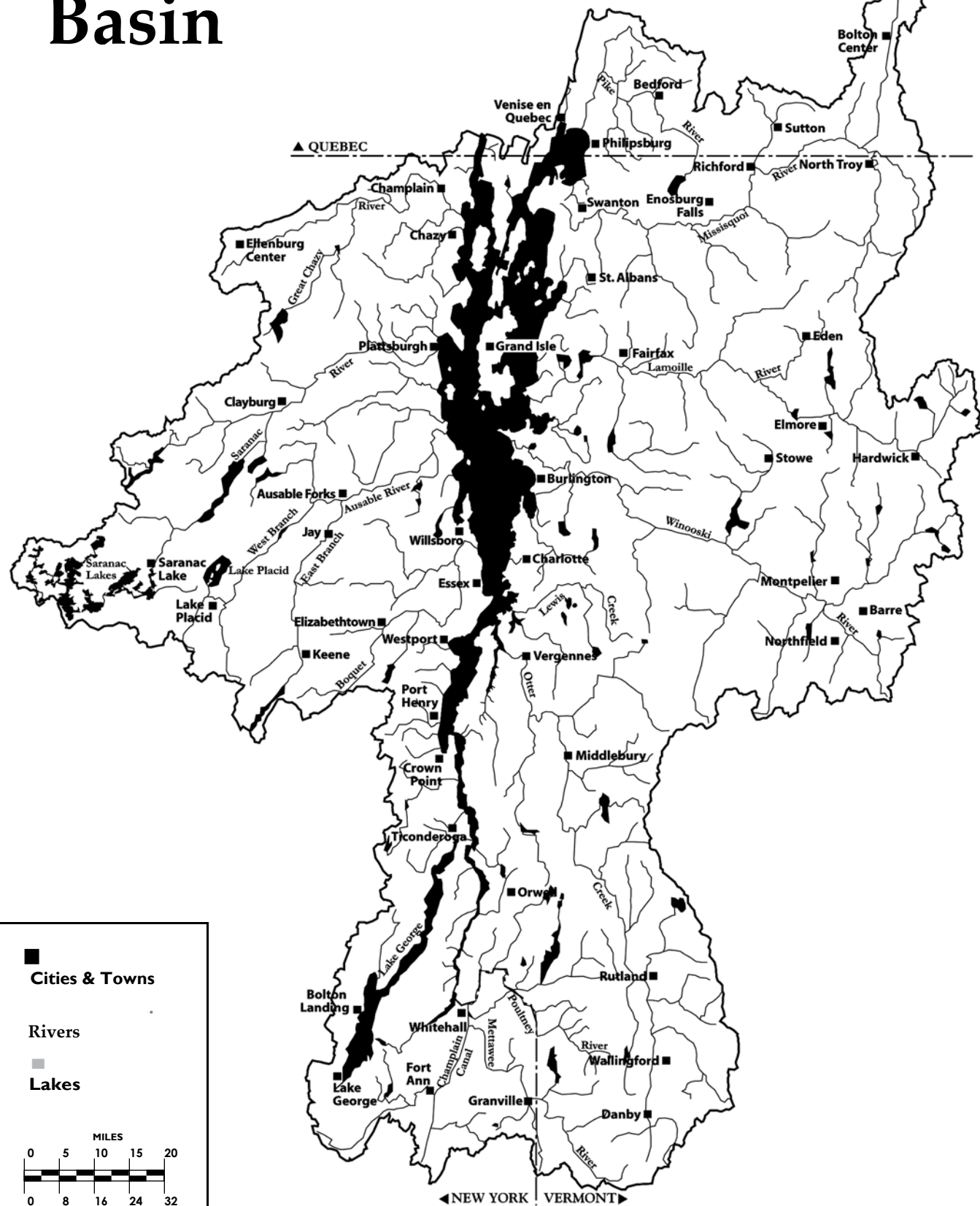
Land Use & Vegetation in the Lake Champlain Basin



Source: T. Millet, LCBP Technical Report No. 24, October 1997

Lake Champlain Basin

N



Did You Know...

- The 2000 United States and Canadian census data recorded 571,000 people living in the Basin.
- The population of the Basin has been growing at an average of 1.2% per year of the last 40 years. The population density of the Basin is 73 people per square mile.
- Approximately one third of the Basin's residents use the Lake as a source of drinking water.
- Lake Champlain flows north to the St. Lawrence River, but during the Ice Age it flowed south, emptying into the Hudson River.
- In 1998-1999 tourist expenditures in the Basin totaled an estimated \$3.8 billion.
- In 2004 Vermont and New York enacted a reciprocal fishing program enabling over a million fishing license holders to fish the majority of Lake Champlain with only one license.
- Besides humans, the Lake's ecosystem includes about 91 species of fish, 312 species of birds, 56 species of mammals, 21 species of amphibians, and 20 of reptiles.
- 12 bird species are listed by New York, Vermont and/or the federal government as endangered or threatened.
- Lake Champlain has 45 marinas.
- Nonpoint sources are estimated to account for about 90% of the total phosphorus load to Lake Champlain, with point sources contributing the remaining 10%.
- On a typical summer day in 1992 over 7,500 motor boats, more than 3,000 sail boats, at least 15 commercial vessels, and thousands of swimmers, windsurfers, kayakers, canoers, scuba divers and other recreationists were enjoying Lake Champlain.

LCBP Goals

The Lake Champlain Basin Program is implementing a plan to insure that the Lake and its drainage basin will be restored, protected and maintained so that current and future generations will enjoy its full benefits. If you would like to learn more about the program or become involved with associated projects, call: 1-800-468-LCBP or visit our web site at www.lcbp.org

Revision: Mac Lynch, 2006

Text: Patrice Farrey

Spot illustrations: George Daly

Design: Don Hanson

Produced through a grant from

EPA in coordination with NEIWPCC

X7-97105501-0

Vital Statistics

TOTAL AREA OF BASIN: 8,234 square miles (21,325 km²), about the size of New Jersey; 56% of the Basin lies in Vermont, 37% in New York, and 7% in Canada.

LENGTH OF LAKE: 120 miles (193km) flowing North from Whitehall, NY to the Richelieu River in Quebec.

WIDTH OF LAKE: 12 miles (19km) at widest point

DEPTH OF LAKE: Average 64 ft with the deepest part being over 400ft (12m). Water levels fluctuate in response to precipitation, temperature, and runoff.

SURFACE WATER: 435 square miles (1,127 km²).

SHORELINE: 587 miles (945 km).

ISLANDS: Over 70 islands.

DRAINAGE: Tributaries that drain the basin contribute more than 90% of the water which enters Lake Champlain.

MAJOR TRIBUTARIES: In New York: the Great Chazy, Saranac, Ausable and Boquet Rivers; In Vermont: the Missisquoi, Lamoille, Winooski and LaPlatte Rivers, and Otter Creek.

AVERAGE AIR TEMPERATURE: 40 F/4 C

GROWING SEASON: averages from 150 days on the shoreline to 105 days in the high altitudes.

AVERAGE ANNUAL PRECIPITATION: 30"(76cm) in the valley, 50"(127cm) in the mountains.

MEAN ANNUAL WATER LEVEL: 95.5 feet above sea level.

GLOSSARY

ecosystem – a biological community together with the physical and chemical environment with which it interacts.

habitat – the specific area that provides the basic requirements of survival for a particular type of plant or animal.

nonpoint source pollution – diffuse sources of pollutants that cannot be attributed to a single discharge point.

point source pollution – discharges from specific identifiable sources.

wetland – the transitional area between land and water. Swamps, bogs and marshes are examples of wetlands.

