

BIOGEOCHEMICAL SIGNATURES THROUGH TIME AS INFERRED FROM WHOLE MICROBIAL GENOMES

AUBREY L. ZERKLE, CHRISTOPHER H. HOUSE[†], and SUSAN L. BRANTLEY

Department of Geosciences, Pennsylvania State University, 212 Deike Building,
University Park, Pennsylvania, 16802

ABSTRACT. Throughout geologic time, a strong feedback has existed between the geosphere and the biosphere; therefore biological evolution and innovation can be linked to the evolution of ancient environments on Earth. Here we deduce geochemical signatures and phylogenetic relationships of prokaryotes from whole genome sequences and use this link to infer geochemical aspects of the biosphere through time. In particular, we have investigated two potential biosignatures for modern and ancient biochemistry: the magnitude of microbial carbon isotopic fractionation, and the use of metals in microbial cells.

The distribution of carbon fixation pathways on microbial phylogenies suggests that: (1) both low and moderate carbon isotope fractionation were established quickly in the early evolution of life; and (2) methanotrophic and ethanotrophic metabolisms capable of producing biomass with extreme ¹³C depletions are not primitive, but rather evolved after the major groups of Prokaryotes had already diverged. The universal importance of the TCA cycle, which results in low carbon isotopic fractionation, indicates it may have evolved especially early making it perhaps the most likely carbon fixation pathway for the last common ancestor (LCA). Low isotopic fractionation by the reductive TCA cycle can be considered consistent with carbon isotope ratios found in 3,800 million year old Isua sediments. Additionally, moderately fractionated biomass from up to 3,500 million years ago can now likely be attributed to carbon fixation by anoxygenic photoautotrophs using the reductive pentose phosphate cycle (Calvin cycle).

In addition to carbon, cells require a number of other elements that could potentially provide biosignatures, including bioactive trace metals. We calculated “model metallomes” for 52 prokaryotes based on the number of atoms of trace metals required to express one molecule of each metallo-enzyme coded for in the corresponding genomes. Our results suggest that the use of metals in prokaryotes as a group generally follows the hierarchy: Fe ≧ Zn > Mn ≧ Mo, Co, Cu ≧ Ni > W, V. However, model metallomes vary with metabolism, oxygen tolerance, optimum growth temperature, and phylogeny. The model metallome of methanogens shows a unique metal signature, suggesting that elevated requirements of nickel and tungsten might be translated to the expressed metallome providing a biosignature for methanogenesis. The model metallomes of diazotrophs and cyanobacteria do not show unique signatures; however changes in enzyme expression under some conditions could still translate into a metal biosignature in the expressed cellular metal content.

In a separate analysis, we made inferences about the timing of the evolution of individual metallo-enzymes based on their function and occurrence in modern organisms. The results suggest that fluctuations in the redox state of the Earth’s oceans and atmosphere have forced changes in the proportions of metals used in biology. In this model, biological use of copper and molybdenum has developed along with bioavailability; biological use of iron and manganese has developed counter to bioavailability; and biological use of zinc, cobalt, and nickel has not changed significantly through time. Using this technique, we estimated a model metallome for the LCA based on the metallo-enzymes we infer to have been present at that time. This metallome for the LCA differs greatly from one extrapolated from the distribution of model metallomes on microbial phylogenies, supporting the idea that gene loss, metal substitution, and lateral gene transfer have been important in shaping the enzymatic composition of extant organisms.

[†]Corresponding author: Christopher House, chouse@geosc.psu.edu

INTRODUCTION

Using Microbial Genomes to Infer Past Biochemistry

Microbial genomes provide information about metabolic capabilities and gene regulation in extant organisms, as well as molecular clues to the events leading to the evolution of these genes in the geologic past (Macalady and Banfield, 2003). In this context, modern prokaryotic genomes can provide a link between the geosphere and the biosphere through the evolutionary signatures they contain. Here we demonstrate this by using genomes to examine the modern phylogenetic distribution of two potentially important biochemical signatures, and interpreting their evolution and possible distribution in ancient life. We used the gene content of prokaryotic genomes to predict cellular metal requirements in modern organisms, and to determine how metal use and biological carbon isotope fractionation have evolved through geologic time. We first refined the topology of the tree of life using whole genome sequences of composite phylogenetic groups of prokaryotes. We then mapped biochemical characteristics onto the topology of this composite tree, and used the result as a foundation for discussion about microbial biochemistry on early earth. As part of this study, we have reviewed and updated the literature on biological carbon isotope fractionation by prokaryotes from carbon dioxide and methane, establishing a range of fractionation factors produced by specific carbon fixation pathways. We have also examined the cellular metallome inferred from gene content, and evaluated the use of a model metallome as a biosignature for specific microbial metabolisms. Ultimately we suggest that such analyses will be useful in interpreting biosignatures in both modern prokaryotic ecosystems and in the rock record.

Biological Carbon Isotope Fractionation

In recent decades, the biological fractionation of carbon isotopes has become an important means of recognizing and studying ancient life (for a review, see Schidlowski, 2001). Autotrophic microorganisms preferentially incorporate ^{12}C into their biomass to a varying degree based partly on their carbon fixation pathway. Therefore, specific carbon fixation pathways recognizably affect the isotopic signature of organic matter preserved in the geologic record.

Among the Bacteria and the Archaea, the two prokaryotic domains of life, there are four known CO_2 carbon fixation pathways: (1) the reductive tricarboxylic acid (TCA) cycle; (2) the 3-hydroxypropionate cycle (3-HP); (3) the reductive pentose phosphate cycle (or Calvin Cycle), using Rubisco (PP); and (4) the reductive acetyl-CoA pathway (AP). There may be additional presently unknown CO_2 carbon fixation pathways, operating in organisms such as the recently described '*Candidatus Chlorothrix halophila*', recovered from a hypersaline microbial mat (Klappenbach and Pierson, 2004).

Cell carbon can also be fixed from methane or other hydrocarbon gases. In the Bacteria, methanotrophy is accomplished using the enzyme methane monooxygenase in either the RuMp pathway (RuMP) or the Serine Pathway (S) (Hanson and Hanson, 1996). In the Archaea, it has been suggested that anaerobic methanotrophy (AOM) occurs using a modified methyl co-enzyme M reductase (Hallam and others, 2003; Krüger and others, 2003). The consumption of non-methane hydrocarbons, such as ethane, is known in the Actinobacteria and in the γ -Proteobacteria (Bokova, 1954; Davis and others, 1956; de Bont, 1976) but the process, ethanotrophy (E), remains considerably less studied than bacterial methanotrophy despite the potential of ethane as a global microbial substrate (D'Hondt and others, 2003).

Figure 1A shows the degree of carbon isotopic fractionation from either CO_2 or CH_4 to cellular biomass observed in organisms using different carbon fixation pathways (data found in tables 1A and 1B; adopted and updated from House and others,

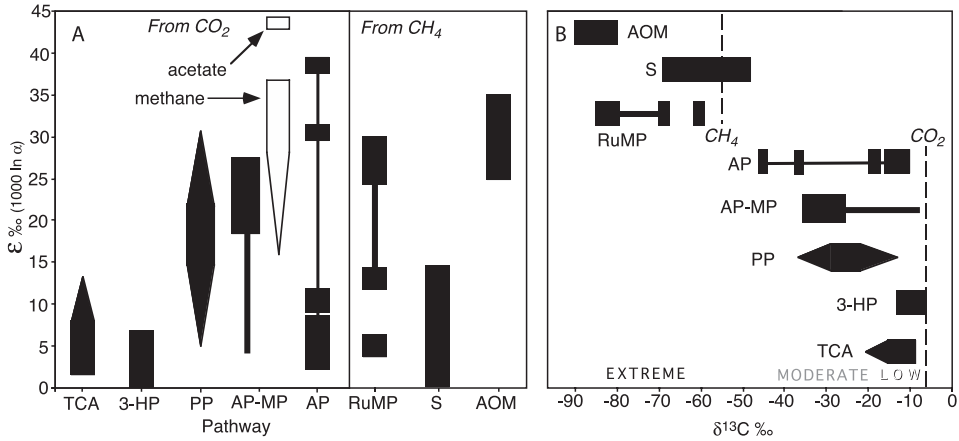


Fig. 1. (A) A generalized summary of published carbon-isotopic fractionation factors from carbon dioxide or methane to biomass by prokaryotes grouped by their known or suspected carbon fixation pathway (modified from House and others, 2003a, with data from tables 1A and 1B). Note that ϵ only defines a relative value of fractionation produced by the specific pathway, and that the actual carbon isotopic compositions produced will vary according to the isotopic composition of the substrate used. The carbon fixation pathways from carbon dioxide are: the reductive TCA cycle (TCA), the 3-hydroxypropionate cycle (3-HP), the reductive pentose phosphate cycle (PP), and the acetyl-CoA pathway in species that form a metabolic product from carbon dioxide such as methane or acetate (AP-MP), or in species that do not (AP). The shape of the bars indicates the estimated relative importance of that pathway in producing a particular fractionation factor. The changing carbon isotopic fractionation that has been observed for methanogens with growth phase (listed in AP-MP) is shown as a thin solid line extending up to the values expected for a culture at stationary growth. Fractionation factors observed in the metabolic products produced (acetate and methane) are shown as white boxes. Other work has shown that larger fractionations to methane can be obtained during methanogenesis in open systems (Botz and others, 1996; Valentine and others, 2004). The published fractionation factors for AP have been connected by a thin solid line; while these organisms all use the acetyl-CoA pathway, they demonstrate a wide range of isotopic fractionations. The carbon fixation pathways from methane are bacterial methanotrophy using the RuMP pathway (RuMP) or the Serine Pathway (S), and anaerobic methanotrophy (AOM). For the RuMP pathway, the different fractionation factors associated with soluble versus particulate methane monooxygenase is connected by a solid line because both forms of the enzyme is found in some species causing them to show a reduction in fractionation at the initiation of stationary growth. (B) The predicted carbon isotopic composition produced by natural microbial cells using each of the pathways shown, assuming fixation of CO₂ with a typical carbon isotopic composition of about -7 per mil, or CH₄ with a composition of -55 per mil. The labels “extreme”, “moderate”, and “low” refer to terminology used in fig. 3 and the corresponding discussion.

2003a; also reviewed by Knoll and Canfield, 1998; van der Meer and others, 2001a). Fractionation factors in figure 1A are given in ϵ (Hayes, 2002):

$$\epsilon = \ln \alpha_{A-B} \times 10^3 \quad (1)$$

where A is the substrate being used, B is the resulting biomass, and α is the fractionation factor, the ratio of the rate of carbon fixation for ¹²C to the rate of carbon fixation for ¹³C, calculated as

$$\alpha_{A-B} = \frac{\delta^{13}C_A + 10^3}{\delta^{13}C_B + 10^3} \quad (2)$$

Note that ϵ only defines a relative value of fractionation produced by the specific pathway, and that the actual carbon isotopic compositions produced will vary according to the isotopic composition of the initial substrate used.

The results in figure 1A indicate that the reductive pentose phosphate cycle, acetyl-CoA pathway, and some methanotrophic pathways can yield larger isotopic fractionations than the reductive TCA cycle and 3-hydroxypropionate cycle (Preuß

TABLE 1A

Published results of growth experiments estimating carbon-isotopic fractionation from carbon dioxide to cell biomass, permil. The carbon fixation pathways are those listed in figure 1. Where available, isotopic fractionation to corresponding metabolic products are given in parentheses (MP; methane in most cases, except acetate for *A. woodii*). Updated from House and others, 2003a.

| TAXON | Pathway | T (°C) | ϵ (‰) from CO ₂ to cells (MP) | Ref | TAXON | Pathway | T (°C) | ϵ (‰) from CO ₂ to cells (MP) | Ref |
|--------------------------------------|---------|--------|---|-----|------------------------------------|---------|--------|--|-----|
| <i>Acetobacterium woodii</i> | AP-MP | 28 | 22.2 (43.6), 29.0 (ND), 15.8 (43.4) | 1 | <i>Methanococcus jannaschii</i> | AP-MP | 85 | 6.2 (25.0), 10.7 (19.6), 13.7 (17.7), 17.7 (19.6) | 2 |
| <i>Acidianus brierleyi</i> | 3-H | 65 | 3.6 | 2 | <i>M. thermolithotrophicus</i> | AP-MP | 41 | 7.4 (26.2) | 2 |
| <i>Agmenellum quadruplicatum</i> | PP | 39 | 15.9 | 3 | | | 45 | 20.5 (21.1), 13.4 (23.0), 4.8 (16.5) | 2 |
| | | 39 | 22.2, 23.9, 19.6 | 4 | | | 51 | 13.5 (27.0) | 2 |
| <i>Alkalicoccus eutrophus</i> | PP | 28 | 28.2 | 1 | | | 60 | 19.7 (26.4) | 2 |
| <i>Ammonifex degensii</i> | ? | 70 | 4.1, 4.2 | 2 | | | 65 | 22.7 (29.0), 25.8 (29.6), 26.0 (30.8), 26.7 (29.5), 24.8 (29.8), 22.0 (23.2), 19.4 (24.9), 8.7 | 2 |
| <i>Anacystis nidulans</i> | PP | 39 | 13.1, 18.0 | 3 | | | | | |
| | | 33 | 16.0, 21.5, 18.0 | 5 | | | | | |
| <i>Aquifex aeolicus</i> | TCA | 85 | 5.4, 5.4 | 2 | | | | | |
| <i>Archaeoglobus fulgidus</i> | AP | 85 | 2.7, 5.8 | 2 | | | | | |
| <i>Archaeoglobus lithotrophicus</i> | AP | 85 | 8.0 | 2 | <i>Methanopyrus kandleri</i> | AP-MP | 70 | 24.9 (27.9) | 2 |
| <i>Candidatus "Brocadia</i> | ? | 35 | 18.7, 17.0, 22.3, 21.2, 15.8 | 6 | | | 100 | 20.3 (30.0), 12.7 (27.6) | 2 |
| anamoxidans" strain Delft | | | | | | | | 17.7 | |
| <i>Chlorella sorokiniana</i> | PP | 39 | 22.6 | 4 | <i>Methanosarcina barkeri</i> | AP-MP | 37 | 19.5 (17.2), 7.9 | 2 |
| <i>Chloroflexus aurantiacus</i> | 3-H | 55 | 7.6 | 7 | <i>Methanothermus fervidus</i> | AP-MP | 85 | 13.1 (30.6) | 2 |
| <i>Chromatium tepidum</i> | PP | 50 | 20.5 | 8 | <i>Microcoleus chthonoplastes</i> | PP | 39 | 17.1 | 4 |
| <i>Coccochloris eiebens</i> | PP | 39 | 12.3, 16.0 | 3 | <i>Osillatoria williamsii</i> | PP | 39 | 5.0 | 3 |
| <i>Desulfobacter hydrogenophilus</i> | TCA | 28 | 10.0, 8.9, 14.2 | 1 | | | 39 | 17.3 | 4 |
| | | 30 | 13.7 | 9 | <i>Pyrobaculum aerophilum</i> | TCA | 100 | 2.9 | 2 |
| | | 30 | 15.4 | 10 | <i>Pyrodicticum occultum</i> | ? | 102 | 2.3 | 2 |
| <i>D. autotrophicum</i> | AP | 28 | 39.3, 38.2 | 1 | <i>Pyrobolus fumarii</i> | ? | 105 | 3.8 | 2 |
| | | 30 | 10.0 | 9 | <i>Rhodospseudomonas capsulata</i> | PP | | 10.8 | 14 |