Perspective

Prokaryotes: The unseen majority

William B. Whitman^{*†}, David C. Coleman[‡], and William J. Wiebe[§]

Departments of ^{*}Microbiology, [‡]Ecology, and [§]Marine Sciences, University of Georgia, Athens GA 30602

ABSTRACT The number of prokaryotes and the total amount of their cellular carbon on earth are estimated to be $4\text{--}6 \times 10^{30}$ cells and 350–550 Pg of C (1 Pg = 10^{15} g), respectively. Thus, the total amount of prokaryotic carbon is 60–100% of the estimated total carbon in plants, and inclusion of prokaryotic carbon in global models will almost double estimates of the amount of carbon stored in living organisms. In addition, the earth's prokaryotes contain 85–130 Pg of N and 9–14 Pg of P, or about 10-fold more of these nutrients than do plants, and represent the largest pool of these nutrients in living organisms. Most of the earth's prokaryotes occur in the open ocean, in soil, and in oceanic and terrestrial subsurfaces, where the numbers of cells are 1.2×10^{29} , 2.6×10^{29} , 3.5×10^{30} , and $0.25\text{--}2.5 \times 10^{30}$, respectively. The numbers of heterotrophic prokaryotes in the upper 200 m of the open ocean, the ocean below 200 m, and soil are consistent with average turnover times of 6–25 days, 0.8 yr, and 2.5 yr, respectively. Although subject to a great deal of uncertainty, the estimate for the average turnover time of prokaryotes in the subsurface is on the order of $1\text{--}2 \times 10^3$ yr. The cellular production rate for all prokaryotes on earth is estimated at 1.7×10^{30} cells/yr and is highest in the open ocean. The large population size and rapid growth of prokaryotes provides an enormous capacity for genetic diversity.

Although invisible to the naked eye, prokaryotes are an essential component of the earth's biota. They catalyze unique and indispensable transformations in the biogeochemical cycles of the biosphere, produce important components of the earth's atmosphere, and represent a large portion of life's genetic diversity. Although the abundance of prokaryotes has been estimated indirectly (1, 2), the actual number of prokaryotes and the total amount of their cellular carbon on earth have never been directly assessed. Presumably, prokaryotes' very ubiquity has discouraged investigators, because an estimation of the number of prokaryotes would seem to require endless cataloging of numerous habitats.

To estimate the number and total carbon of prokaryotes on earth, several representative habitats were first examined. This analysis indicated that most of the prokaryotes reside in three large habitats: seawater, soil, and the sediment/soil subsurface. Although many other habitats contain dense populations, their numerical contribution to the total number of prokaryotes is small. Thus, evaluating the total number and total carbon of prokaryotes on earth becomes a solvable problem.

Aquatic Environments. Numerous estimates of cell density, volume, and carbon indicate that prokaryotes are ubiquitous in marine and fresh water (e.g., 3–5). Although a large range of cellular densities has been reported ($10^4\text{--}10^7$ cells/ml), the mean values for different aquatic habitats are surprisingly similar. For the continental shelf and the upper 200 m of the open ocean, the cellular density is about 5×10^5 cells/ml. A

portion of these cells are the autotrophic marine cyanobacteria and *Prochlorococcus* spp., which have an average cellular density of 4×10^4 cells/ml (6). The deep (>200 m) oceanic water contains 5×10^4 cells/ml on average. From global estimates of volume, the upper 200 m of the ocean contains a total of 3.6×10^{28} cells, of which 2.9×10^{27} cells are autotrophs, whereas ocean water below 200 m contains 6.5×10^{28} cells (Table 1).

The upper 10 cm of sediment in the open ocean is included in the oceanic habitat because, as a result of animal mixing and precipitation, it is essentially contiguous with the overlying water column. Most of the marine sediment is found in the continental rise and abyssal plain, so the numbers of prokaryotes were calculated from an arithmetic average of the cellular densities in the studies cited by Deming and Baross (ref. 9; Table 1). The Nova Scotian continental rise was excluded from this calculation because of its unusual hydrology (10).

There are fewer estimates of the number of prokaryotes in freshwaters and saline lakes (5). Given an average density of 10^6 cells/ml, the total number of cells in freshwaters and saline lakes is 2.3×10^{26} . This value is three orders of magnitude below the numbers of prokaryotes in seawater.

In the polar regions, a relatively dense community of algae and prokaryotes forms at the water–ice interface in annual sea ice (11). In Antarctic sea ice, the estimated number of prokaryotes (2.2×10^{24} cells) was based on the mean cell numbers of Delille and Rosiers (12) and the mean areal extent of seasonal ice (13). If the population size in the Arctic is similar (14), the global estimate for both polar regions is 4×10^{24} cells, only a fraction of the total number of prokaryotes.

Soil. Soil is a major reservoir of organic carbon on earth and an important habitat for prokaryotes. Prokaryotes are an essential component of the soil decomposition subsystem, in which plant and animal residues are degraded into organic matter and nutrients are released into food webs (15). Many studies indicate that the number of prokaryotes in forest soils is much less than the number in other soils. The total number of prokaryotes in forest soil was estimated from detailed direct counts from a coniferous forest ultisol (16), which were considered representative of forest soils in general (Table 2). For other soils, including grasslands and cultivated soils, the numbers of prokaryotes appear about the same, e.g., the number of prokaryotes in Negev desert soil is comparable to the number in cultivated soil (19). Therefore, the numbers of prokaryotes in all other soils were estimated from the unpublished field studies of E. A. Paul for cultivated soils (cited in ref. 18).

Subsurface. The subsurface is defined here as terrestrial habitats below 8 m and marine sediments below 10 cm. Few direct enumerations of subsurface prokaryotes have been made, largely because of the difficulty in obtaining uncontaminated samples. Nevertheless, circumstantial evidence suggests that the subsurface biomass of prokaryotes is enormous (20). For instance, groundwater from deep aquifers and formation

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in

accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1998 by The National Academy of Sciences 0027-8424/98/956578-6\$2.00/0
PNAS is available online at <http://www.pnas.org>.

[†]To whom reprint requests should be addressed: e-mail: whitman@uga.cc.uga.edu.

Table 1. Number of prokaryotes in aquatic habitats

Habitat	Volume,* cm ³	Cells/ml, X 10 ⁵	Total no. of cells, X 10 ²⁶
Marine			
Continental shelf	2.03 X 10 ²⁰	5	1.0
Open ocean			
Water, upper 200 m	7.2 X 10 ²²	5	360
Water, below 200 m	1.3 X 10 ²⁴	0.5	650
Sediment, 0–10 cm	3.6 X 10 ¹⁹	4600	170
Fresh			
Lakes	1.25 X 10 ²⁰	10	1.3
Rivers	1.2 X 10 ¹⁸	10	0.012
Saline lakes	1.04 X 10 ²⁰	10	1.0
Total			1180

*Marine, freshwater, and saline lake volumes were calculated from refs. 7 and 8.

water from petroleum deposits contain 10³–10⁶ prokaryotic cells/ml (21, 22).

Unconsolidated sediments represent most of the marine subsurface and about 20% of the terrestrial subsurface (23). The number and sizes of subsurface prokaryotes in unconsolidated sediments of the deep ocean and the continental shelf and slope (24–30) and the terrestrial coastal plain (31, 32) have been determined. Because the terrestrial values fall within the range of the marine values, arithmetic averages were calculated to create a depth profile to 600 m (Table 3). For deeper sediments to 4 km, the number of prokaryotes was extrapolated from the formula of Parkes *et al.* (33). At 4 km, the average temperature reaches 125°C (34), which is close to the upper temperature limit for prokaryotic life.

Of the 3.8 X 10³⁰ prokaryotes calculated to be in the unconsolidated subsurface sediments, 97% or 3.7 X 10³⁰ occur at depths shallower than 600 m (Table 3). The estimated number of prokaryotes for deeper sediments is only 0.13 X 10³⁰ cells. This value is uncertain because it is based on extrapolation. In addition, the accuracy also depends on whether or not the data used to calculate the depth profile are representative of the entire subsurface. Because most of these data were obtained from regions of the Pacific Ocean, the depth profile is likely to be most accurate for those sediments.

The estimated number of terrestrial subsurface prokaryotes (Table 3, 2.5 X 10²⁹) is a minimum value because it is limited

Table 2. Number of prokaryotes in soil

Ecosystem type*	Area, X 10 ¹² m ²	No. of cells,† X 10 ²⁷
Tropical rain forest	17.0	1.0
Tropical seasonal forest	7.5	0.5
Temperate evergreen forest	5.0	0.3
Temperate deciduous forest	7.0	0.4
Boreal forest	12.0	0.6
Woodland and shrubland	8.0	28.1
Savanna	15.0	52.7
Temperate grassland	9.0	31.6
Desert scrub	18.0	63.2
Cultivated land	14.0	49.1
Tundra and alpine	8.0	20.8
Swamps and marsh	2.0	7.3
Total	123.0	255.6

*From ref. 73.

†For forest soils, the number of prokaryotes in the top 1 m was 4 X 10⁷ cells per gram of soil, and in 1–8 m, it was 10⁶ cells per gram of soil (16). For other soils, the number of prokaryotes in the top 1 m was 2 X 10⁹ cells per gram of soil, and in 1–8 m, it was 10⁸ cells per gram of soil (18). The boreal forest and tundra and alpine soils were only 1 m deep. A cubic meter of soil was taken as 1.3 X 10⁶ g.

Table 3. Total number of prokaryotes in unconsolidated subsurface sediments

Depth interval,* m	Cells/cm ³ , X 10 ⁶	No. of cells, X 10 ²⁸		
		Deep oceans†	Continental shelf and slope‡	Coastal plains§
0.1	220.0¶	66.0	14.5	4.4
10	45.0¶	121.5	26.6	8.1
100	6.2¶	18.6	4.1	1.2
200	19.0¶	57.0	12.5	3.8
300	4.0¶	12.0	2.6	0.8
400	7.8¶		10.1	3.2
600	0.95		3.7	1.2
1,200	0.61		3.2	1.0
2,000	0.44		2.6	0.9
3,000	0.34			0.7
Total		275.1	79.9	25.3

Grand Total: 380 X 10²⁸ 3.8 X 10³⁰

*Depth intervals are designated by the upper boundary. Thus, "0.1" represents 0.1–10 m and "3,000" represents 3,000–4,000.

†Corresponds to seismic layer I (23).

‡Corresponds to subcontinental sediments (23).

§Corresponds to geosyncline sediments of Mesozoic origin (23).

¶Calculated from the arithmetic averages.

Calculated by extrapolation of the formula of Parkes *et al.* (33).

to unconsolidated sediments, which represent only 20% of the terrestrial subsurface. Two other approaches can be used to estimate the total number of terrestrial subsurface prokaryotes. The first approach, originally used by Gold (20), is based on the assumption that the average porosity of the terrestrial subsurface is 3%. Assuming that the percentage of the total pore space occupied by prokaryotes is 0.016% (35), the average volume of a subsurface prokaryotic cell is 1.07 X 10⁻¹² cm³ (36), and the volume of the upper 4 km of the terrestrial subsurface is 4.9 X 10²³ cm³, the total number of terrestrial subsurface prokaryotes is 2.2 X 10³⁰ cells. Considering the general nature of these assumptions, the agreement within an order-of-magnitude of the estimate in Table 1 provides some confidence in the latter estimate.

Alternatively, the number of terrestrial subsurface prokaryotes can be estimated from groundwater data. Based on values from seven sites and four studies (31, 37–39), the average number of unattached cells in groundwater is 1.54 X 10⁵ cells/ml. The total volume of groundwater in the upper 4 km of the earth's surface is 9.5 X 10²¹ cm³ (40), and thus the number of unattached prokaryotes in groundwater is 1.46 X 10²⁷ cells. However, the number of prokaryotes in aquifer

sediments is probably many orders of magnitude greater than the number unattached in the groundwater per se. For an aquifer 30–200 m deep, only 0.058% of the prokaryotes are unattached (calculated from the data of refs. 31, 41, and 42). This value appears to be representative of groundwater from other deep aquifers (22, 37), which implies that the terrestrial subsurface contains about 2.5 X 10³⁰ prokaryotic cells. This estimate contains two major uncertainties. First, about 55% of the earth's groundwater is found below 750 m (40), and the extrapolation of values from the groundwater and aquifers above 750 m may not be applicable. Second, the ratio of unattached prokaryotes in aquifers was calculated from unconsolidated sediments, and the ratio may vary in other types of aquifers where the physical properties of the rocks and sediments are very different.

In summary, the subsurface is a major habitat for prokaryotes, and the number of subsurface prokaryotes probably exceeds the numbers found in other components of the biosphere. The greatest uncertainty is in the estimate for the terrestrial subsurface because this estimate is based on only a few measurements. However, even for the terrestrial subsurface, two independent methods suggest that the number of

prokaryotes is very large, about $2.5\text{--}25 \times 10^{29}$ cells. Thus, the total number of subsurface prokaryotes is probably $3.8\text{--}6.0 \times 10^{30}$ cells.

Other Habitats. Although they were found not to constitute a large fraction of the total number of prokaryotes, other habitats are of interest in their own right.

Animals. Many vertebrate and invertebrate animals contain dense populations of prokaryotes that play important roles in nutrition and disease. To estimate the total number of prokaryotes on and within animals, the numbers of prokaryotes in each individual animal and the population size of the animal must be known. Unfortunately, these values are only known for a small number of mostly domestic animals.

In mammals and birds, prokaryotes are abundant on the skin and within the gastrointestinal tract. Within the gastrointestinal tract, most of the prokaryotes are anaerobes in the colon, cecum, or rumen (43, 44), and the total number found within animals whose population sizes are known can be readily calculated (Table 4). For comparison, the numbers of prokaryotes on the skin of humans can be calculated. The density of prokaryotes is about $10^3\text{--}10^4$ cells/cm², except in the groin and axilla, where it is 10^6 cells/cm² (57). Based on the surface area of an adult (58), the total number of prokaryotes on the skin of an individual is about 3×10^8 cells, a value far below the number of prokaryotes in the colon (Table 4).

Insects, such as termites, cockroaches, and crane flies, harbor dense prokaryotic populations in their hindguts (53, 59, 60). Because the number of termites in the world has been estimated and the number of prokaryotes for at least one type of termite has been measured (53, 55), it is possible to estimate the total number of prokaryotes in termites (Table 4). Although huge, this value is much smaller than the total number of prokaryotes found in many other habitats.

Although the number of prokaryotes in the gastrointestinal tracts of animals is enormous, it is unlikely to represent a large fraction of the total prokaryotes on earth. For example, the number of prokaryotes in the bovine rumen is 4–6 orders of magnitude less than the numbers found in soil, the subsurface, and sea water. Therefore, although the numbers of prokaryotes are known for only a few groups of animals, it is unlikely that animals contain a major fraction of the total number of prokaryotes.