TABLE 1A
(continued)

| TAXON | Pathway | T (°C) | ϵ (%) from CO ₂ to cells (MP) | Ref | TAXON | Pathway | T (°C) | ϵ (%) from CO ₂ to cells (MP) | R ef |
|-------------------------------------|---------|--------|---|-----|-----------------------------------|---------|--------|---|------|
| <i>Desulfotomaculum acetoxidans</i> | AP | 30 | 9.9 | 10 | <i>Rhodospirillum rubrum</i> | PP | 20 | 21.1 | 15 |
| | | 30 | 30.5 | 9 | | | | | |
| | | 30 | 29.7 | 10 | | | | | |
| <i>Ferroglobus placidus</i> | AP | 85 | 3.5 | 2 | <i>Schizothrix calcicola</i> | PP | 39 | 12.7 | 14 |
| <i>Hydrogenobacter thermophilus</i> | TCA | 70 | 5.5 | 2 | <i>Sulfolobus solfataricus</i> | 3-H | 85 | 0.2 | 2 |
| <i>Metallosphaera sedula</i> | 3-H | 65 | -3.0 | 11 | <i>Synechococcus</i> | PP | 57 | 22.5 | 5 |
| | | 65 | 3.1 | 2 | | | | 15.5, 16.4, 18.6, 18.6, | 16 |
| <i>Methanobacterium</i> sp. | AP-MP | 37 | 25.1 (36.4), | 12 | | | | 17.2, 17.1, 15.6, | |
| | | | 23.2 (34.9) | | | | | 16.9, 18.9, 18.1 | |
| | | 46 | 18.8 (33.7), | 12 | <i>Synechococcus lividus</i> | PP | 47 | 12.8 | 4 |
| | | | 24.5 (35.7) | | | | | 11.1 | |
| <i>Methanobacterium thermoauto.</i> | AP-MP | 65 | 19.1 (34.1) | 13 | <i>Thermocrinis ruber</i> | TCA | 85 | 3.3 | 17 |
| | | 65 | 15.0 (24.6) | 2 | <i>Thermoproteus neutrophilus</i> | TCA | 85 | 8.7 | 1 |
| <i>Methanococcus igneus</i> | AP-MP | 85 | 20.2 (28.4) | 2 | | | | 2.0 | 2 |
| | | | | | <i>Thiobacillus novellus</i> | PP | 30 | 5.1 | 2 |

References: 1 = Preuß and others, 1989; 2 = House and others, 2003a; 3 = Calder and Parker, 1973; 4 = Pardue and others, 1976; 5 = Mizutani and Wada, 1982; 6 = Schouten and others, 2004; 7 = van der Meer and others, 2001b; 8 = Madigan and others, 1989; 9 = Londry and Des Marais, 2003; 10 = Londry and others, 2004; 11 = van der Meer and others, 2001a; 12 = Belyaev and others, 1983; 13 = Fuchs and others, 1979; 14 = Quandt and others, 1977; 15 = Sirevåg and others, 1977; 16 = Popp and others, 1998; 17 = Jahnke and others, 2001

TABLE 1B

Published results of growth experiments estimating carbon-isotopic fractionation from dissolved inorganic carbon (DIC) or methane to cell biomass, per mil. The carbon fixation pathways are those listed in figure 1. Updated from House and others, 2003a.

| TAXON | Pathway | T (°C) | ϵ (‰) from DIC to Cells | Ref | TAXON | Pathway | T (°C) | ϵ (‰) from CH ₄ to Cells | Ref |
|--|---------|--------|----------------------------------|-----|-------------------------------------|---------|-----------|---|-----|
| <i>Candidatus</i> "Brocadia anammoxidans" strain Delft | ? | 35 | 27.1, 25.3, 29.5, 24.2, 30.6 | 6 | ANME-1 | AOM | ~25 to 35 | 23 | |
| <i>Chlamydomonas reinhardtii</i> | PP | 20 | 36.8 | 15 | ANME-2 | AOM | ~30 to 35 | 23, 24 | |
| <i>Chlorobium limicola</i> | TCA | 30 | 9.5 | 14 | CEL 1923 | RuMP | 31.2 | 25 | |
| <i>Chlorobium phaeovibrioides</i> | TCA | 30 | 10.4, 9.5 | 14 | | | 25.7 | | |
| <i>Chlorobium thiosulfatophilum</i> | TCA | 20 | 20.1 | 15 | | | 24.1 | | |
| <i>Chlorobium vibrioforme</i> | TCA | 30 | 12.2, 10.8, 10.7 | 14 | | | 25 | | |
| | | 28 | 11.4 | 18 | | | 30 | | |
| <i>Chloroflexus aurantiacus</i> | 3-H | 55 | 13.7 | 19 | <i>Methylococcus capsulatus</i> | RuMP | 35 | 22.1 | |
| | | 55 | 12.7 | 18 | | | 37 | 31.3, 28.9, 24.4, 19.5, 19.6, 17.9, 17.6, 19.3, 15.8, 18.6, 14.5, 3.8, 28.4, 27.6, 18.2 | |
| <i>Chromatium strain D</i> | PP | 20 | 31.0 | 15 | | | | | |
| <i>Chromatium vinosum</i> | PP | 35 | 32.7 | 20 | | | 37 | 14.4, 12.4, 13.9, 13.6, 12.8, 24.1, 26.0 | |
| | | 30 | 27.3 | 14 | | | | | |
| <i>Metallosphaera sedula</i> | 3-H | 65 | 2.0 | 11 | | | | | |
| <i>Nitrosomonas europaea</i> | PP | | 13.8, 13.2 | 21 | | | | | |
| <i>Oscillochloris trichoides</i> | PP | 28 | 20.1 | 18 | <i>Methylobionas methanica</i> | RuMP | 29.3 | 26 | |
| <i>Thiobacillus neapolitanus</i> | PP | | 25.8 | 22 | <i>Methylophilus methylotrophus</i> | RuMP | 5.6 | 25 | |
| <i>Thiocapsa roseopersicina</i> | PP | 28 | 22.0 | 18 | <i>Methylostinus trichosporium</i> | S | 9.9 | 26 | |
| <i>Thiomicrospira sp. L-12</i> | PP | | 25.5 | 22 | | | 30 | 5.2, -3.7, 7.6, -0.6, 9.2, -7.3, 14.7, -7.8 | |
| <i>Thiomicrospira crunogena</i> | PP | | 24.5, 23.3 | 21 | | | | | |

References: as in table 1A; additionally 18 = Ivanovsky and others, 1999; 19 = Holo and Sirevåg, 1986; 20 = Wong and others, 1975; 21 = Mirkin and Ruby, 1991; 22 = Ruby and others, 1987; 23 = Orphan and others, 2002; 24 = Orphan and others, 1999; 25 = Jahnke and others, 1999; 26 = Summons and others, 1994

and others, 1989; House and others, 2003a; Londry and Des Marais, 2003). House and others (2003a) also showed that biomass produced by acetyl-CoA pathway utilizing microbes can vary over a remarkably wide range, with ϵ values of 2.7 to 8.0 permil for the Archaeoglobales and ϵ values of 4.8 to 26.7 permil for the methanogens. The magnitude of carbon isotopic fractionation observed in methanogens is dependent on the growth status of the culture, with isotopic fractionation increasing as the cultures proceeded toward stationary growth phase. Furthermore, Londry and Des Marais (2003) have shown that different sulfate reducers using the acetyl-CoA pathway can demonstrate widely different fractionations (ϵ values of both 10.0 and 30.5‰). The fractionation associated with bacterial methanotrophy by RuMP has been shown to be up to about 30 permil for the particulate methane monooxygenase, but lower for the soluble methane monooxygenase (Summons and others, 1994; Jahnke and others, 1999). Finally, the fractionation associated with anaerobic methane oxidation has been estimated to be around 32 permil based on studies of natural cell carbon from seep environments (Orphan and others, 2001a; Orphan and others, 2002).

Figure 1B shows the predicted carbon isotopic composition produced by natural microbial cells using each of the pathways shown, assuming fixation of CO_2 with a typical carbon isotopic composition of about -7 permil, or CH_4 with a composition of -55 permil. The global range of carbon isotopic compositions for methane is between -50 to -110 permil (Whiticar, 1999), but is typically heavier (about -55 ‰) in marine seeps subject to microbial methane oxidation (Whiticar, 1999; Orphan and others, 2004). This figure provides only an approximate guide to the expected carbon isotopic signatures that each metabolic pathway should produce in a natural setting, because the actual isotopic composition of either CO_2 or CH_4 can be highly variable in nature. Nevertheless, there is expected to be demonstrably different values of carbon isotopic composition found in natural microbial biomass. The most ^{13}C -depleted carbon is produced by cells that have grown on natural biogenic methane. Methanotrophic cells using either an aerobic pathway (RuMP) or an anaerobic pathway (AOM) typically have cellular biomass that is around -85 permil (Orphan and others, 2001b, 2002) due to both a ^{13}C -depleted methane source and isotopic fractionation during carbon fixation. Of the CO_2 fixation pathways, the acetyl-Co A pathway can produce the most ^{13}C -depleted biomass, but in other cases produces significantly less ^{13}C -depleted biomass (from $\delta^{13}\text{C}$ of about -45 ‰ to about -10 ‰). In methanogens and acetogens, this pathway produces biomass that is highly ^{13}C -depleted ($\delta^{13}\text{C}$ about -30 ‰). The same level of ^{13}C -depletion is found for natural microbial cells using the reductive pentose phosphate cycle with type I rubisco ($\delta^{13}\text{C}$ around -30 ‰), and somewhat less ^{13}C -depletion is found for cells using the reductive pentose phosphate cycle with type II rubisco ($\delta^{13}\text{C}$ around -20 ‰; Roeske and O'Leary, 1984, 1985; Robinson and Cavanaugh, 1995). Finally, the two remaining carbon fixation pathways, the 3-hydroxypropionate cycle and the reductive TCA cycle, produce the least ^{13}C -depleted biomass (with $\delta^{13}\text{C}$ values expected between 0 ‰ to -20 ‰).

The Microbial Metallome

In addition to carbon and other traditional macro-nutrients such as nitrogen and phosphorous, organisms require more than 20 elements for cellular growth. In fact, a biological cell can be compositionally described by: (1) its genome, which contains a blueprint of all molecules required to carry out cellular processes (DNA); (2) its proteome, the individual proteins encoded for by the genome and expressed; or (3) its metallome, the twenty-some elements (including carbon, nitrogen, phosphorous, as well as bioactive trace metals) that make up both DNA and proteins and are involved in cellular reactions (Frausto da Silva and Williams, 2001). Bioactive metals are essential nutrients that function as structural elements and catalytic centers in metalloproteins and metal-activated enzymes involved in virtually all cell functions, including DNA and

RNA synthesis, respiration and photosynthesis, electron transport, and cellular detoxification (Holm and others, 1996; Frausto da Silva and Williams, 2001). Iron, for example, is almost universally required for life, forming iron-sulfur proteins involved in mitochondrial electron transport, and important heme proteins used for electron transfer, oxidase formation, storage, and transport. Other metals have more limited biological uses associated with specific microbial physiologies. Molybdenum, for example, is an important enzyme in the nitrogen cycle, used in nitrate reductase and the nitrogenase enzyme necessary for nitrogen fixation (Burgess and Lowe, 1996; Eady, 1996; Hille, 1996; Kisker and others, 1997); nickel is found in enzymes catalyzing carbon reduction reactions of acetogenic and methanogenic organisms (Hausinger, 1987; Ragsdale and Kumar, 1996; Eitinger and Friedrich, 1997; Taha and others, 1997); and tungsten is most commonly used by hyperthermophiles and methanogenic archaea (Johnson and others, 1996; L'vov and others, 2002).

Conflicting ideas exist as to the extent of variance in the metallome of prokaryotic organisms. Previous work indicates that *E. coli* exert tight genetic control over cellular metal concentrations, even for relatively low-toxicity metals such as zinc (Outten and O'Halloran, 2001). These researchers conclude that heterotrophic microbial cells cultured in the laboratory contain a background metal content, or "metal quota", that is conserved regardless of the metal content of the media. Other studies (for example Raven, 1988) suggest that organisms with specialized metabolic functions, such as nitrogen fixation (diazotrophy), may require up to 100 times more of a specific trace metal to sustain growth under nitrogen-limited conditions. Specialized prokaryotic trace metal requirements in excess of background cellular metal content could potentially provide a useful fingerprint of physiological function in poorly characterized microbial communities in the environment.

In addition to providing a possible link between cellular chemistry and physiology, the metallome also offers the most direct chemical link between biological systems and the environment. An element that is widely used by biological systems must be 'biologically available', or present in an easily extractable form, from the atmosphere or from solutions. Conditions such as temperature and the oxidation state of the atmosphere and oceans have fluctuated through geologic time, considerably affecting the bioavailability of redox sensitive elements. Some researchers suggest that decreased organic matter production and carbon isotope fractionation during the mid-Proterozoic could have been due to metal limitation of nitrogen fixation and primary productivity in a sulfidic ocean (Canfield, 1998; Anbar and Knoll, 2002). The trace metal preferences and sensitivities of cyanobacteria could also reflect evolution in a sulfidic environment (Saito and others, 2003). Modern biological systems have evolved sophisticated methods of acquiring essential metals under metal-limited conditions, and can significantly alter geochemical cycling in natural environments. For example, in terrestrial settings bacteria can enhance weathering rates through the production of chelators to promote solubilization of metals from minerals in soils (Kalinowski and others, 2000; Liermann and others, 2000; Brantley and others, 2001). In response to low concentrations of some essential metals in the surface of modern oceans due to phytoplankton depletion (Bruland, 1980, 1989; Martin and others, 1989; Rue and Bruland, 1995), various microorganisms release strong complexing agents and catalyze redox reactions that modify the bioavailability of these metals and promote their rapid cycling in the upper water column (for a review see Morel and Price, 2003).

We examine the trace metal content of the modern prokaryotic metallome predicted by the presence or absence of genes coding for metallo-enzymes in whole genome sequences. We utilize these data to determine if prokaryotic metallomes could provide physiological biosignatures, and to clarify the evolution of biological chemistry

and its inter-dependence with the Earth surface environment. Due to the nearly universal biological role of iron, it is assumed that all organisms require cellular iron concentrations far in excess of other redox-sensitive metals. For this reason, we focused only on the predicted metallome contribution of non-iron trace metals (Zn, Mn, Cu, Mo, Co, Ni, W, and V) from prokaryotic genomes. Iron-containing enzymes were considered only if they contained an additional trace metal. In a second analysis, we estimated the approximate timing of evolution of metallo-enzymes (including those containing iron) from their functions and occurrences in modern organisms. This view of metallome evolution is then compared to modeled changes in the bioavailability of these metals through time from the literature (Saito and others, 2003).