Leaves. Although prokaryotes associated with plant roots are measured with other soil prokaryotes for methodological reasons, leaves and other plant tissues also harbor large populations of prokaryotes. Leaf area can be estimated from the leaf area index. The numbers of prokaryotes on leaves are highly variable, but the viable count (cfu or colony-forming units) rarely exceeds $10^4\text{--}10^6$ cfu/cm² (61–64). An upper limit for the number of prokaryotes on leaves can be estimated by assuming a dense population and a high leaf area index. Assuming a leaf area index of 10, which is typical of many

forests, the maximum number of prokaryotes would be about 10^{11} cfu/m². A forest soil contains about 6×10^{13} cells/m² (see Table 2). Even if the viable counts are 1–10% of the direct counts, the maximum number of prokaryotes on leaves is unlikely to exceed the number in soil. In fact, in a temperate forest, the number of prokaryotes on leaves is a small fraction of the number in the underlying soil (65).

Air. By volume, the atmosphere represents the largest compartment of the biosphere, and prokaryotes have been detected at altitudes as high as 57–77 km (66). Nevertheless, the total number of airborne prokaryotes appears to be quite low. For the bottom 3 km of the atmosphere, the total number of prokaryotes over land is about 5×10^{19} cfu (calculated from refs. 67–69), a value so low that it is unlikely that airborne prokaryotes represent a large fraction of the total number of prokaryotes.

Carbon Content. The amount of carbon in prokaryotes can be estimated from the cell numbers in soil, aquatic systems, and the subsurface. In the soil and subsurface, the cellular carbon is assumed to be one-half of the dry weight. In soil, the average dry weight of a prokaryotic cell is 2×10^{-13} g or 200 fg (18). Thus, the total prokaryotic cellular carbon in soil is 26×10^{15} g of C or 26 Pg of C (Table 5). In the subsurface, there is only one measurement of the average dry weight of cells, that of 172 fg for cells from a terrestrial aquifer (36). This value yields an estimate of the terrestrial prokaryotic cellular carbon of 22–215 Pg of C (Table 5). The estimate for the marine subsurface, 303 Pg of C (Table 5), may be compared with 56 Pg of C, the value obtained by Parkes *et al.* (33). The difference, 5.4-fold, is due in part to how the depth integrations were calculated. Parkes *et al.* (33) used logarithmic extrapolations rather than arithmetic averages, which decreased their estimated number of cells by 3-fold. They also estimated the amount of carbon per cell at 65 fg of C rather than the 86 fg of C used here. The remaining difference occurs because the current estimate is based in part on additional marine and terrestrial data.

For aquatic systems, the average cellular carbon and volume has been a matter of considerable discussion, and the range in average cellular carbon reported is 5–20 fg of C/cell (5, 17, 70–72). To obtain the estimate of 2.2 Pg of C (Table 5), the average cellular carbon for sedimentary (9) and planktonic prokaryotes (17, 70–72) was assumed to be 10 and 20 fg of C/cell, respectively. If the average cellular carbon is assumed to be 5 fg of C/cell, the total amount of prokaryotic cellular carbon would be 0.6 Pg of C.

Discussion. The total carbon of prokaryotes on earth is enormous, approximately 60–100% of the total carbon found in plants (Tables 5 and 6). Inclusion of this carbon in global models will greatly increase estimates of the amount of carbon stored in living organisms. In addition, prokaryotes contain large amounts of N, P, and other essential nutrients. For instance, assuming a C/N/P ratio in prokaryotes of

Table 4. Total number of prokaryotes in some representative animals

Animal	Organ	Cells/ml or cells/g	Organ contents*	No. of animals†	No. of cells, $\times 10^{23}$	Refs.
Human	Colon	3.2×10^{11}	220 g	5.6×10^9	3.9	45, 46
Cattle	Rumen	2.1×10^{10}	106 liter	1.3×10^9	29.0	47, 48
Sheep and goats	Rumen	4.4×10^{10}	12 liter	1.7×10^9	9.0	47, 48
Pigs	Colon	$5.4 \times 10^{10}\ddagger$	9 liter	8.8×10^8	4.3	49, 50
	Cecum	$2.8 \times 10^{10}\ddagger$	1 liter	8.8×10^8	0.3	49, 50
Domestic birds§	Cecum	9.5×10^{10}	2 g	1.3×10^{10}	0.024	51, 52
Termites	Hindgut	$2.7 \times 10^6\parallel$		2.4×10^{17}	6.5	53

*Organ contents in volume or grams of wet weight. For comparison, the volume of the human colon is 0.5 liter. For domestic birds, weight wet was calculated from a volume of 2 ml assuming that 1 ml = 1 g wet weight.

†Values from the *FAO Production Yearbook* (54), except for the termites value which was from ref. 55.

‡The direct count was assumed to be 2.7×10^6 viable count (56).

§Includes chickens, ducks, and turkeys.

¶Per termite.

Table 5. Number and biomass of prokaryotes in the world

Environment	No. of	
	prokaryotic cells, X 10 ²⁸	Pg of C in prokaryotes*
Aquatic habitats	12	2.2
Oceanic subsurface	355	303
Soil	26	26
Terrestrial subsurface	25–250	22–215
Total	415–640	353–546

*Calculated as described in the text.

1:0.24:0.025 (74), the entire prokaryotic pool for N and P is 85–130 Pg of N and 9–14 Pg of P. In all plants, assuming C/N and C/P ratios for the 471 Pg of plant C in forests and woodlands of 156 and 1340, respectively, and C/N and C/P ratios for the 88 Pg of plant C in other ecosystems of 12.5 and 125, respectively (73), the amounts of N and P are 10 Pg and 1.05 Pg, respectively. Thus, the plant pool for these nutrients is an order of magnitude smaller than the total prokaryotic pool. In fact, the amount of N and P in soil prokaryotes, 6.2 Pg and 0.65 Pg, respectively, is nearly equal to the amount in terrestrial plants even though terrestrial plants contain much more carbon. Other essential nutrients are probably distributed similarly, and prokaryotes may represent the largest living reservoir for these elements on earth.

The abundance of prokaryotic carbon and other elements may be compared with the statement of Kluyver that about one-half of the “living protoplasm” on earth is microbial (2). Because most of the plant biomass is made up of extracellular material such as cell walls and structural polymers, the protoplasmic biomass of prokaryotes probably far exceeds that of plants, and Kluyver’s well-accepted estimate is probably much too conservative.

From the estimate of prokaryotic carbon in soil and aquatic habitats, it is possible to set some limits for the average growth or turnover rates for these populations. Assuming an efficiency of carbon assimilation of 0.2 (75, 76), the amount of “net productivity” necessary to support the turnover of prokaryotes in the upper 200 m of the ocean is four times their carbon content or 0.7–2.9 Pg of C (depending on the amount of carbon per cell). Given that about 85% of the net productivity is consumed in the upper 200 m (73) and assuming that all of this carbon is used by prokaryotes, the average turnover rate

cannot exceed 15–60 yr⁻¹, and the average generation time cannot be less than 6–25 days. For the upper 200 m of the open ocean, the reported average generation time is 2.5–27 days (3). Similar calculations for the deep ocean (below 200 m) and soil suggest that the average turnover rate for prokaryotes cannot exceed approximately 1.2 and 0.4 yr⁻¹, respectively. The value for soil is not greatly different from current estimates for the upper portion of the soil of 0.4–2 yr⁻¹ (77–79). Thus, our estimates of the prokaryotic cellular carbon in the upper ocean and soil are consistent with published productivity estimates.