MATERIALS AND METHODS

Estimation of a Consensus Tree of Life

We estimated the topology of the tree of life from the combined gene content of taxonomic groups of prokaryotes. For this, we downloaded data matrices containing the presence or absence of homologous gene families in individual prokaryotes, as calculated in House and others (2003b). We then combined the data matrices for individual organisms in each group to produce a representative data matrix scoring the presence or absence of gene families in a specified percentage of individual organisms in that group, including gene families present in greater than 1, 10, 15, 20, 25, 50, and 99 percent of the organisms in the group (combined matrices available at <http://www.geosc.psu.edu/~chouse/ajs1.html>). We constructed phylogenetic trees based on the presence and absence of informative gene families in composite genomes, as previously described for individual organisms (Fitz-Gibbon and House, 1999; House and others, 2003b). Parsimony and distance analyses were performed on the combined matrices using PAUP version 4.0b (Swofford, 2002); compatibility and threshold parsimony analyses were applied using the Phylip software package (Felsenstein, 1993). We compared these results to published trees based on rRNA (Fox and others, 1980; Woese, 1987), concatenated protein datasets (Hansmann and Martin, 2000; Brown and others, 2001; Brochier and others, 2002; Matte-Tailiez and others, 2002), and other gene content methods (Gerstein, 1998; Gerstein and Hedgyi, 1998; Snel and others, 1999; Tekaiia and others, 1999; Lin and Gerstein, 2000; Wolf and others, 2001; Bansal and Meyer, 2002; Clarke and others, 2002; Korbel and others, 2002; Li and others, 2002; Blank, 2004; Yang and others, 2005), and produced a consensus microbial tree of life based on this comparison. This tree does not represent a strict consensus, but simply our best estimate of phylogenetic relationships based on these disparate studies.

Predicted Carbon Isotope Fractionation

We produced a compilation of experimentally determined carbon isotope fractionation values as a function of carbon fixation pathways from data compiled in House and others, 2003a (data from Calder and Parker, 1973; Wong and others, 1975; Pardue and others, 1976; Quandt and others, 1977; Sirevåg and others, 1977; Fuchs and others, 1979; Mizutani and Wada, 1982; Belyaev and others, 1983; Holo and Sirevåg, 1986; Ruby and others, 1987; Preuß and others, 1989; Mirkin and Ruby, 1991; Popp and others, 1998; Jahnke and others, 2001; House and others, 2003a) and updated with additional published literature (Madigan and others, 1989; Summons and others, 1994; Jahnke and others, 1999; van der Meer and others, 2001a, 2001b; Orphan and others, 2001a, 2002; Londry and Des Marais, 2003; Londry and others, 2004; tables 1A and 1B; fig. 1A). We superimposed the results onto our consensus microbial tree of life based on the phylogenetic distribution of carbon fixation pathways (Andreesen and Gottschalk, 1969; Sirevag, 1974; Fuchs and Stupperich, 1978; Fuchs and others, 1980;

Schauder and others, 1987, 1989; Holo, 1989; Altekar and Rajagopalan, 1990; Windhovel and Bowien, 1990; Meijer and others, 1991; Beh and others, 1993; Strauss and Fuchs, 1993; Kandler, 1994; Vorholt and others, 1995; Delwiche and Palmer, 1996; Ishii and others, 1996; Wahlund and Tabita, 1997; Ishii, 1998; Shively and others, 1998; Menendez and others, 1999; Watson and others, 1999; Wirsen and others, 2002; Hügler and others, 2003; Macalady and Banfield, 2003). We then used parsimony to extrapolate carbon isotopic fractionation at various nodes on the tree, including the last common ancestor (LCA).

Computing the Model Metallome

We compiled a list of known metallo-enzymes utilizing the bioactive metals iron, zinc, copper, molybdenum, manganese, cobalt, nickel, tungsten, and vanadium, and the number of atoms of metal required per molecule of enzyme, when known (data tables available at <http://www.geosc.psu.edu/~chouse/ajs1.html>; Burgess and Lowe, 1996; Dismukes, 1996; Hille, 1996; Holm and others, 1996; Johnson and others, 1996; Lipscomb and Sträter, 1996; Ragsdale and Kumar, 1996; Berman and others, 2000; Frausto da Silva and Williams, 2001; L'vov and others, 2002; and references therein). We created a mysql database of genome content for the complete genomes of 52 prokaryotes from data downloaded from the NCBI site (www.ncbi.nih.gov or <ftp.ncbi.nih.gov>), which contains reannotated microbial genomes for maximum similarity. For each gene in a genome, the NCBI database contains information including, but not limited to, fields for the gene name, the Clusters of Orthologous Groups (COG) code (Tatusov and others, 1997), and the COG product. We then used multiple arrangements of the common names and synonyms of each of the metal-containing enzymes we identified to search within our mysql genome database for the presence of a gene or genes encoding each metallo-enzyme. The presence of genes coding for metal-containing enzymes in each genome was then used to compute a “model metallome” for each organism.

In computing the model metallome, we made a series of assumptions. First, for each gene known to code for a metallo-enzyme, we assumed that one enzyme molecule would be expressed. We then defined the model metallome, T_{Me} , for this one enzyme per gene system by summing all the metal atoms in the expressed enzymes, such that if Me_i is the total number of atoms of any individual metal, i , in that model metallome, then

$$T_{Me} = \sum_i Me_i \quad (3)$$

We calculated the fractional metal contribution of each metal i to the model metallome (F_i), by the following equation:

$$F_i = \frac{Me_i}{T_{Me}} \quad (4)$$

Due to the nearly ubiquitous use of iron in biological systems, iron-containing enzymes were considered only if they contained an additional metal cofactor. In that case, only the number of non-Fe atoms appeared in the T_{Me} summation. Thus, strictly, F_i is the fractional metal contribution to a model metallome where Fe atoms have been excluded. Values of Me_i , T_{Me} , and F_i for individual metals are shown for each organism in table 2.

In any natural system, the microbial metallome can be altered either by changing the genome through the addition or loss of genes encoding for a metallo-enzyme, or by changing the proportion of genes expressed by the organism. In order to make generalizations about T_{Me} across a wide range of microbial cellular functions, we

assumed that all enzymes encoded in the genome are produced in equal proportions within the cell at all times. Thus, the “model metallome” represents the metals in all metallo-enzymes that can be produced by an organism (excluding Fe-only enzymes). At any given time under any given environmental condition, the actual expressed cellular metallome in the organism may differ substantially from the theoretical “model metallome” we calculate, due to changes in gene expression.

We then sought to look at how the model metallome varied across the consensus phylogenetic tree. We tested sets of individual F_i values of organisms within phylogenetic groups for similarity using the chi-square (Pearson goodness of fit) test for general association of variables. The test was used to determine whether or not the organisms have statistically similar model metallomes within the groups, and statistically different model metallomes between groups. This particular statistical analysis is generally weakened by the occurrence of counts of less than five for the variable in question; therefore tests performed using each individual F_i count were checked against tests combining F_i counts less than five, commonly F_{Ni} , F_V , and F_W . The model metallome for phylogenetic groups was statistically significant in all but two cases: the model metallomes for the firmicutes fit into three statistically significant categories that were unrelated to previously recognized phylogenetic groupings, and the model metallome of methanogenic archaea fit into two statistically significant groups, *Methanosarcina* sp. and chemoautotrophic methanogens (excluding *M. thermoautotrophicum*).

We extrapolated these model metallomes on a consensus tree representing our analysis of combined gene families considered along with the most widely accepted microbial tree topologies (see discussion below). We calculated the extrapolated metallomes for each node by computing the mean F_i for each metal of all branches diverging from the node. For example (see fig. 7A), the F_i values assigned to the branching point of the firmicutes and actinobacteria is calculated as an average of the F_i values for the three groups of firmicutes and the F_i values for the actinobacteria. In turn, the F_i values of earlier nodes on the tree were calculated based on the values determined for the lineages that diverge from them.

COMPOSITE ORGANISM PHYLOGENY AND A CONSENSUS TREE OF LIFE

We estimated the topology of the tree of life from the combined gene content of taxonomic groups of prokaryotes. Trees based on gene families present in ≥ 20 percent of individual genomes in each taxonomic group are shown in figure 2A (maximum parsimony) and in figure 2B (compatibility). We chose to present the trees based on twenty percent because using gene families present in ≥ 25 percent of genomes in a group resulted in composite genomes with very few gene families unique to each group, and thus produced primarily unresolved trees. Trees produced with gene families present in ≥ 15 percent and ≥ 20 percent of genomes in a group were nearly identical. Additionally, a cutoff point of 20 percent produced the most consistent trees between the various tree-building methods we employed.

In general, different tree-building methods have different strengths and weaknesses, leading to differing tree topologies. Therefore, based on comparisons of our results with trees in the literature constructed from ribosomal RNA, the concatenation of large protein-data sets, and other gene content methods (see references above), we developed a tree best representing the present consensus on microbial relationships (fig. 2C). For the consensus tree, the two groups of gram positive bacteria (the firmicutes and the actinobacteria) were united based on their overall similar cell structure, in spite of the limited molecular support for this grouping (Brown and others, 2001; Wolf and others, 2001; Fu and Fu-Liu, 2002; Blank, 2004). Our resulting tree of life, shown in figure 2C, is broadly similar to that based on rRNA (for example: Fox and others, 1980; Woese, 1987) with the following exceptions: (1) the methanogenic archaea are united together based on whole genome analysis (Slesarev and

TABLE 2

The organisms considered in this study, their groupings in discussed analyses (including phylogenetic groups, optimum growth temperature, oxygen tolerance, and any specialized metabolisms considered), the number of atoms of each individual metal in a given model metallome (Me_i), the total model metallome (T_{Me}), and the fractional contribution of each metal to the model metallome (F_p , in %). Abbreviations used are defined below.

| Microorganism | Group | Opt | T | O ₂ pref | Met | Me _{Zn} | F _{Zn} | Me _{Mn} | F _{Mn} | Me _{Cu} | F _{Cu} | Me _{Ni} | F _{Ni} | Me _{Co} | F _{Co} | Me _{Mg} | F _{Mg} | Me _W | F _W | Me _V | F _V | T _{Me} |
|---|--------------------------------|-------|------|---------------------|--------|------------------|-----------------|------------------|-----------------|------------------|-----------------|------------------|-----------------|------------------|-----------------|------------------|-----------------|-----------------|----------------|-----------------|----------------|-----------------|
| <i>Aeropyrum pernix</i> K1 | Crenarchaeota | Hyp | A | | A | 12 | 35% | 7 | 21% | 3 | 9% | 0 | 0% | 5 | 15% | 6 | 18% | 1 | 3% | 0 | 0% | 34 |
| <i>Agrobacterium tumefaciens</i> C58 | α-Proteobacteria | Mes | A | | Nf | 26 | 38% | 18 | 26% | 7 | 10% | 2 | 3% | 7 | 10% | 6 | 9% | 2 | 3% | 1 | 1% | 69 |
| <i>Aquifex aeolicus</i> VF5 | | Hyp | M | | M | 15 | 33% | 12 | 26% | 7 | 15% | 1 | 2% | 4 | 9% | 6 | 13% | 1 | 2% | 0 | 0% | 46 |
| <i>Archaeoglobus fulgidus</i> DSM4304 | | Hyp | AN | | AN | 13 | 35% | 9 | 24% | 5 | 14% | 1 | 3% | 3 | 8% | 4 | 11% | 2 | 5% | 0 | 0% | 37 |
| <i>Bacillus halodurans</i> C-1 25 | Firmicutes(1) | Mes | A | | A | 18 | 32% | 19 | 33% | 5 | 9% | 1 | 2% | 9 | 16% | 4 | 7% | 1 | 2% | 0 | 0% | 57 |
| <i>Bacillus subtilis</i> 168 | Firmicutes(1) | Mes | A | | A | 12 | 29% | 16 | 39% | 4 | 10% | 1 | 2% | 4 | 10% | 3 | 7% | 1 | 2% | 0 | 0% | 41 |
| <i>Bruceella meliensis</i> | α-Proteobacteria | Psy | A | | A | 32 | 41% | 18 | 23% | 9 | 12% | 2 | 3% | 9 | 12% | 6 | 8% | 1 | 1% | 1 | 1% | 78 |
| <i>Campylobacter jejuni</i> NCTC 11168 | δ _ε -Proteobacteria | Mes | M | | M | 16 | 38% | 10 | 24% | 3 | 7% | 1 | 2% | 5 | 12% | 6 | 14% | 1 | 2% | 0 | 0% | 42 |
| <i>Caulobacter crescentus</i> CB15 | α-Proteobacteria | Mes | A | | A | 29 | 48% | 10 | 16% | 6 | 10% | 1 | 2% | 8 | 13% | 6 | 10% | 1 | 2% | 0 | 0% | 61 |
| <i>Chlorobium tepidum</i> TLS | | Therm | AN | | Nf | 15 | 37% | 10 | 24% | 3 | 7% | 1 | 2% | 7 | 17% | 3 | 7% | 1 | 2% | 1 | 2% | 41 |
| <i>Chlostridium acetobutylicum</i> ATCC 824 | Firmicutes(2) | Mes | AN | | AN | 32 | 46% | 19 | 28% | 4 | 6% | 1 | 1% | 5 | 7% | 6 | 9% | 2 | 3% | 0 | 0% | 69 |
| <i>Clostridium perfringens</i> | Firmicutes(2) | Mes | AN | | AN | 30 | 56% | 8 | 15% | 4 | 7% | 0 | 0% | 6 | 11% | 5 | 9% | 1 | 2% | 0 | 0% | 54 |
| <i>Corynebacterium glutamicum</i> | Actinobacteria | Mes | A/FA | | A | 31 | 51% | 12 | 20% | 6 | 10% | 1 | 2% | 6 | 10% | 4 | 7% | 1 | 2% | 0 | 0% | 61 |
| <i>Deinococcus radiodurans</i> R1 | | Mes | A | | A | 26 | 40% | 19 | 29% | 5 | 8% | 1 | 2% | 6 | 9% | 6 | 9% | 1 | 2% | 1 | 2% | 65 |
| <i>Escherichia coli</i> K-12 Strain MG1655 | γ-Proteobacteria | Mes | A/AN | | AN | 27 | 42% | 11 | 17% | 7 | 11% | 1 | 2% | 7 | 11% | 10 | 16% | 1 | 2% | 0 | 0% | 64 |
| <i>Fusobacterium nucleatum</i> ATCC 25586 | | Mes | AN | | AN | 25 | 57% | 3 | 7% | 3 | 7% | 1 | 2% | 9 | 20% | 2 | 5% | 1 | 2% | 0 | 0% | 44 |
| <i>Haemophilus influenzae</i> Road KW20 | γ-Proteobacteria | Mes | FAN | | FAN | 27 | 56% | 6 | 13% | 4 | 8% | 1 | 2% | 4 | 8% | 6 | 13% | 0 | 0% | 0 | 0% | 48 |
| <i>Halobacterium</i> sp. NRC-1 | | Mes | AN | | AN | 23 | 61% | 3 | 8% | 3 | 8% | 0 | 0% | 3 | 8% | 5 | 13% | 1 | 3% | 0 | 0% | 38 |
| <i>Halobacterium pylori</i> 26695 | δ _ε -Proteobacteria | Mes | M | | M | 16 | 44% | 8 | 22% | 4 | 11% | 1 | 3% | 4 | 11% | 3 | 8% | 0 | 0% | 0 | 0% | 36 |
| <i>Lactococcus lactis</i> IL1403 | Firmicutes(3) | Mes | A | | A | 20 | 50% | 12 | 30% | 2 | 5% | 0 | 0% | 4 | 10% | 5 | 0 | 0 | 0 | 0 | 0% | 40 |
| <i>Mesorhizobium loti</i> MAFF303099 | α-Proteobacteria | Mes | A | | Nf | 30 | 38% | 21 | 26% | 6 | 8% | 2 | 3% | 8 | 10% | 10 | 13% | 2 | 3% | 1 | 1% | 80 |
| <i>Methanococcus jannaschii</i> DSM 2661 | Methanoarchaea(1) | Hyp | AN | | Meth | 10 | 33% | 5 | 17% | 3 | 10% | 4 | 13% | 2 | 7% | 3 | 10% | 2 | 7% | 1 | 3% | 30 |
| <i>Methanopyrus kandleri</i> AV 1 9 | Methanoarchaea(1) | Hyp | AN | | Meth | 24 | 51% | 9 | 19% | 3 | 6% | 4 | 9% | 2 | 4% | 4 | 9% | 1 | 2% | 0 | 0% | 47 |
| <i>Methanosarcina acetivorans</i> C2A | Methanoarchaea(2) | Mes | AN | | Meth | 21 | 43% | 10 | 20% | 3 | 6% | 3 | 6% | 5 | 10% | 4 | 8% | 3 | 6% | 0 | 0% | 49 |
| <i>Methanosarcina mazei</i> Goel | Methanoarchaea(2) | Mes | AN | | Meth | 20 | 40% | 11 | 22% | 2 | 4% | 3 | 6% | 6 | 12% | 5 | 10% | 3 | 6% | 0 | 0% | 50 |
| <i>Methanobacterium thermoautotrophicum</i> | | Therm | AN | | Meth | 10 | 29% | 9 | 26% | 2 | 6% | 3 | 9% | 3 | 9% | 6 | 18% | 1 | 3% | 0 | 0% | 34 |
| <i>Mycobacterium leprae</i> | Actinobacteria | Mes | A | | A | 18 | 38% | 11 | 23% | 4 | 9% | 0 | 0% | 6 | 13% | 7 | 15% | 1 | 2% | 0 | 0% | 47 |
| <i>Neisseria meningitidis</i> Z2491 | β-Proteobacteria | Mes | A | | A | 21 | 45% | 8 | 17% | 6 | 13% | 1 | 2% | 6 | 13% | 4 | 9% | 1 | 2% | 0 | 0% | 47 |
| <i>Nostoc</i> sp. PCC 7120 | Cyanobacteria | Mes | A | | Nf, Ph | 24 | 46% | 8 | 15% | 5 | 10% | 2 | 4% | 5 | 10% | 7 | 13% | 1 | 2% | 0 | 0% | 52 |
| <i>Pseudomonas aeruginosa</i> PA01 | γ-Proteobacteria | Mes | A | | A | 33 | 45% | 14 | 19% | 10 | 14% | 2 | 3% | 4 | 5% | 8 | 11% | 1 | 1% | 1 | 1% | 73 |
| <i>Pyrobaculum aerophilum</i> IM2 | Crenarchaeota | Hyp | FA | | FA | 19 | 39% | 10 | 20% | 4 | 8% | 2 | 4% | 5 | 10% | 7 | 14% | 2 | 4% | 0 | 0% | 49 |
| <i>Pyrococcus abyssii</i> | Pyrococcales | Hyp | AN | | AN | 18 | 44% | 12 | 29% | 2 | 5% | 0 | 0% | 4 | 10% | 4 | 10% | 1 | 2% | 0 | 0% | 41 |

TABLE 2A
(continued)

| Microorganism | Group | Opt T | O ₂ pref | Mett | Me _{Zn} | F _{Zn} | Me _{Mn} | F _{Mn} | Me _{Cu} | F _{Cu} | Me _{Ni} | F _{Ni} | Me _{Co} | F _{Co} | Me _{V0} | F _{Me} | Me _W | F _W | Me _Y | F _Y | T _{Me} |
|--|-------------------|-------|---------------------|--------|------------------|-----------------|------------------|-----------------|------------------|-----------------|------------------|-----------------|------------------|-----------------|------------------|-----------------|-----------------|----------------|-----------------|----------------|-----------------|
| <i>Pyrococcus furiosus</i> DSM 3638 | Pyrococcales | Hyp | AN | | 16 | 35% | 15 | 33% | 4 | 9% | 0 | 0% | 4 | 9% | 4 | 9% | 3 | 7% | 0 | 0% | 46 |
| <i>Pyrococcus horkoshii</i> OT3 | Pyrococcales | Hyp | AN | | 9 | 47% | 2 | 11% | 3 | 16% | 0 | 0% | 2 | 11% | 3 | 16% | 0 | 0% | 0 | 0% | 19 |
| <i>Ralstonia solanacearum</i> | β-Proteobacteria | Mes | A | | 34 | 44% | 14 | 18% | 11 | 14% | 2 | 3% | 7 | 9% | 8 | 10% | 1 | 1% | 0 | 0% | 77 |
| <i>Salmonella typhimurium</i> LT2 | γ-Proteobacteria | Mes | A | | 36 | 46% | 12 | 15% | 7 | 9% | 0 | 0% | 9 | 12% | 12 | 15% | 2 | 3% | 0 | 0% | 78 |
| <i>Sinorhizobium meliloti</i> 1021 | α-Proteobacteria | Mes | A | Nf | 24 | 32% | 18 | 24% | 9 | 12% | 3 | 4% | 7 | 9% | 10 | 14% | 2 | 3% | 1 | 1% | 74 |
| <i>Staphylococcus aureus</i> Mu50 | Firmicutes(1) | Mes | A/FA | | 14 | 31% | 16 | 36% | 4 | 9% | 1 | 2% | 4 | 9% | 4 | 9% | 2 | 4% | 0 | 0% | 45 |
| <i>Streptococcus pyogenes</i> M1 | Firmicutes(2) | Mes | A | | 14 | 45% | 5 | 16% | 4 | 13% | 0 | 0% | 5 | 16% | 3 | 10% | 0 | 0% | 0 | 0% | 31 |
| <i>Streptomyces coelicolor</i> A3(2) | Actinobacteria | Mes | A | | 28 | 39% | 17 | 24% | 8 | 11% | 2 | 3% | 7 | 10% | 7 | 10% | 1 | 1% | 1 | 1% | 71 |
| <i>Sulfolobus solfataricus</i> P2 | Crenarchaeota | Therm | A | | 13 | 30% | 12 | 28% | 4 | 9% | 2 | 5% | 6 | 14% | 5 | 12% | 1 | 2% | 0 | 0% | 43 |
| <i>Sulfolobus tokodaii</i> 7 | Crenarchaeota | Therm | A | | 8 | 32% | 5 | 20% | 2 | 8% | 1 | 4% | 3 | 12% | 5 | 20% | 1 | 4% | 0 | 0% | 25 |
| <i>Synechocystis</i> sp. PCC 6803 | Cyanobacteria | Mes | A | Nf, Ph | 27 | 39% | 22 | 31% | 5 | 7% | 2 | 3% | 7 | 10% | 6 | 9% | 1 | 1% | 0 | 0% | 70 |
| <i>Thermoanaerobacter tengcongensis</i> MB4T | Firmicutes(3) | Therm | AN | | 27 | 45% | 18 | 30% | 3 | 5% | 2 | 3% | 8 | 13% | 2 | 3% | 0 | 0% | 0 | 0% | 60 |
| <i>Thermoplasma acidophilum</i> | Thermoplasmatales | Therm | A | | 10 | 30% | 9 | 27% | 2 | 6% | 1 | 3% | 4 | 12% | 5 | 15% | 2 | 6% | 0 | 0% | 33 |
| <i>Thermoplasma volcanium</i> GSS1 | Thermoplasmatales | Therm | A | | 13 | 33% | 10 | 26% | 4 | 10% | 1 | 3% | 5 | 13% | 4 | 10% | 2 | 5% | 0 | 0% | 39 |
| <i>Thermotoga maritima</i> MS138 | Thermoplasmatales | Therm | AN | | 18 | 49% | 9 | 24% | 3 | 8% | 0 | 0% | 5 | 14% | 2 | 5% | 0 | 0% | 0 | 0% | 37 |
| <i>Vibrio cholerae</i> N16961 | γ-Proteobacteria | Mes | A | | 26 | 48% | 11 | 20% | 4 | 7% | 0 | 0% | 5 | 9% | 7 | 13% | 1 | 2% | 0 | 0% | 54 |
| <i>Xanthomonas axonopodis</i> pv. citri 306 | γ-Proteobacteria | Mes | A | | 31 | 52% | 13 | 22% | 4 | 7% | 0 | 0% | 5 | 8% | 5 | 8% | 1 | 2% | 1 | 2% | 60 |
| <i>Xanthomonas campestris</i> ATCC 33913 | γ-Proteobacteria | Mes | A | | 34 | 52% | 13 | 20% | 5 | 8% | 0 | 0% | 5 | 8% | 6 | 9% | 1 | 2% | 1 | 2% | 65 |
| <i>Xylella fastidiosa</i> 9a5c | γ-Proteobacteria | Mes | A | | 24 | 55% | 7 | 16% | 4 | 9% | 0 | 0% | 5 | 11% | 4 | 9% | 0 | 0% | 0 | 0% | 44 |
| <i>Yersinia pestis</i> CO-92 Biovar Orientalis | γ-Proteobacteria | Mes | A | | 23 | 43% | 10 | 19% | 5 | 9% | 2 | 4% | 6 | 11% | 6 | 11% | 1 | 2% | 0 | 0% | 53 |
| AVERAGES | | | | | 22 | 42% | 11 | 22% | 5 | 9% | 1 | 2% | 5 | 11% | 5 | 11% | 1 | 2% | 0.2 | 0% | 51 |

Psy = psychrophile; Mes = mesophile; Therm = thermophile; Hyp = hypothermophile; A = aerobic; AN = anaerobic; M = microaerophilic; F = facultative; NF = nitrogen fixer; Meth = methanogen; Ph = oxygenic photoautotroph

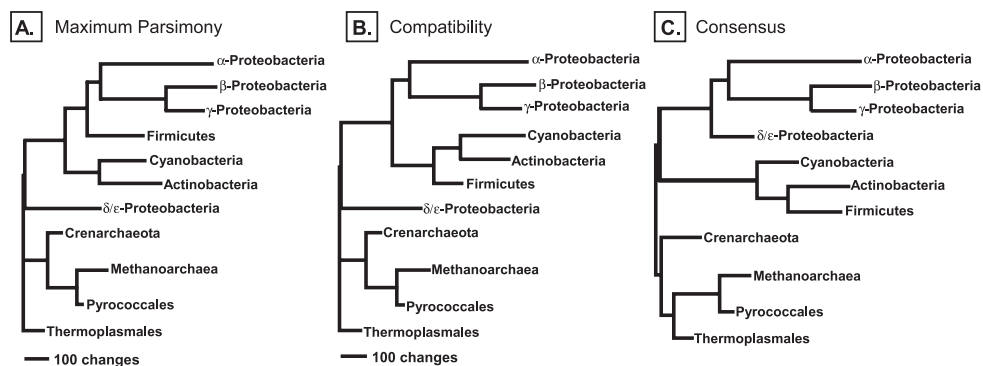


Fig. 2. (A) Composite organism genomic tree generated by maximum parsimony. (B) Composite organism genomic tree generated by compatibility. (C) Consensus tree representing our analysis of combined gene families considered along with the most widely accepted microbial tree topologies.

others, 2002; House and others, 2003b; Yang and others, 2005); and (2) cyanobacteria have been placed as a sister group to the gram positive bacteria (Gerstein, 1998; Hansmann and Martin, 2000; Blank, 2004; this study). Our trees of composite organisms (fig. 2A and 2B) differ from the consensus tree (fig. 2C) in that our genomic trees failed to unite the Proteobacteria, and placed the Thermoplasmatales at the base of the tree, presumably due to difficulties in analyzing genomes of differing size (House and Fitz-Gibbon, 2002).

INFERRED CARBON ISOTOPE SIGNATURES OF ANCIENT ORGANISMS

Based on the observation that the reductive TCA and 3-hydroxypropionate cycles result in low carbon isotope fractionation between CO_2 and biomass, the reductive pentose-phosphate cycle and acetyl-CoA pathway result in moderate fractionation between CO_2 and biomass, and pathways using other lighter substrates result in extremely ^{13}C depleted biomass (as in fig. 1), we superimposed the distribution of modern carbon fixation pathways and the resulting carbon isotope signatures onto a modified version of our consensus tree of life (fig. 3). From extrapolation by parsimony, we then inferred the carbon fixation and approximate isotopic signature for the common ancestors of various ancient prokaryotic groups, including the last common ancestor (LCA). For ethanotrophy (E), a biological process for which fractionation factors have not been measured, we assumed some degree of biological fractionation from natural ethane, which is itself isotopically depleted. For example, measured $\delta^{13}\text{C}$ values for ethane in deep sea sediments are around -60 permil due to biological fractionation during synthesis (Oremland and others, 1988; Waseda and Didyk, 1995; Paull and others, 2000).