Results from a similar analysis for the subsurface pro-

karyotes are problematic. Assuming that 1 Pg of C/yr, or about 1% of the total net productivity, reaches the subsurface and that the net burial rate is 0.06 Pg of C/yr (73), only 0.94 Pg of C/yr is available to support the subsurface community of

Table 6. Relationship of plant and prokaryotic biomass to primary productivity

Ecosystem	Net primary productivity,* Pg of C/yr	Total carbon content, Pg of C		
		Plant*	Soil and aquatic prokaryotes	Subsurface prokaryotes
Terrestrial	48	560	26	22–215
Marine	51	1.8	2.2	303

*From ref. 73.

prokaryotes. If the efficiency of carbon assimilation is 0.20, then the calculated average turnover time is 1–2 X 10³ yr, far

longer than found in other ecosystems. At present, a number of plausible explanations for this apparent anomaly exist. (i) The average turnover time could be on the order of 1,000 yr.

If this were the case, most of the subsurface prokaryotes must be metabolically inactive and probably nonviable. Circumstantial evidence suggests that this is not the case, and viability of subsurface prokaryotes is within the range observed for prokaryotes from surface sediments and soils (cf. 24, 31). Sulfate reduction, methanogenesis, and other activities have also been detected in cores from the subsurface (24). Thus, although it is likely that the relative metabolic activity and rate of carbon consumption of subsurface bacteria are lower than that found on the surface, activity must still be sufficient to maintain culture viability. (ii) Lithoautotrophic processes may provide an additional source of energy for growth of subsurface prokaryotes. Although lithoautotrophy has been demonstrated in some geological formations, current evidence suggests that most of the subsurface biomass is supported by organic matter deposited from the surface (80–82). Because the data are so limited, future studies could revise this view. (iii) The subsurface biomass may be overestimated. The estimate of subsurface carbon is based on a conversion factor derived from data at one site, which may not be representative. However, given that some of the smallest cells so far described in nature contain 5 fg of C, the magnitude of this error is unlikely to be more than 10- to 20-fold. (iv) The efficiency of carbon assimilation may be underestimated. Pure culture studies with rich media suggest that the efficiency of carbon assimilation can be as high as 0.85 (83). However, the error associated with this factor cannot be more than 4-fold. These points, when considered together, emphasize that our current understanding of subsurface prokaryotes is incomplete. Because of their numerical importance, more extensive examination of this habitat is imperative.

The large population size of prokaryotes implies that events that are extremely rare in the laboratory could occur frequently in nature. For instance, prokaryotes have an enormous potential to accumulate mutations and, thus, to acquire genetic diversity. However, the population size itself is not altogether an accurate measure of the potential for mutational change, which must also include the growth rates of the populations. Large, slowly growing populations may produce fewer cells and fewer mutational events than smaller, rapidly growing populations do. Even with the uncertainties for the average growth rates for many natural populations discussed above, it is still possible to estimate the cellular production rates and hence the frequency of these rare events (Table 7). Although subsurface prokaryotes predominate numerically, their cellular productivity is comparable to that of the much smaller but more rapidly growing population associated with domestic animals

Table 7. Annual cellular production of prokaryotes in various habitats

Habitat	Population size	Turnover time, days	Cells/yr, X 10 ²⁹
Marine heterotrophs			
Above 200 m	3.6 X 10 ²⁸	16*	8.2
Below 200 m	8.2 X 10 ²⁸	300*	1.1
Marine autotrophs	2.9 X 10 ²⁷	1.5†	7.1
Soil	2.6 X 10 ²⁹	900*	1.0
Subsurface	4.9 X 10 ³⁰ *	5.5 X 10 ⁵ *	0.03
Domestic mammals	4.3 X 10 ²⁴ ‡	1§	0.02

*The value or mean of the range discussed in the text.

†Based on the median generation time of *Prochlorococcus* (84).

‡Sum of the number of prokaryotes in cattle, sheep, goats, and pigs

6582 Perspective: Whitman *et al.*
Table 4.
§From ref. 85.

(Table 7). The highest cellular productivity is found in the open ocean (Table 7). Thus, mutations and other rare genetic events are more likely to occur in the population of marine prokaryotes than in populations in other habitats.

Genes that are widely distributed in prokaryotes have a tremendous opportunity for mutational change, and the evolution of conserved genes must be otherwise greatly constrained. Assuming a prokaryotic mutation rate of 4×10^{-7} mutations per gene per DNA replication (86, 87), four simultaneous mutations in every gene shared by the populations of marine heterotrophs (in the upper 200 m), marine autotrophs, soil prokaryotes, or prokaryotes in domestic animals would be expected to occur once every 0.4, 0.5, 3.4, or 170 hr, respectively. Similarly, five simultaneous mutations in every gene shared by all four populations would be expected to occur every 60 yr. The capacity for a large number of simultaneous mutations distinguishes prokaryotic from eukaryotic evolution and should be explicitly considered in methods of phylogenetic analyses.

For essentially asexual, haploid organisms such as prokaryotes, mutations are a major source of genetic diversity and one of the essential factors in the formation of novel species. Given prokaryotes' enormous potential to acquire genetic diversity, the number of prokaryotic species may be very large. Recent estimates for the number of prokaryotic species range from 10^5 to 10^7 (88). However, the current definition of a prokaryotic species, which includes strains whose genomic DNAs form hybrids with a change in the melting temperature (ΔT_m) of less than 5°C (89), may be misleading. Application of the same definition to eukaryotes would lead to the inclusion of members of many taxonomic tribes into the same species (90). Similarly, phylogenetic groups such as humans, orangutans and gibbons would also belong to the same species (91). Thus, a simple comparison of the number of eukaryotic and prokaryotic species greatly underestimates prokaryotic diversity. Given prokaryotes' numerical abundance and importance in biogeochemical transformations, the absence of detailed knowledge of prokaryotic diversity is a major omission in our knowledge of life on earth.

We are grateful to our colleagues, whose understanding, generosity, and sense of humor made this project possible. They include, but are not limited to, M. Azain, B. Binder, J. F. Dowd, R. P. Freeman-Lynde, T. C. Hazen, T. Hollibaugh, S. Kayar, M. Lee, S. Martin, M. Moran, W. J. Payne, L. Pomeroy, J. B. Risatti, and J. Russell. We acknowledge support from National Science Foundation Grants BIR-94-13235 (W.B.W. and D.C.C.), DEB 96-32854 (D.C.C.), and DEB 94-12089 (W.J.W.) and Department of Energy Grant DE-FG02-97ER20269 (W.B.W.).