The results shown in figure 3 suggest the following: (1) both low and moderate carbon isotope fractionation were established quickly in the early evolution of life; and (2) methanotrophic and ethanotrophic metabolisms capable of producing biomass with extreme ^{13}C depletions are not primitive, but rather evolved after the major groups of Prokaryotes had already diverged. Based on the distribution of carbon fixation pathways in figure 3, it appears more or less equally parsimonious to assign low or moderate carbon isotope fractionation to the base of the Bacteria and the base of the Archaea, and thus to the last common ancestor. This conclusion follows because it is difficult to extrapolate the relative antiquities of the reductive TCA cycle, acetyl CoA pathway, and reductive pentose phosphate cycle. For example, the last common ancestor of the Archaea could have exhibited high fractionation through either a

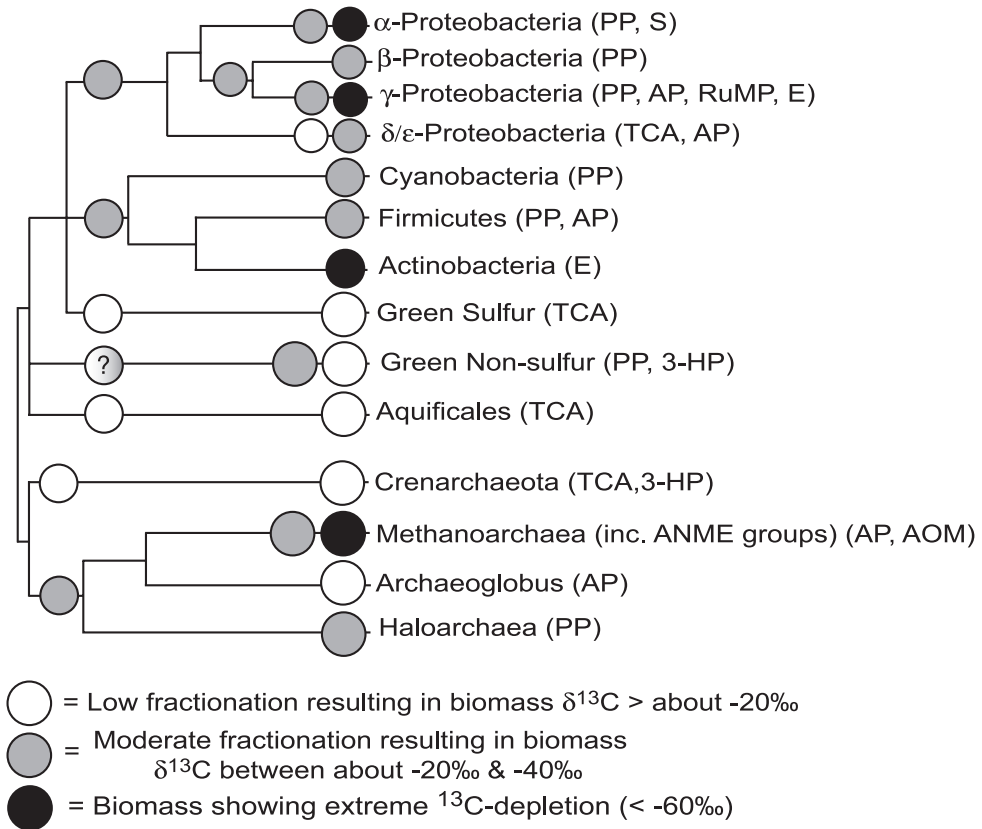


Fig. 3. Carbon isotopic signatures from the literature (fig. 1; tables 1A and 1B) superimposed on our consensus tree of life (fig. 2C). For this analysis, we additionally added single organisms with known carbon fixation pathways that did not have representative whole genome sequences or did not fit into previously assigned phylogenetic groupings. Carbon fixation pathways are noted as in fig. 1, with the addition of ethanotrophy (E). Note that we positioned the Haloarchaea as a sister group to the Methanoarchaea and *Archaeoglobus*, within the Euryarchaeota. This placement is primarily based on arguments presented in House and others (2003b), however it is currently the subject of debate (for example, Korbel and others, 2002; Yang and others, 2005). Additionally, the anaerobic methane oxidizing archaea (ANME groups) have been included with the Methanoarchaea (Orphan and others, 2001a). Ancient isotopic signatures are extrapolated using simple parsimony.

variant of the acetyl CoA pathway as found in methanogens, or through the reductive pentose phosphate cycle, as found in the haloarchaea (Rawal and others, 1988).

It has been suggested that the acetyl Co-A pathway is the oldest carbon fixation pathway (Huber and Wächtershäuser, 1997; Martin and Russell, 2004). Alternatively, the wide phylogenetic distribution of the reductive pentose phosphate cycle (Rawal and others, 1988; Ivanovsky and others, 1999), and the ubiquitous occurrence of diverged forms of Rubisco in both Bacteria and Archaea support the antiquity of this pathway (Finn and Tabita, 2003; Hügler and others, 2003). The Rubisco enzyme, however, could have had a complex history, originally evolving as an oxygenase (Hanson and Tabita, 2001). The TCA cycle, which seems to have evolved originally as a reductive pathway (Romano and Conway, 1996), is universally important in biosynthesis and therefore must have evolved especially early making it perhaps the most likely carbon fixation pathway for the last common ancestor (Wächtershäuser, 1990; Morowitz and others, 2000). For the purposes of extrapolating carbon isotopic signatures to

the LCA, if the reductive TCA cycle is the oldest carbon fixation pathway, then the LCA would demonstrate low fractionation.

We expect that carbon isotopic signatures of biomass identified experimentally in modern organisms using different carbon fixation pathways have been manifested in isotopic compositions of ancient microorganisms, and subsequently preserved in microfossils (House and others, 2000; Ueno and others, 2001; Kaufman and Xiao, 2003) and other organic matter in the rock record (for example, Hayes and others, 1983). Carbon isotopic studies of Precambrian sediments have traced a distinctive isotopic signature of highly fractionated biomass, resulting from biological carbon fixation, to ~3,500 million years ago (Schidlowski and others, 1983; Hayes and others, 1992). For example, recent studies of Buck Reef Chert, South Africa, have shown that photosynthetic mats were an important microbial ecosystem in shallow water environments at 3,416 million years ago (Tice and Lowe, 2004). These laminated cherts contain carbonaceous matter with a carbon isotopic composition of -20 to -30 permil, consistent with fractionation by the reductive pentose phosphate cycle (Tice and Lowe, 2004). It is now possible to partially explain the isotopic fractionations found throughout the Archean through anoxygenic photoautotrophy. In the past, such an explanation was hindered (House, ms, 1999) because the known deeply diverging lineages of anoxygenic phototrophs, *Chloroflexus* and *Chlorobium*, do not use the reductive pentose phosphate cycle; and therefore do not demonstrate sufficiently large isotopic fractionations during carbon fixation to explain the isotopic record prior to the evolution of Cyanobacteria or photosynthetic Proteobacteria (like *Chromatium*). However, it is now known that *Oscillochloris* (Keppen and others, 2000), a Green Non-sulfur Bacteria, uses the reductive pentose phosphate cycle for carbon fixation (Ivanovsky and others, 1999). This lineage, along with the recently described 'Candidatus Chlorothrix' which has been shown not to use the 3-hydroxypropionate cycle (Klappenbach and Pierson, 2004), may be modern examples of the kinds of anoxygenic photoautotrophs that built microbial mats 3,500 million years ago producing organic material with moderately high ^{13}C -depletions. Furthermore, the prospect that anoxygenic photoautotrophs, such as *Oscillochloris*, were mat builders in the early Archean is consistent with reported "cyanobacteria-like" microfossils of that Eon (Schopf and Packer, 1987; Schopf, 1993; Schopf and others, 2002), as the morphology of *Oscillochloris* is remarkably similar to that of filamentous cyanobacteria (Gorlenko and Pivovarova, 1977; Keppen and others, 1994).

Compared to younger Archean sediments, a relatively low fractionation has been observed in turbidites from the 3,800 million years old Isua terrain in central West Greenland (Rosing, 1999), and a large amount of ^{13}C -depletion has been reported in graphite inclusions in apatite from the 3,850 million years old Akilia Island metasediments in southwest Greenland (Mojzsis and others, 1996). For the Isua terrain, the carbon isotopic composition of the organic matter in turbidite sediments (about -19‰ ; Rosing, 1999) and inclusions in various mineral phases (down to about -18‰ ; Ueno and others, 2002) could be consistent with fractionation by a primitive autotroph using the reductive TCA cycle (van der Meer and others, 2000). However, since the carbon isotopic results from the more ancient Akilia Island seem to show greater ^{13}C -depletion, an explanation of this carbon is more complicated (also see Knoll, 2003). The Akilia Island results would suggest that: 1) the last common ancestor of life dates from earlier than 3,850 million years ago; 2) processes other than biologic carbon fixation have contributed to the apparently large magnitude fractionations reflected in the ancient Akilia samples; and/or 3) the reductive TCA cycle is not the oldest carbon fixation pathway. (Note: during final preparation of this manuscript, Lepland and others, 2005, published concerns about the relevance of Akilia apatites as

a recorder of the past record of life on Earth based on their inability to reproduce the earlier results of Mojzsis and others, 1996).

For example, it is possible that life originated and diversified extensively before the 3,800 to 3,900 million year old sedimentary units were deposited. Growing evidence suggests that the Earth was habitable before the oldest preserved sediments. Zircons from arc volcanism, dated from 4,200 to 4,400 million years ago, have been recovered from Jack Hills, Western Australia, suggesting that oceans were present at that time (Mojzsis and others, 2001; Wilde and others, 2001). Additionally, studies of the history of lunar impacts suggest that the Earth was comparatively peaceful from ~4,400 to ~3,900 million years ago, between the completion of an accretionary event ~4,450 million years ago and a late intensive bombardment at ~3,900 million years ago (Ryder, 2003), allowing for a period of prior planetary habitability.

MODEL METALLOMES OF MODERN AND ANCIENT ORGANISMS

Comparing Model Metallomes from Modern Genome Sequences

We hypothesized that the prokaryotic metallome could provide a further biosignature of microbial physiology in the environment. To test this hypothesis, we examined the model metallomes of modern prokaryotes predicted by the presence or absence of genes encoding for metallo-enzymes in whole genome sequences, and compared these values across specific metabolisms, oxygen tolerance, optimum growth temperature, and phylogeny (fig. 4 and fig. 5). When making broad comparisons across groups of organisms, if we suppose equal gene expression and a relatively constant T_{Me} , then a higher F_i value for a specific metal means that the organism generally requires more atoms of that metal, if all enzymes are being expressed equally.

Results of these comparisons suggest that the use of metals in prokaryotes when analyzed as a group follows the general hierarchy $Fe \gg Zn > Mn \gg Co, Cu, Mo \gg Ni > W, V$ (fig. 4 and fig. 5). As explained above, iron is nearly universally required for life and therefore not even considered in the model metallome as we define it. Zinc (F_{Zn} 30 – 50%) forms metallopeptidases, metallophosphatases, and many other zinc polymerases involved in DNA and RNA binding and synthesis (Lipscomb and Sträter, 1996). Manganese (F_{Mn} 20 – 30%) catalyzes several redox reactions, hydration-dehydration, isomerizations, phosphorylation-dephosphorylation, phosphoryl transfer, and dioxygen production in photosynthesis (Dismukes, 1996; Yachandra and others, 1996). Cobalt (F_{Co} 10 – 20%) forms the cobalamin cofactor of the essential vitamin B12, which is required for a number of enzymes (Marsh, 1999). Copper (F_{Cu} 10 – 20%) forms the common 'blue' copper proteins used in electron transfer, often associated with oxidative enzymes and energy capture (Guss and Freeman, 1983; Solomen and others, 1996). And molybdenum (F_{Mo} 5 – 20%), in addition to its involvement in nitrogen metabolism (in nitrogenase and nitrate reductase), is used in a wide variety of other enzymes that catalyze an oxygen atom transfer or two-electron reaction, including xanthine oxidase, sulfate reductase, and fumarate dehydrogenase (Burgess and Lowe, 1996; Eady, 1996; Hille, 1996; Kisker and others, 1997; Frausto da Silva and Williams, 2001). Nickel, tungsten, and vanadium have much more limited biological uses, as discussed below.

We compared model metallomes for organisms with specialized metabolic functions to determine if their use of specific metallo-enzymes produces a unique metal signature that could be apparent in the expressed cellular metallome (fig. 4A). We chose to compare organisms that perform methanogenesis, nitrogen fixation, and oxygenic photosynthesis, as these are metabolic functions that are important in elemental cycling in the environment and are each associated with one or more specific metallo-enzymes.

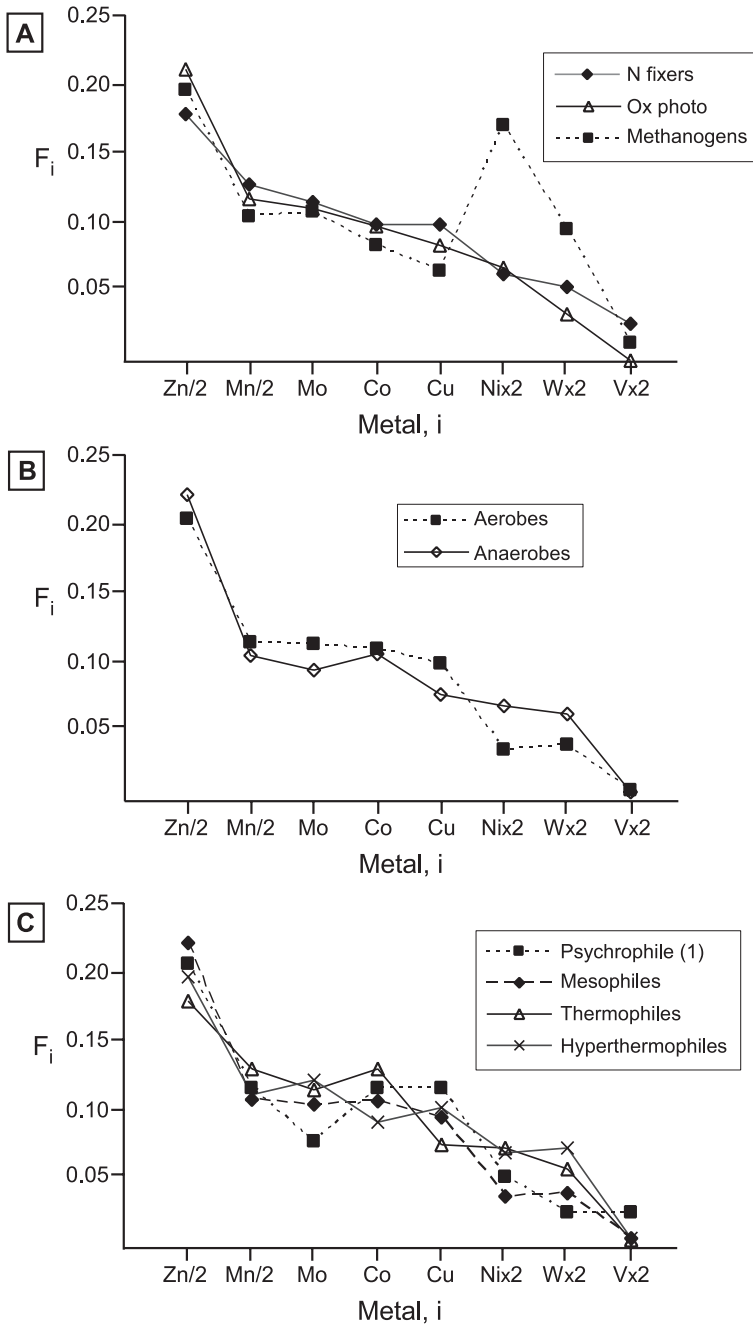


Fig. 4. Average F_i values predicted by whole genomes compared across: (A) specific metabolisms; (B) oxygen tolerance; and (C) optimum growth temperature. Note the scaling (Zn/2, Mn/2, Nix2, Wx2, and Vx2).

Model metallome comparisons indicate that of the metabolisms we tested, only the model metallome of methanogens contains a unique signature. Methanogenic archaea have significantly elevated F_{Ni} and F_W ($F_{Ni} \sim 9\%$ and $F_W \sim 5\%$ for methano-

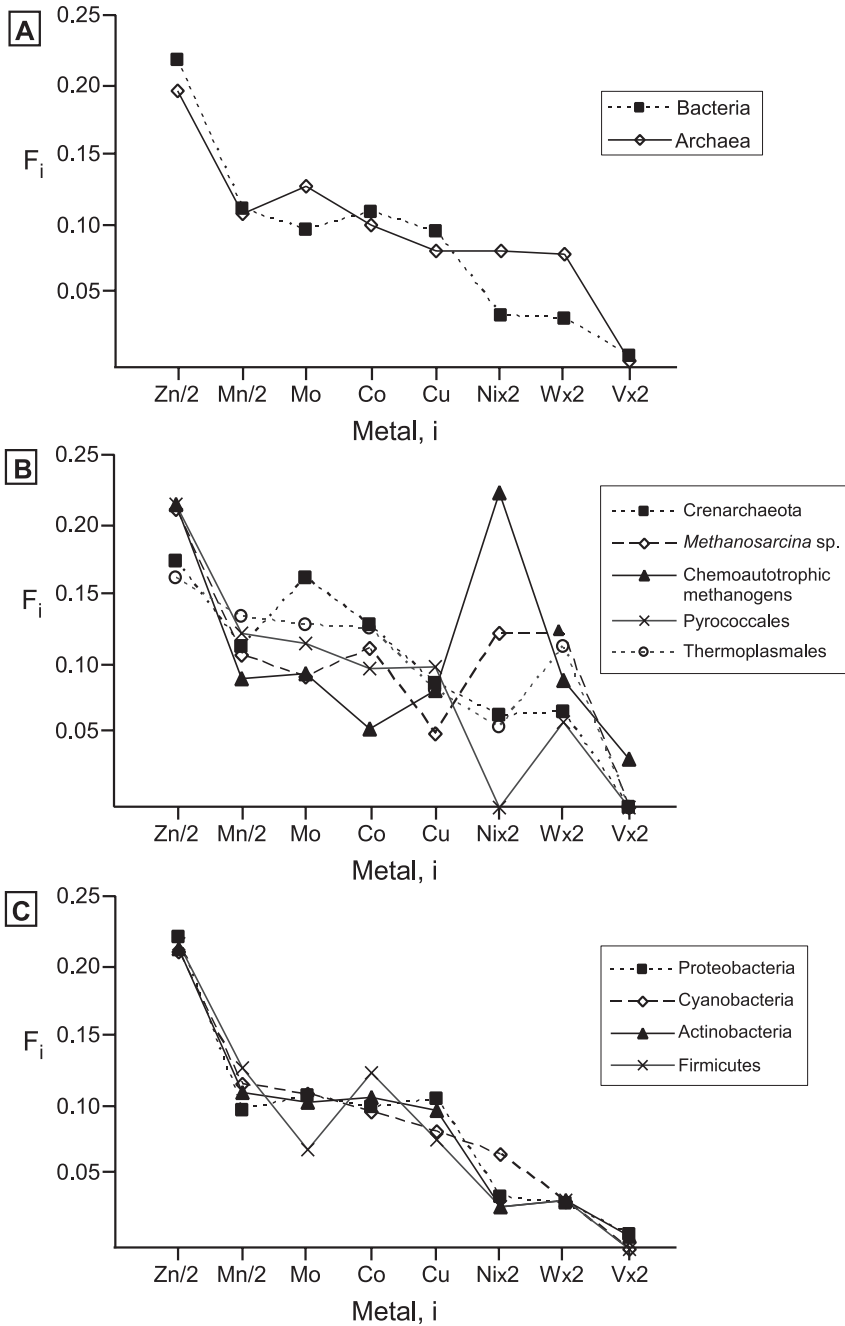


Fig. 5. Average F_i values compared across phylogenies: (A) Bacteria and Archaea; (B) individual groups of archaea; and (C) individual groups of bacteria. Note the scaling as in figure 4 (Zn/2, Mn/2, Nix2, Wx2, and Vx2).

gens, versus $F_{Ni} \sim 2\%$ and $F_W \sim 2\%$ for non-methanogens; fig. 4A and fig. 6A). Nickel is extremely important in methanogenesis, forming the cofactor of two primary enzymes involved in methane production, methyl-coenzyme M reductase and acetyl-CoA syn-

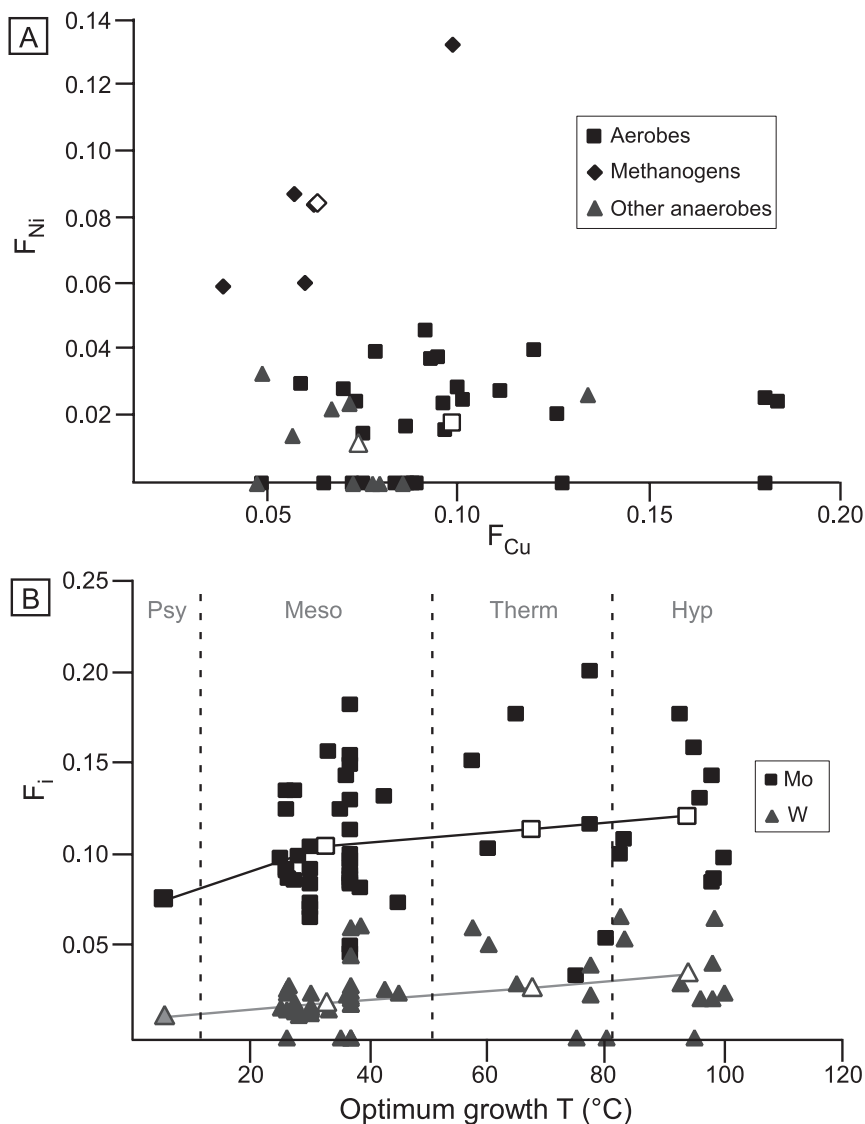


Fig. 6. (A) F_{Cu} versus F_{Ni} values for aerobes, methanogens, and non-methanogenic anaerobes. Average values for each group are shown as open symbols. (B) F_{Mo} and F_{W} for organisms plotted against optimum growth temperature (for organisms with a range of optimum growth temperatures, the mean temperature was used). Average values for each group (abbreviated as in table 1) are connected with a line and shown as open symbols.

these. It is also used in carbon monoxide dehydrogenase, which can catalyze the reduction of CO_2 to carbon CO and subsequent assembly of acetyl-CoA, key to pathway of carbon fixation for acetogenic methanogens (Ragsdale and Kumar, 1996). In non-methanogens nickel is additionally used in urease, some hydrogenases, and rarely in superoxide dismutase (Frausto da Silva and Williams, 2001). Hyperthermophilic methanogens use tungsten in aldehyde ferredoxin oxidoreductase (AOR) enzymes, and in some thermophilic methanogens tungsten is also used preferentially to molyb-

denum in formylmethanofuran dehydrogenase, catalyzing the first step in the conversion of CO_2 to CH_4 (L'vov and others, 2002).

Our examinations indicate the model metallomes of nitrogen fixers and cyanobacteria, on the other hand, do not contain unique metal signatures. We calculate only slightly higher F_{Mo} values in nitrogen-fixing microorganisms (F_{Mo} ~12% for diazotrophs, versus F_{Mo} of 10% for non N-fixers), despite their requirement for molybdenum in the Fe-Mo cofactor of the nitrogenase enzyme responsible for nitrogen-fixation. In the absence of molybdenum, some organisms can form alternative nitrogenases containing vanadium and iron or iron only (Eady, 1996; Burgess and Lowe, 1996; Howard and Rees, 1996), so that diazotrophs additionally have a small but insubstantial F_{V} (1 – 2%) not found in most non N-fixers, as the only other well-understood use of vanadium in biology is in haloperoxidases (Frausto da Silva and Williams, 2001). Due to the substantial requirement of manganese in photosystem II, which is responsible for oxygenic photosynthesis (using 4 atoms of manganese per molecule of enzyme; Yachandra and others, 1996), we expected that oxygenic phototrophs might exhibit elevated F_{Mn} values over non-phototrophs. However, the two cyanobacteria in our database have a mean 23 percent F_{Mn} , an increase of only one percent over non-phototrophs.

High F_{Ni} and F_{W} values in methanogenic archaea support the idea that these organisms may contain elevated cellular nickel and tungsten concentrations in excess of other organisms, providing a biosignature for methanogens utilizing the known Ni- and W-containing enzymes in the environment. In fact initial measurements of the cellular metal content of *Methanosarcina* demonstrate high nickel and tungsten concentrations (House, unpublished data). Furthermore, in our calculations of a model metallome we ignored gene expression, assuming that all enzymes were expressed to the same extent in all organisms at all times. However, in different environments under different conditions, specific proteins will be expressed in higher quantity than others in a given organism. For example, Raven (1990) calculated an increase in both manganese and iron requirements for photosynthesis at low photon flux densities and with different ATP sources and CO_2 assimilation mechanisms. Additionally, growth in the absence of fixed nitrogen requires a high concentration of the nitrogenase enzyme, and as a result some diazotrophs accumulate 10 to 20 percent of the total cell protein as nitrogenase under nitrogen-limited conditions (Martinez-Argudo and others, 2004). If the metallo-enzymes responsible for specific metabolic capabilities such as nitrogen fixation and oxygenic photosynthesis are expressed in excess of other enzymes under conditions when these metabolisms are active, then this would result in an amplification of the metal in the expressed cellular metallome beyond that predicted by our model. Thus the expressed metallome could still potentially contain a biosignature for the activation of these metabolisms in nature.

Next we compared the model metallomes of strict aerobes and strict anaerobes (fig. 4B). This comparison indicates that anaerobes have elevated F_{Ni} and F_{W} and slightly lower F_{Cu} and F_{Mo} than aerobes (average for anaerobes: F_{Ni} and $F_{\text{W}} > 3\%$, $F_{\text{Cu}} \sim 8\%$, $F_{\text{Mo}} \sim 9\%$; average for aerobes F_{Ni} and $F_{\text{W}} \sim 2\%$, $F_{\text{Cu}} \sim 10\%$, $F_{\text{Mo}} \sim 11\%$). A closer examination of F_{Ni} for individual organisms (fig. 6A) clearly shows that the increased nickel signature for anaerobes is due to the bias in our anaerobic database for methanogenic archaea, which have elevated F_{Ni} as explained above. This examination also shows that most anaerobes have F_{Cu} less than 9 percent, while aerobes have F_{Cu} values up to twice that (fig. 6A). The two exceptions were *M. jannaschii* and *A. fulgidus*. This is difficult to explain for *A. fulgidus*; however, *M. jannaschii* has a relatively small total number of identified metallo-enzymes, most likely due to poor genome annotation (T_{Me} of 30 for this organism, much less than the average T_{Me} of 51 for all

organisms examined), such that even though it only requires 3 atoms of copper in our model (in amine oxidase and nitrate reductase), this leads to a higher overall F_{Cu} .

We also compared the model metallomes of organisms with varying optimum growth temperatures (fig. 4C). Increasing trends in F_{Mo} and F_W are seen with increasing thermophily (fig. 4C and fig. 6B). F_{Mo} increases from around 8 percent in the one psychrophile we examined to greater than 12 percent in hyperthermophiles. At high temperatures, molybdenum forms the primary cofactor of enzymes such as formyl methanofuran dehydrogenase (L'vov and others, 2002). F_W is ~4 percent for hyperthermophiles, ~3 percent for thermophiles, ~2 percent for mesophiles, and ~1 percent for the one psychrophile examined. As mentioned above, tungsten is used in AOR enzymes, best known from the hyperthermophile *Pyrococcus furiosus* (Mukund and Adams, 1991; Chan and others, 1995; Kletzin and others, 1995). Hyperthermophilic archaea also contain two other AOR-type tungstoenzymes that catalyze the oxidation of aldehydes, formaldehyde ferredoxin oxidoreductase and glyceraldehyde ferredoxin-oxidoreductase (Johnson and others, 1996). In some thermophilic organisms, tungsten can also be substituted for molybdenum in molybdo-enzymes such as formylmethanofuran dehydrogenase (L'vov and others, 2002). Despite the small increase in F_W , tungsten is such a minor contributor to the total metallome ($\leq 6\%$ for all organisms in this study), that unless the proportion of tungsten-containing proteins expressed in an organism or group of organisms (that is hyperthermophiles) is found to be higher than that of other metallo-enzymes, tungsten content will not be a useful biomarker for thermophiles. Preliminary studies indicate cellular tungsten concentrations in all hyperthermophiles studied thus far are less than 0.1 parts per billion (Cameron and others, 2004).

Similar comparisons can be made between model metallomes across phylogenetic groups (fig. 5). These comparisons suggest relatively constant F_{Zn} , F_{Mn} , F_{Cu} , and F_{Co} , with fluctuating F_{Mo} , F_{Ni} , and F_W . The average archaeal metallome predicts higher F_{Mo} (~13%), F_{Ni} (~4%), and F_W (~4%), when compared to the average bacterial metallome (F_{Mo} 10%, F_{Ni} ~2%, and F_W ~2%; fig. 5A). Individual groups of Archaea have much more widely variant model metallomes (fig. 5B) than individual groups of Bacteria (fig. 5C). All archaeal and bacterial groups shown in figures 5B and 5C have statistically significant model metallomes, except for the firmicutes, which had to be separated into three groups unrelated to previous phylogenetic groupings. Mean values for F_{Mo} , F_{Co} , F_{Ni} , and F_W for Archaeal groups all vary significantly (fig. 5B). Again, the most clear signals are elevated F_{Ni} and F_W for *Methanosarcina* sp. and chemoautotrophic methanogens, and elevated F_W for the methanogens and thermophilic Thermoplasmatales. Compared to the groups of Archaea, groups of Bacteria are fairly uniform, with only the Firmicutes (a mean of all three statistically significant groups) and cyanobacterial F_{Ni} differing significantly.

Evolution of the Metallome from Modern Genome Sequences

Model metallomes of phylogenetic groups can be mapped onto our consensus tree of life (fig. 2C), and used to infer evolutionary changes (fig. 7). Extrapolation to the base of the tree of life by consecutive calculation of each mean F_i from the F_i of converging branches projects a model metallome for the LCA: F_{Zn} ~40 percent, F_{Mn} ~20 percent, F_{Mo} , F_{Co} , and F_{Cu} ~10 percent each, F_{Ni} and F_W ~3 percent (fig. 7). The predicted model metallome of the LCA is very similar to the average model metallome we calculated from genomes of modern anaerobes in this study, as in figure 4 (F_{Zn} = 44%, F_{Mn} = 21%, F_{Mo} = 9%, F_{Co} = 11%, F_{Cu} = 8%, F_{Ni} = 3% and F_W = 3%; fig. 8). This similarity is supported by the isotopic evidence for life more than 3,800 million years ago (see discussion above), prior to the oxygenation of the Earth's atmosphere (Rye and Holland, 1998).