- Romankevich, E. A. (1988) *Geokhimiya* **(2)**, 292–306.
- Kluyver, A. J. & van Niel, C. B. (1956) *The Microbe's Contribution to Biology* (Harvard Univ. Press, Cambridge, MA), p. 3.
- Ducklow, H. W. & Carlson, C. A. (1992) *Adv. Microb. Ecol.* **12**, 113–181.
- Simon, M. (1994) *Arch. Hydrobiol.* **130**, 283–302.
- Fry, J. C. (1988) in *Methods in Aquatic Bacteriology*, ed. Austin, B. (Wiley, London), pp. 27–72.
- Campbell, L. & Vaulot, D. (1993) *Deep-Sea Res. I* **40**, 2043–2060.
- Garrison, T. (1994) *Essentials of Oceanography* (Wadsworth, New York), p. 353.
- Mitsch, W. J., Mitsch, R. H. & Turner, R. E. (1994) in *Global Wetlands: Old World and New*, ed. Mitsch, W. J. (Elsevier, New York), pp. 3–56.
- Deming, J. W. & Baross, J. A. (1993) in *Organic Geochemistry: Principles and Applications*, eds. Engel, M. H. & Macko, S. A. (Plenum, New York), pp. 119–144.
- Aller, J. Y. & Aller, R. C. (1986) *Deep-Sea Res.* **33**, 755–790.
- Vincent, W. F. (1988) *Microbial Ecosystems of Antarctica* (Cambridge Univ. Press, New York).
- Delille, D. & Rosiers, C. (1995) *Polar Biol.* **16**, 27–34.
- Gordon, A. L. (1981) *J. Geophys. Res.* **86**, 4193–4197.
- Smith, R. E. H., Clement, P. & Cota, G. F. (1989) *Microb. Ecol.* **17**, 63–76.
- Coleman, D. C. & Crossley, D. A., Jr. (1996) *Fundamentals of Soil Ecology* (Academic, San Diego).
- Richter, D. D. & Markewitz, D. (1995) *Bioscience* **45**, 600–609.
- Bratbak, G. & Dundas, I. (1984) *Appl. Environ. Microbiol.* **48**, 755–757.
- Paul, E. A. & Clark, F. E. (1989) *Soil Microbiology and Biochemistry* (Academic, New York).
- Fließbach, A., Sarig, S. & Steinberger, Y. (1994) *Arid Soil Res. Rehabil.* **8**, 353–362.
- Gold, T. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 6045–6049.
- Nilsen, R. K., Beeder, J., Thorstenson, T. & Torsvik, T. (1996) *Appl. Environ. Microbiol.* **62**, 1793–1798.
- Pedersen, K. (1993) *Earth-Sci. Rev.* **34**, 243–260.
- Ronov, A. B. & Yaroshevsky, A. A. (1969) in *Geophysical Monograph 13: The Earth's Crust and Upper Mantle*, ed. Hart, P. J. (Am. Geophys. Union, Washington, DC), pp. 37–57.
- Cragg, B. A., Harvey, S. M., Fry, J. C., Hebert, R. A. & Parkes, R. J. (1992) *Proc. Ocean Drill. Program: Sci. Results* **127/128 (Pt. 1)**, 761–776.
- Cragg, B. A. (1994) *Proc. Ocean Drill. Program: Sci. Results* **135**, 147–150.
- Cragg, B. A. & Kemp, A. E. S. (1995) *Proc. Ocean Drill. Program: Sci. Results* **138**, 599–604.
- Cragg, B. A. & Parkes, R. J. (1994) *Proc. Ocean Drill. Program: Sci. Results* **139**, 509–516.
- Cragg, B. A., Parkes, R. J., Fry, J. C., Hebert, R. A., Wimpenny, J. W. T. & Getliff, J. M. (1990) *Proc. Ocean Drill. Program: Sci. Results* **112**, 607–619.
- Cragg, B. A., Parkes, R. J., Fry, J. C., Weightman, A. J., Rochelle, P. A., Maxwell, J. R., Kastner, M., Hovland, M., Whiticar, M. J. & Sample, J. C. (1995) *Proc. Ocean Drill. Program: Sci. Results* **146 (Pt. 1)**, 399–411.
- Cragg, B. A., Parkes, R. J., Fry, J. C., Weightman, A. J., Maxwell, J. R., Kastner, M., Hovland, M., Whiticar, M. J., Sample, J. C. & Stein, R. (1995) *Proc. Ocean Drill. Program: Sci. Results* **146 (Pt. 2)**, 139–144.
- Hazen, T. C., Jimenez, L., Lopez de Victoria, G. & Fliermans, C. B. (1991) *Microb. Ecol.* **22**, 293–304.
- Chapelle, F. H., Zeliber, J. L., Jr., Grimes, D. J. & Knobel, L. L. (1987) *Water Resources Res.* **23**, 1625–1632.
- Parkes, R. J., Cragg, B. A., Bale, S. J., Getliff, J. M., Goodman, K., Rochelle, P. A., Fry, J. C., Weightman, A. J. & Harvey, S. M. (1994) *Nature (London)* **371**, 410–413.
- Garland, G. D. (1971) *Introduction to Geophysics: Mantle, Core, and Crust* (Saunders, Philadelphia).
- Harvey, R. W., Smith, R. L. & George, L. (1984) *Appl. Environ. Microbiol.* **48**, 1197–1202.
- Balkwill, D. L., Leach, F. R., Wilson, J. T., McNabb, J. F. & White, D. C. (1988) *Microb. Ecol.* **16**, 73–84.
- Pedersen, K. & Ekendahl, S. (1990) *Microb. Ecol.* **20**, 37–52.
- Stevens, T. O., McKinley, J. P. & Fredrickson, J. K. (1993) *Microb. Ecol.* **25**, 35–50.
- Pedersen, K., Arlinger, J., Hallbeck, L. & Pettersson, C. (1996) *Mol. Ecol.* **5**, 427–436.
- Berner, E. K. & Berner, R. A. (1987) *The Global Water Cycle* (Prentice-Hall, Englewood Cliffs, NJ).
- Balkwill, D. L. (1989) *Geomicrobiol. J.* **7**, 33–52.
- Sinclair, J. L. & Ghiorse, W. C. (1989) *Geomicrobiol. J.* **7**, 15–31.
- Drasar, B. S. & Barrow, P. A. (1985) *Intestinal Microbiology* (Am. Soc. Microbiol., Washington, DC).
- Smith, H. W. (1965) *J. Pathol. Bacteriol.* **89**, 95–122.
- van Houte, J. & Gibbons, R. J. (1966) *Antonie van Leeuwenhoek* **32**, 212–222.
- Cummings, J. H., Banwell, J. G., Segal, I., Coleman, N., Englyst, H. N. & Macfarlane, G. T. (1990) *Gastroenterology* **98**, A408.
- Hungate, R. E. (1966) *The Rumen and Its Microbes* (Academic, New York), p. 34.
- Habel, R. E. (1975) in *The Anatomy of the Domestic Animals*, ed. Getty, R. (Saunders, Philadelphia), 5th Ed., p. 861.
- Butine, T. J. & Leedle, J. A. Z. (1989) *Appl. Environ. Microbiol.* **55**, 1112–1116.
- Dukes, H. H. (1955) *Physiology of Domestic Animals* (Comstock Pub. Assoc., Ithaca, NY), 7th Ed.
- Barnes, E. M., Mead, G. C., Barnum, D. A. & Harry, E. G. (1972) *Br. Poult. Sci.* **13**, 311–326.