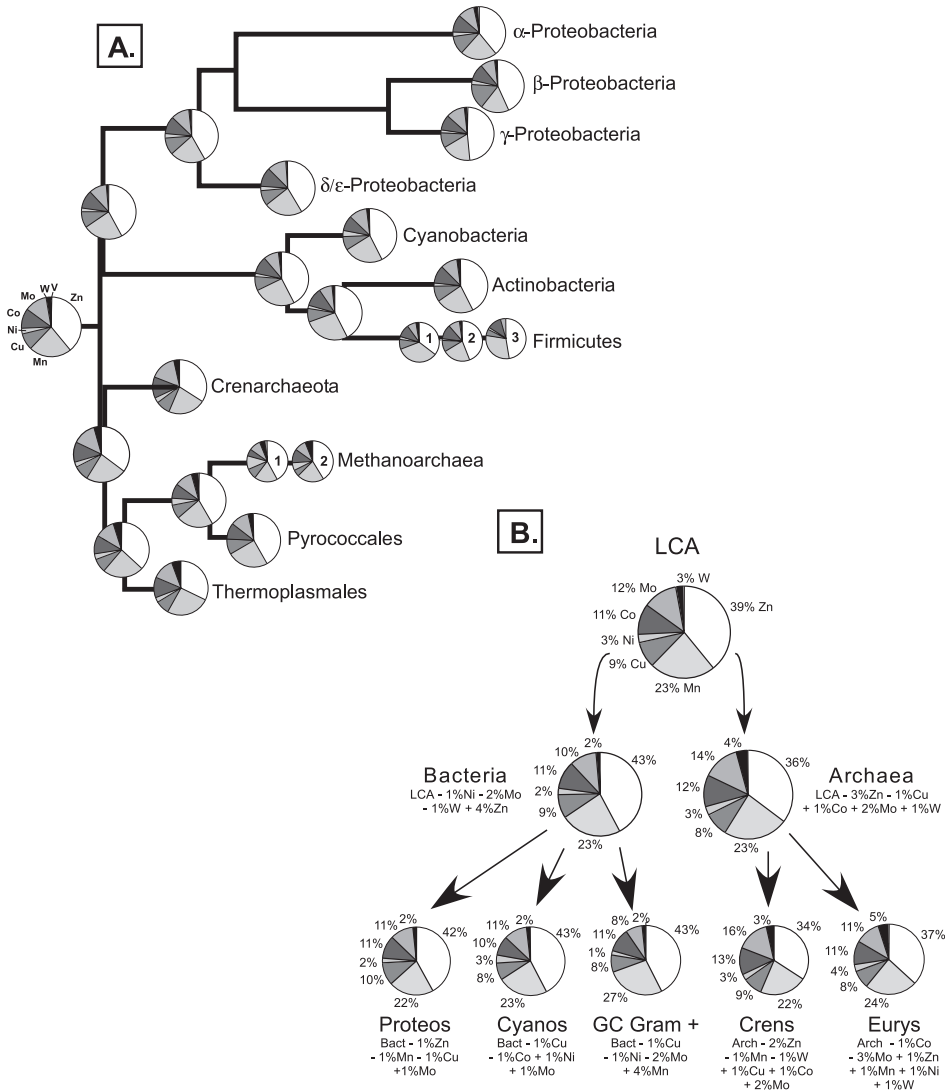


Fig. 7. (A) Model metallomes extrapolated onto a consensus tree of life. (B) Possible scenario for metallome evolution based on this extrapolation.

A hypothesis for the possible evolution of trace metal use in the microbial metallome can be formed by assuming that Bacteria and Archaea with mean F_1 values evolved from the last common ancestor, and then specific groups of Bacteria and Archaea evolved from these organisms (fig. 7). The predicted changes in the metallome with each evolutionary step may have been caused by: 1) changes in trace metal availability due to fluctuating redox conditions in the organism's environment; 2) competition between evolving organisms for available trace metals; 3) opportunistic evolution of enzymatic uses for widely available but previously unused trace metals; 4) gain of genes encoding for metallo-enzymes by lateral gene transfer; and/or 5) loss of genes encoding for metallo-enzymes along with unnecessary metabolic functions. Early in the evolution of life, it is likely that changes in trace metal availability due to

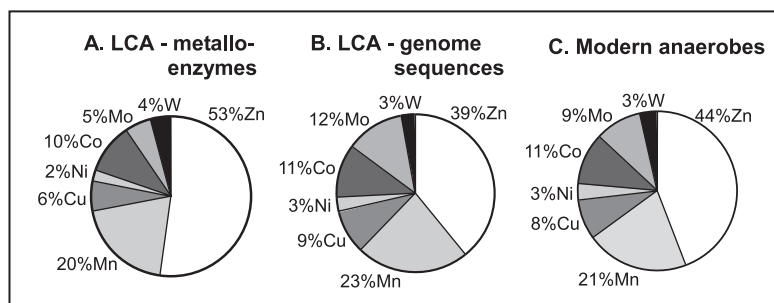


Fig. 8. Model metallomes for: (A) the LCA, based on metallo-enzymes inferred to be present at this time, from figure 9; (B) the LCA, based on extrapolation of model metallomes from our consensus tree in figure 7; and (C) the average of modern anaerobes, from figure 5.

variable redox conditions and competition with other organisms were important in driving biological metal use. As complex life evolved and developed specialized lifestyles, lateral gene transfer and loss of extraneous genetic material likely became important in shaping trace metal utilization (for example, Gogarten and others, 2002).

Evolution of the Metallome from Metallo-enzyme Functions

Another way to view the evolution of the use of metals in biology is by considering the timing of evolution of modern metallo-enzymes, based on their individual function and occurrence in life today (fig. 9). Although some enzymes are believed to have initially developed for one biological use and later modified and adapted to fit a different function (for example nitrogenase; Silver and Postgate, 1973; Normand and Bousquet, 1989), we assume a unique origin for each individual metallo-enzyme, and assign it an inferred timing of evolution based on the rules presented in table 3. The metallo-enzymes can then be mapped onto a geochemical timeline based on the inferred timing of major evolutionary events and we can examine how the use of these metals has developed with changes in global ocean redox chemistry (fig. 9). Note: a growing body of evidence indicates that beginning around 1,800 million years ago, after an initial pulse of oxygen into the atmosphere, the Earth's oceans experienced a prolonged period of deep water euxinia, before a second pulse of O_2 brought atmospheric levels to near current values (Canfield, 1998; Shen and others, 2002, 2003; Arnold and others, 2004; Poulton and others, 2004; shown as a dark hatch in fig. 9B). This could have dramatically altered the bioavailability of many redox-sensitive metals during this time (fig. 9B; Saito and others, 2003). Some researchers have speculated that this crisis in metal bioavailability could have forced a change in the use of specific metals in metallo-enzymes (Anbar and Knoll, 2002; Saito and others, 2003). However, as we have no conclusive evidence for biological innovation at this time, we can only speculate as to what these changes might have been. Therefore, although the bioavailability of metals during the mid-Proterozoic calculated by Saito and others (2003) is shown in figure 9B, our interpretation of the evolution of metal use with bioavailability considers only a single oxygenation event.

The metals we examined fall into one of three categories: (1) metals with biological uses that developed along with their bioavailability (Cu and Mo); (2) metals with biological uses that developed counter to their bioavailability (Fe and Mn); and (3) metals with biological uses that did not fluctuate significantly through time (Zn, Co, Ni, W, and V).

The biological uses of both copper and molybdenum have increased along with their bioavailability since oxygenation of the Earth's atmosphere. We speculate the

maximum number of metallo-enzymes utilizing copper doubled from 10 to 20 enzymes with oxygenation. Today copper is commonly used in 'blue proteins' involved in electron transfer, especially in oxidative enzymes and energy capture (Guss and Freeman, 1983). Copper enzymes are also commonly used in oxidase reactions involving small non-metal compounds that only became available to any extent with the rise of dioxygen, including phenols, amines, ferrous ions, and oxidized nitrogen species (Solomen and others, 1996). Similarly, the maximum number of metallo-enzymes utilizing molybdenum more than doubled from 13 to 27 enzymes with oxygenation. Today, molybdenum enzymes catalyze reactions ranging from six-electron reductions (for example nitrogenase) to oxygen atom transfer. The increased use of copper and molybdenum in oxygen transformation is consistent with an increase in the biological availability of these metals with complete oxidation of the oceans (fig. 9). Under oxidizing conditions such as the modern ocean, copper is moderately available as Cu^{2+} , and molybdenum forms the highly mobile molybdate anion, MoO_4^{2-} , which is unreactive in seawater, exhibiting a conservative profile with an average concentration of 105 nanomolar (Emerson and Huested, 1991).

In contrast to Cu and Mo, biological use of iron and manganese has increased counter to their bioavailability since oxidation of Earth's atmosphere (fig. 9). We speculate that the maximum number of metallo-enzymes using iron has nearly tripled from 26 to 74 enzymes with oxygenation, while the maximum number of enzymes using manganese has increased from 13 to 18, despite the predicted decrease in the availability of these metals in Earth systems with progressive oxidation of the atmosphere (fig. 9). Despite the low solubility under oxidizing conditions of minerals containing these elements, both Fe and Mn are ubiquitous in natural environments. Reduced ferrous iron (Fe^{2+}) is readily available under anoxic conditions such as on the early Earth, where it could have easily been used to form early metallo-enzymes (Holland, 1973). However, in a transitional sulfidic ocean iron would have been highly complexed with sulfides, and might have existed in concentrations three orders of magnitude lower than Archean seawater (Lewis and Landing, 1991, 1992; Saito and others, 2003). In modern oxygenated oceans iron is estimated to exist at no more than nanomolar concentrations, due to the formation of insoluble Fe-hydroxides (Gordon and others, 1982; Landing and Bruland, 1987; Martin and Gordon, 1988; Bruland and others, 1991; Donat and Bruland, 1995). Despite the decreasing availability of iron under surface conditions with the progressive oxidation of the Earth, the use of iron in transformations of molecular oxygen and in dealing with oxygen radical toxicity seems to have increased with the rise of oxygen in the Earth's atmosphere. Iron's use in enzymes dealing with oxygen radicals (for example, superoxide dismutase), and in multiple reductase enzymes (sulfite, nitrite, *et cetera*) could have evolved in organisms living in anoxic environments where iron was still abundant. In contrast, the wide use of iron in molecular oxygen transformations in most dioxygenases suggests that mechanisms for acquiring iron in oxic environments might have evolved very early in life as an adaptation to rising oxygen in the Earth's atmosphere (Raven, 1995). For instance, some microorganisms secrete ligands, called siderophores, designed almost exclusively to bind and take up Fe^{3+} (for example, Neilands, 1984; Hughes and Poole, 1989). These siderophores are low-molecular weight organic molecules, commonly catechols or hydroxamates, with formation constants for ferric iron as high as 10^{51} (Hider, 1984; Hughes and Poole, 1989; Winkelmann, 1991; Hersman and others, 1995).

Manganese has similar redox sensitivity, and can therefore substitute into some iron-containing enzymes (such as superoxide dismutase). Other manganese enzymes are used in a wide variety of functions, most notably dioxygen production, as in photosystem II, catalase, superoxide dismutase, and in the organelles of eukaryotes

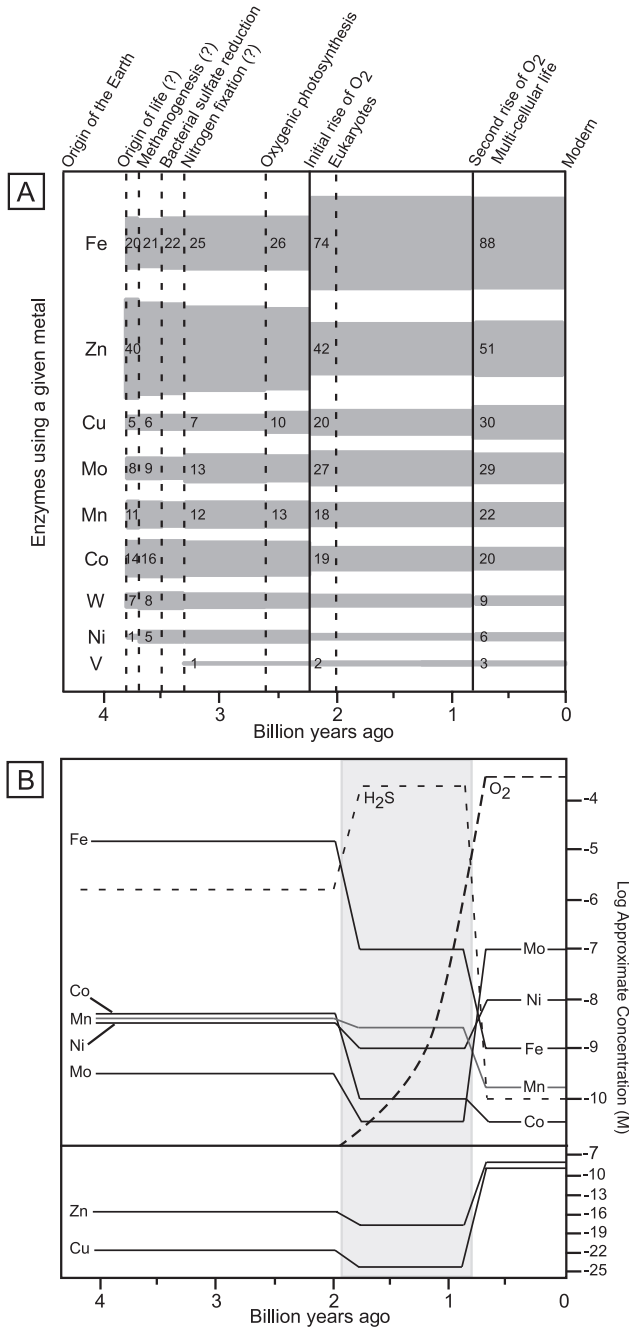


Fig. 9. (A) A second way to view the evolution of the use of metals in biology, considering the timing of evolution of known modern metallo-enzymes based on their individual function and occurrence in life today (using rules outlined in table 3). The inferred evolutionary timing of these enzymes is based on the estimated timing of major biological innovations (from Young, 1992; Golubic and others, 1995; Beaumont and Robert, 1999; Brocks and others, 1999; Summons and others, 1999; Shen and others, 2001; Habicht and others, 2002; Battistuzzi and others, 2004) and the estimated timing of oxygenation of the Earth’s atmosphere (Holland, 1984; Knoll and others, 1986; Holland and others, 1990; Des Marais and others, 1992; Canfield and Teske, 1996; Rye and Holland, 1998). The numbers on the bars refer to the number of metallo-enzymes containing a given metal inferred to be in use during that time (for example, 88

TABLE 3

Rules used to assign the timing of evolution for individual metallo-enzymes.

| | |
|---|---|
| 1 | Enzymes involved in basic cellular machinery, such as those involved in DNA repair and replication, RNA transcription, protein translation, respiration and electron transport, and carbon assimilation are considered to have been present in very early life. |
| 2 | Enzymes involved in specific microbial metabolisms (such as methanogenesis, bacterial sulfate reduction, nitrogen fixation, and oxygenic photosynthesis) are considered to have evolved at the same time as the metabolism. |
| 3 | Enzymes involved in other anaerobic microbial metabolisms, such as anaerobic respiration and fermentation, are considered ancient. |
| 4 | Enzymes involved in aerobic metabolisms and iron acquisition, enzymes requiring or transforming molecular oxygen or oxygen radicals (such as H ₂ O ₂), and enzymes used to transform oxidized nitrogen and sulfur species, are considered to have evolved after the rise of oxygen to near atmospheric levels at ~2.2 billion years ago. |
| 5 | Enzymes that are only found in plants or animals are considered to have evolved after the diversification of multicellular organisms at 542 million years ago. |
| 6 | Enzymes that have only been documented in one species of Bacteria or Archaea, or enzymes that were not found in any of the organisms considered in this study, are considered to have evolved recently, unless enzymatic function suggested otherwise. |

(Frausto da Silva and Williams, 2001). These functions might have evolved later in life after organisms began to produce molecular oxygen and learn to deal with free oxygen radicals. Furthermore, over some pH ranges manganese minerals tend to have higher solubility than iron minerals even under oxidizing conditions.