52. Salanitro, J. P., Fairchilds, I. G. & Zgornicki, Y. D. (1974) *Appl. Microbiol.* **27**, 678–687.
53. Schultz, J. E. & Breznak, J. A. (1978) *Appl. Environ. Microbiol.* **35**, 930–936.
54. Food and Agricultural Organization (1994) *FAO Production Yearbook* (Food and Agricultural Organization of the United Nations, Rome), Vol. 48.
55. Zimmerman, P. R., Greenberg, J. P., Wandiga, S. O. & Crutzen, P. J. (1982) *Science* **218**, 563–565.
56. Eller, C., Crabill, M. R. & Bryant, M. P. (1971) *Appl. Microbiol.* **22**, 522–529.
57. Willett, H. P. (1992) in *Zinsser Microbiology*, eds. Joklik, W. K., Willett, H. P., Amos, D. B. & Wilfert, C. M. (Appleton & Lange, Norwalk, CT), 20th Ed., pp. 393–400.
58. Marples, M. J. (1965) *The Ecology of the Human Skin* (Charles C. Thomas, Springfield, IL), p. 6.
59. Cruden, D. L. & Markovetz, A. J. (1984) *Arch. Microbiol.* **138**, 131–139.
60. Klug, M. J. & Kotarski, S. (1980) *Appl. Environ. Microbiol.* **40**, 408–416.
61. Corpe, W. A. & Rheem, S. (1989) *FEMS Microbiol. Lett.* **62**, 243–250.
62. Dickinson, C. H., Austin, B. & Goodfellow, M. (1975) *J. Gen. Microbiol.* **91**, 157–166.
63. Ruinen, J. (1961) *Plant Soil* **15**, 81–109.
64. Van Outryve, M. F., Gossele, F. & Swings, J. (1989) *Microb. Ecol.* **18**, 175–186.
65. Holm, E. & Jensen, V. (1972) *Oikos* **23**, 248–260.
66. Imshenetsky, A. A., Lysenko, S. V. & Kazakov, G. A. (1978) *Appl. Environ. Microbiol.* **35**, 1–5.
67. Fulton, J. D. (1966) *Appl. Microbiol.* **14**, 241–244.
68. Fulton, J. D. (1966) *Appl. Microbiol.* **14**, 245–250.
69. Lighthart, B. & Stetzenbach, L. D. (1994) in *Atmospheric Microbial Aerosols*, eds. Lighthart, B. & Mohr, A. J. (Chapman & Hall, New York), pp. 68–98.
70. Lee, S. H. & Fuhrman, J. A. (1987) *Appl. Environ. Microbiol.* **53**, 1298–1303.
71. Cho, B. C. & Azam, F. (1990) *Mar. Ecol. Prog. Ser.* **63**, 253–259.
72. Carlson, C. A., Bales, N. R., Hansell, D. A. & Ducklow, H. W. (1998) *EOS* **79**, OS171 (abstr.).
73. Schlesinger, W. H. (1997) *Biogeochemistry* (Academic, New York), 2nd Ed.
74. Duboc, P., Schill, N., Menoud, L., van Gulik, W. & von Stockar, U. (1995) *J. Biotechnol.* **43**, 145–158.
75. Cole, J. J., Findlay, S. & Pace, M. L. (1988) *Mar. Ecol. Prog. Ser.* **43**, 1–10.
76. Pomeroy, L. R. & Wiebe, W. J. (1993) *Marine Microb. Food Webs* **7**, 101–118.
77. Coleman, D. C. (1995) in *Food Webs: Integration of Patterns and Dynamics*, eds. Polis, G. A. & Winemiller, K. O. (Chapman & Hall, New York), pp. 39–50.
78. Jenkinson, D. S., Ladd, J. N. & Rayner, J. H. (1980) in *Soil Biochemistry*, eds. Paul, E. A. & Ladd, J. N. (Dekker, New York), Vol. 5, pp. 415–471.
79. Schnurer, J., Clarholm, M. & Rosswall, T. (1985) *Soil Biol. Biochem.* **17**, 611–618.
80. Ekeendahl, S. & Pedersen, K. (1994) *Microbiology* **140**, 1565–1573.
81. Stevens, T. (1997) *FEMS Microbiol. Rev.* **20**, 327–337.
82. Wellsbury, P., Goodman, K., Barth, T., Cragg, B. A., Barnes, S. P. & Parkes, R. J. (1997) *Nature (London)* **388**, 573–576.
83. Payne, W. J. & Williams, M. L. (1976) *Biotechnol. Bioengineer.* **18**, 1653–1655.
84. Binder, B. J., Chisholm, S. W., Olsen, R. J., Frankel, S. L. & Worden, A. Z. (1996) *Deep-Sea Res. II* **43**, 907–931.
85. Russell, J. B. & Bruckner, G. G. (1991) in *Microbiology of Animals and Animal Products*, ed. Woolcock, J. B. (Elsevier, New York), pp. 1–17.
86. Hutchinson, F. (1996) in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, ed. Neidhardt, F. C. (Am. Soc. Microbiol., Washington, DC), 2nd Ed., pp. 2218–2235.
87. Jacobs, K. L. & Grogan, D. W. (1997) *J. Bacteriol.* **179**, 3298–3303.
88. Hammond, P. M. (1995) in *Microbial Diversity and Ecosystem Function*, eds. Allsopp, D., Colwell, R. R. & Hawksworth, D. L. (CAB International, Wallingford, Oxon, U.K.), pp. 29–71.
89. Wayne, L. G., Brenner, D. J., Colwell, R. R., Grimont, P. A. D., Kandler, O., Krichevsky, M. I., Moore, L. H., Moore, W. E. C., Murray, R. G. E., Stackebrandt, E., *et al.* (1987) *Int. J. System. Bacteriol.* **37**, 463–464.
90. Sibley, C. G., Ahlquist, J. E. & Monroe, B. L., Jr. (1988) *Auk* **105**, 409–423.
91. Sibley, C. G. & Ahlquist, J. E. (1987) *J. Mol. Evol.* **26**, 99–121.

REVIEW

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, re-supply 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.)

Microbial Ecology

organic matter (i.e., they are chemoautotrophs). Finally, in the absence of oxygen, a third set of opportunistic microbes uses NO_3^- and NO_2^- 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

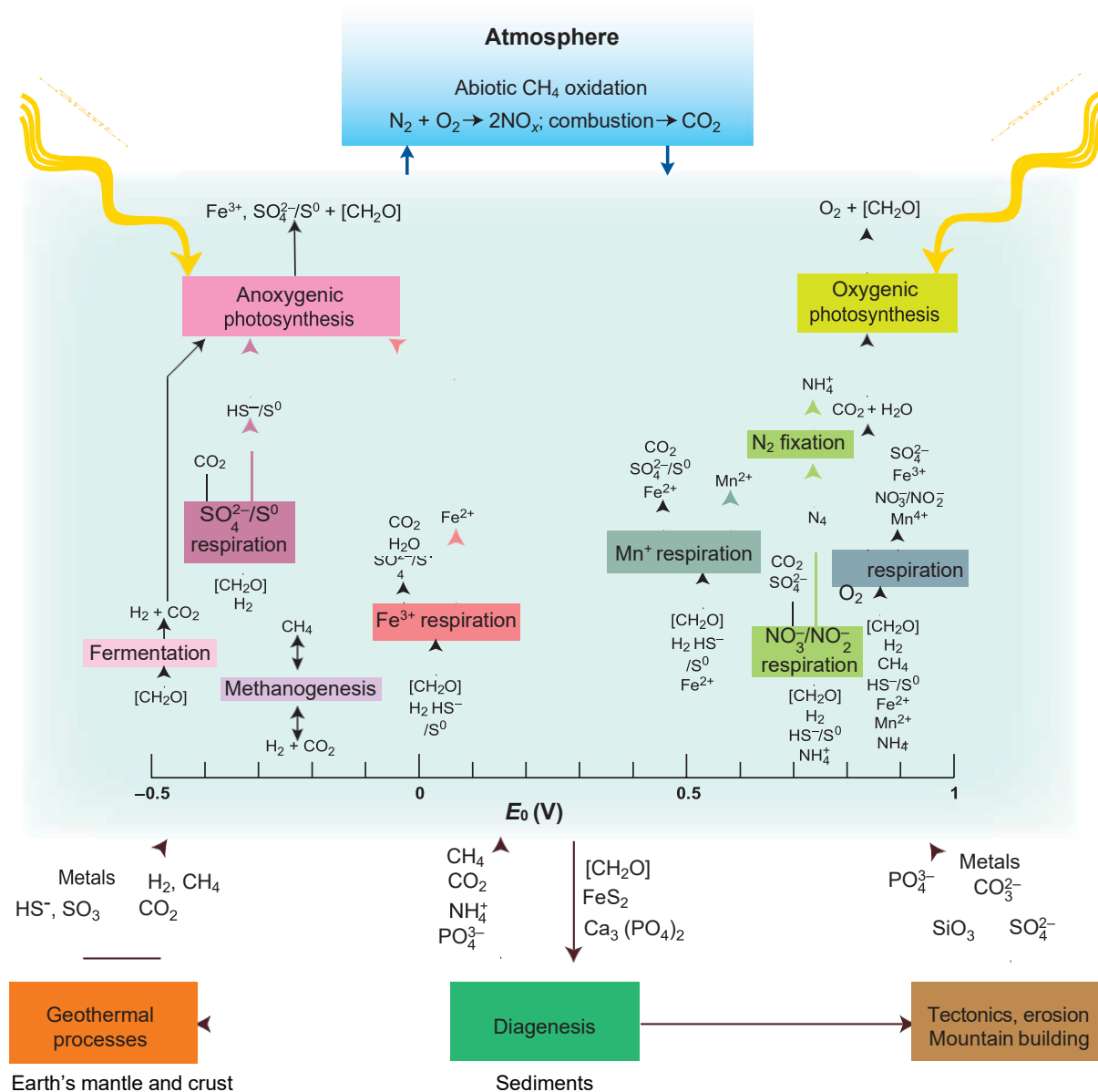


Fig. 1. A generalized biosphere model showing the basic inputs and outputs of energy and materials. Geochemical (abiotic) transformations are represented at the top (atmospheric) and bottom (tectonic and geothermal) compartments, while microbially driven biochemical processes are represented in the middle, 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.



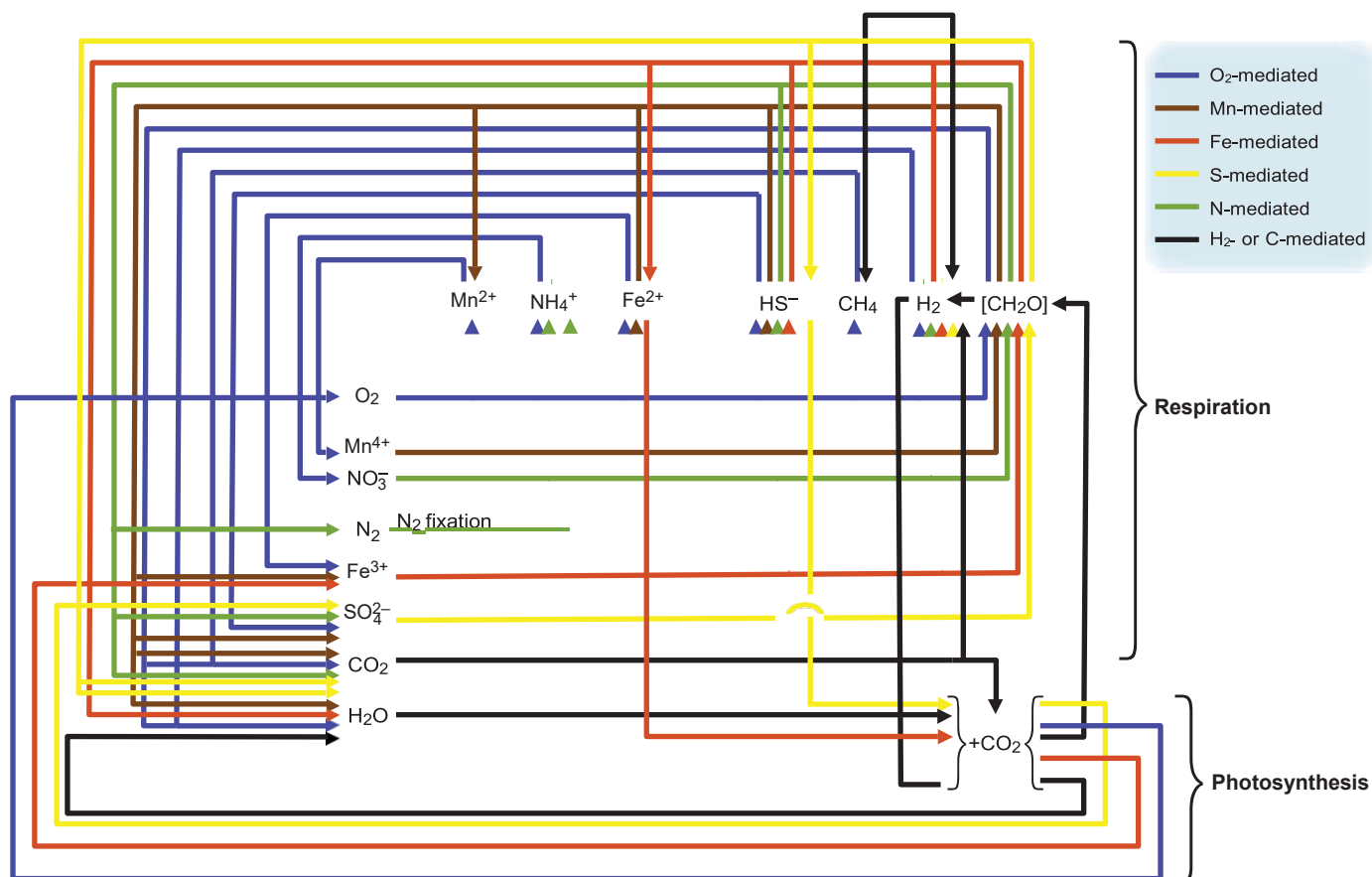


Fig. 2. A schematic diagram depicting a global, interconnected network of the biologically mediated cycles for hydrogen, carbon, nitrogen, oxygen, sulfur, and iron. A large portion of these microbially mediated processes are associated only with anaerobic habitats.

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, nonphototrophic 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 *d*-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 Archaeal 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

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 sub-optimally, 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-

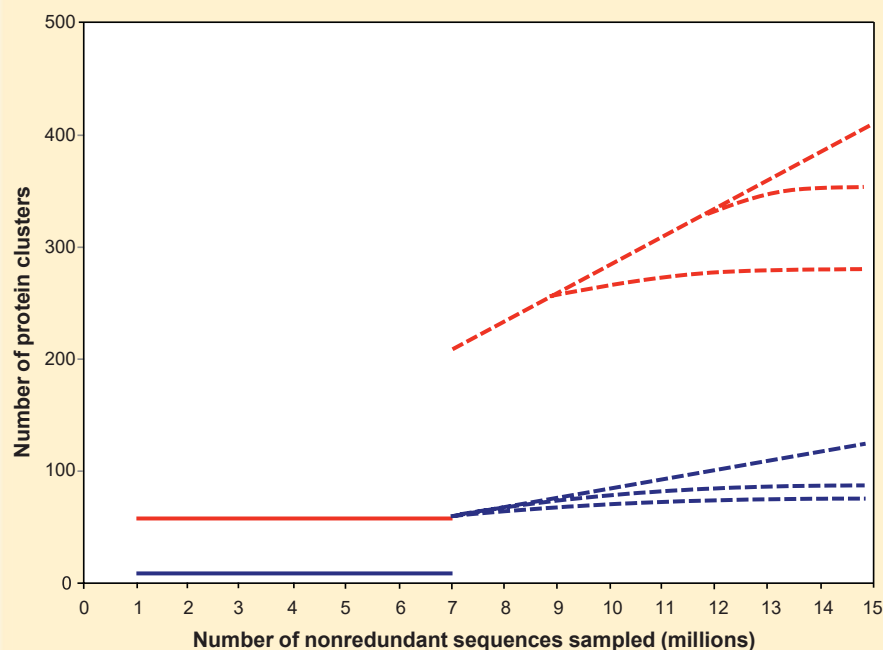


Fig. 3. Observed increases in new protein clusters with increasing sequence sampling [modified from Yooseph *et al.* (42)]. The number of new protein clusters discovered increases linearly with the number of nonredundant sequences sampled. We project hypothetical saturation profiles for the protein families. 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 Yooseph *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.

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

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

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

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