Biological uses of other metals (Zn, Co, Ni, W, and V) do not seem to have fluctuated significantly through time (fig. 9). This may be because organisms established very specific roles for these metals early in life, or because the metals have been subsequently replaced in metallo-enzymes by more bioavailable cofactors. Zinc, for example, is used in many ubiquitous enzymes such as metallopeptidases, metallophosphatases, and other zinc polymerases involved in DNA and RNA binding and synthesis (Lipscomb and Sträter, 1996). Zinc is readily available in modern freshwater and marine systems, where it can exist at up to 0.45 micromolar concentrations, nearly half of which is the bioavailable Zn²⁺ ion (Turner and others, 1981; Stumm and Morgan, 1996). However Saito and others (2003) calculate that zinc would have been scarce in early anoxic oceans, and suggest that during this time prokaryotes (specifically

Fig. 9. (A) (continued). Fe-containing enzymes were considered to exist on Earth today), and the thickness of the bars is scaled to reflect the proportions of each metal used among total metal-containing enzymes at any given time. (B) Approximate concentrations of the considered metals and other species in seawater with progressive oxygenation of the oceans, based on assumptions made by Saito and others (2003), except molybdenum concentrations, which were derived from other sources (Bertine and Turekian, 1973; Collier, 1985; Lewis and Landing, 1992; Stumm and Morgan, 1996; Morford and Emerson, 1999; Anbar and Knoll, 2002). The hatched area refers to the period of time during which the Earth's oceans likely experienced a prolonged period of deep water euxinia (based on the most recent dates from Poulton and others, 2004).

cyanobacteria) either used less zinc due to its scarcity or made use of alternate metal centers. They further suggest that organisms may have obtained zinc enzymes later through lateral gene transfer.

Cobalt and nickel-containing enzymes are most commonly used in metabolisms based on methane, carbon monoxide, and hydrogen, such as those likely operating on early Earth. Nickel is rarely used in higher organisms, and most nickel enzymes are confined to anaerobic prokaryotes, especially methanogens. The main biological use of cobalt is as the cobalamin cofactor of vitamin B12, which it is speculated to have evolved 2.7 to 3.5 billion years ago based on its synthesis without an oxidative step (Scott, 1998, 2001, 2003). Therefore the use of cobalt in B12 likely evolved in early life when Co was more highly available than today (fig. 9B). Cobalt can also replace zinc in the carbonic anhydrase of some eukaryotic phytoplankton (Roberts and others, 1997), and possibly in *Synechococcus* (Smith and others, 1999).

From figure 9, we calculated a model metallome for the LCA, as before, based on the number of atoms of each metal required per metallo-enzyme inferred to be present in early life. This model metallome for the LCA can be compared to our previous model metallome for the LCA based on phylogenetic extrapolation from modern genome sequences, from figure 7. The enzyme-predicted model metallome (fig. 8A) suggests quite a different set of trace metal requirements for the LCA than the model metallome based upon the genome (fig. 8B). The enzyme-predicted model metallome has a much higher F_{Zn} , with lower F_i for all other trace metals examined, especially copper and molybdenum. However, the model metallome based upon the genome is much more similar to the average model metallome we calculated from genomes of modern anaerobes in this study (fig. 4 and fig. 8C). This similarity could be a consequence of the technique used to extrapolate the model genome-based metallome of the LCA. Or, if the second model of metallome evolution based on the known functions of modern metallo-enzymes is more accurate, this could illustrate the point that the modern prokaryotic genome has been greatly affected by gene loss, lateral gene transfer, and/or substitution of alternate metal centers with changing metal availability. Gene loss can occur in organisms with lack of necessity for expression with metabolic function. Analysis of proteins and whole genome sequences in modern prokaryotes support the idea that lateral gene transfer has extensively affected the genetic composition of early life (for example Gogarten and others, 2002). Additionally, numerous metallo-enzymes form functional analogues with metal substitutions, including nitrogenase, superoxide dismutase, and carbonic anhydrase.

CONCLUSIONS

We utilized the information contained within whole genome sequences to investigate the modern distribution of two potentially important microbial biosignatures (carbon isotopic signatures and cellular metal content), and to interpret their evolution and possible distribution in early life. Specifically, we used the gene content of prokaryotic genomes to refine the topology of the tree of life, to predict cellular metal requirements in modern organisms, and to determine how metal use and biological carbon isotope fractionation have evolved through geologic time.

Previous work indicates that the reductive TCA and 3-hydroxypropionate cycles result in low carbon isotope fractionation between CO_2 and biomass, the reductive pentose-phosphate cycle and acetyl-CoA pathway result in moderate fractionation between CO_2 and biomass, and pathways using other lighter substrates result in extremely ^{13}C depleted biomass. The distribution of organisms using these carbon fixation pathways on microbial phylogenies suggests that: (1) both low and moderate carbon isotope fractionation between inorganic carbon dioxide and cellular biomass were established quickly in the early evolution of life; and (2) methanotrophic and ethanotrophic metabolisms capable of producing biomass with extreme ^{13}C depletions

due to growth on a fractionated carbon substrates are not primitive, but rather evolved after the major groups of Prokaryotes had already diverged. The most likely carbon fixation pathway for the last common ancestor is the universally important TCA cycle, which produces low carbon isotopic fractionation in biomass. Low fractionation by the reductive TCA cycle can be considered consistent with carbon isotope ratios found in 3,800 million year old Isua sediments. Additionally, moderately fractionated biomass from up to 3,500 million years ago can now likely be attributed to carbon fixation by anoxygenic photoautotrophs using the reductive pentose phosphate cycle.

To investigate microbial trace metal use from genome sequences, we calculated "model metallomes" for 52 prokaryotes based on the number of atoms of trace metals required to express one molecule of each metallo-enzyme coded in their corresponding genomes. Our results suggest that the use of metals in prokaryotes follows the general hierarchy: Fe \gg Zn > Mn \gg Mo, Co, Cu \gg Ni > W, V. Model metallomes vary with metabolism, oxygen tolerance, optimum growth temperature, and phylogeny. The model metallome of methanogens contains a unique metal signature, suggesting that elevated requirements of nickel and tungsten may be translated to the expressed metallome, providing a biosignature for methanogenesis. The model metallomes of diazotrophs and cyanobacteria do not contain unique metal signatures; however changes in enzyme expression under some conditions might still translate into a biosignature in the expressed cellular metal content.

Metallome comparisons also indicate that the model metallomes of aerobic organisms generally require more copper than those of anaerobic organisms, while tungsten and molybdenum show increasing importance in model metallomes with increasing optimum growth temperature. Most taxonomic groups we tested have statistically significant model metallomes; however, individual groups of Archaea have much more widely variant model metallomes than individual groups of Bacteria. Finally, we estimated a model metallome for the LCA by extrapolating model metallomes of modern phylogenetic groups on a tree of life based on our analysis of composite organisms along with the most widely accepted tree topologies. This extrapolated model metallome for the LCA is very similar to the average model metallome of anaerobic organisms calculated in this study.

In a separate analysis, we made inferences about the timing of the evolution of individual metallo-enzymes based on their function and occurrence in modern organisms. The results suggest that fluctuations in the redox state of the Earth's oceans and atmosphere have forced changes in the proportions of metals used in biology. In this model, biological use of copper and molybdenum has developed along with bioavailability; biological use of iron and manganese has developed counter to bioavailability; and biological use of zinc, cobalt, and nickel has not changed significantly through time. The model metallome of the LCA suggested by this method is much different than that extrapolated from modern genomes, supporting the idea that gene loss, metal substitution, and lateral gene transfer have been important in shaping the enzymatic composition of extant organisms.

ACKNOWLEDGMENTS

Graduate support for this project was provided by the Penn State Biogeochemical Research Initiative for Education (BRIE), sponsored by NSF (IGERT) grant DGE-9972759. The work was also funded by the Penn State Astrobiology Research Center, through the NASA Astrobiology Institute. We thank L. Kump, H. Carrick, L. Liermann, and two anonymous reviewers for useful discussion and comments.

REFERENCES

Altekar, W., and Rajagopalan, R., 1990, Ribulose biphosphate carboxylase activity in halophilic Archaeobacteria: Archives of Microbiology, v. 153, p. 169–174.

- Anbar, A. D., and Knoll, A. H., 2002, Proterozoic ocean chemistry and evolution: A bioinorganic bridge?: *Science*, v. 297, p. 1137–1142.
- Andreesen, J. R., and Gottschalk, G., 1969, The occurrence of a modified Entner-doudoroff pathway in *Clostridium acetium*: *Archives of Microbiology*, v. 69, p. 160–170.
- Arnold, G. L., Anbar, A. D., Barling, J., and Lyons, T. W., 2004, Molybdenum isotope evidence for widespread anoxia in mid-proterozoic oceans: *Science*, v. 304, p. 87–90.
- Bansal, A. K., and Meyer, T. E., 2002, Evolutionary analysis by whole-genome comparisons: *Journal of Bacteriology*, v. 184, p. 2260–2272.
- Battistuzzi, F. U., Feijao, A., and Hedges, S. B., 2004, A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land: *BMC Evolutionary Biology*, v. 4, p. 43.
- Beaumont, V., and Robert, F., 1999, Nitrogen isotope ratios of kerogens in Precambrian cherts: a record of the evolution of atmosphere chemistry?: *Precambrian Research*, v. 96, p. 63–82.
- Beh, M., Strauss, R., Huber, R., Stetter, K. O., and Fuchs, G., 1993, Enzymes of the reductive citric acid cycle in the autotrophic eubacterium *Aquifex pyrophilus* and in the archaeobacterium *Thermoproteus neutrophilus*: *Archives of Microbiology*, v. 160, p. 306–311.
- Belyaev, S. S., Wolkin, R., Kenealy, W. R., DeNiro, M. J., Epstein, S., and Zeikus, J. G., 1983, Methanogenic bacteria from the Bondyuzhskoe Oil field: general characterization and analysis of stable-carbon isotopic fractionation: *Applied and Environmental Microbiology*, v. 45, p. 691–697.
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E., 2000, The Protein Data Bank: *Nucleic Acids Research*, v. 28, p. 235–242.
- Bertine, K. K., and Turekian, K. K., 1973, Molybdenum in marine deposits: *Geochimica et Cosmochimica Acta*, v. 48, p. 605–615.
- Blank, C. E., 2004, Evolutionary timing of the origins of mesophilic sulphate reduction and oxygenic photosynthesis: a phylogenomic dating approach: *Geobiology*, v. 2, p. 1–20.
- Bokova, E. N., 1954, Oxidation of ethane and propane by certain species of mycobacteria: *Mikrobiologiya*, v. 23, p. 15–21.
- Botz, R., Pokojski, H. D., Schmitt, M., and Thomm, M., 1996, Carbon isotope fractionation during bacterial methanogenesis by CO₂ reduction: *Organic Geochemistry*, v. 25, p. 255–262.
- Brantley, S. L., Liermann, L., Bau, M., and Wu, S., 2001, Uptake of trace metals and rare earth elements from hornblende by a soil bacterium: *Geomicrobiology Journal*, v. 18, p. 37–61.
- Brochier, C., Baptiste, E., Moreira, D., and Philippe, H., 2002, Eubacterial phylogeny based on translational apparatus proteins: *Trends in Genetics*, v. 18, p. 1–5.
- Brock, J. J., Logan, G. A., Buick, R., and Summons, R. E., 1999, Archean molecular fossils and the early rise of eukaryotes: *Science*, v. 285, p. 1033–1036.
- Brown, J. R., Donady, C. J., Italia, M. J., Marshall, W. E., and Stanhope, M. J., 2001, Universal trees based on large combined protein sequence data sets: *Nature Genetics*, v. 28, p. 281–285.
- Bruland, K. W., 1980, Oceanographic distributions of cadmium, zinc, nickel, and copper in the North Pacific: *Earth and Planetary Science Letters*, v. 47, p. 176–198.
- 1989, Complexation of Zinc by Natural Organic-Ligands in the Central North Pacific: *Limnology and Oceanography*, v. 34, p. 269–285.
- Bruland, K. W., Donat, J. R., and Hutchins, D. A., 1991, Interactive influences of bioactive trace metals on biological production in oceanic waters: *Limnology and Oceanography*, v. 36, p. 1555–1577.
- Burgess, B. K., and Lowe, D. J., 1996, Mechanism of molybdenum nitrogenase: *Chemical Review*, v. 96, p. 2983–3011.
- Calder, J. A., and Parker, P. L., 1973, Geochemical implications of induced changes in C13 fractionation by blue-green algae: *Geochimica et Cosmochimica Acta*, v. 37, p. 133–140.
- Cameron, V., Zhang, Z., House, C. H., and Brantley, S. L., 2004, A search for hydrothermal tungsten ligands, in Wanty, R., and Seal, R., editors, *Proceedings of the 11th International Symposium on Water-Rock Interaction*: London, Taylor and Francis Group, p. 1269–1273.
- Canfield, D. E., 1998, A new model for Proterozoic ocean chemistry: *Nature*, v. 396, p. 450–453.
- Canfield, D. E., and Teske, A., 1996, Late Proterozoic rise in atmospheric oxygen concentration inferred from phylogenetic and sulphur-isotope studies: *Nature*, v. 382, p. 127–132.
- Chan, M. K., Mukund, S., Kletzin, A., Adams, M. W. W., and Rees, D. C., 1995, Structure of a hyperthermophilic tungstopterin enzyme, aldehyde ferredoxin oxidoreductase: *Science*, v. 267, p. 1462–1469.
- Clarke, G. D., Beiko, R. G., Ragan, M. A., and Charlebois, R. I., 2002, Inferring genome trees by using a filter to eliminate phylogenetically discord sequences and a distance matrix based on mean normalized BLASTP scores: *Journal of Bacteriology*, v. 184, p. 2072–2080.
- Collier, R. W., 1985, Molybdenum in the Northeast Pacific Ocean: *Limnology and Oceanography*, v. 30, p. 1351–1354.
- Davis, J. B., Chase, H. H., and Raymond, R. L., 1956, *Mycobacterium paraffinicum* n. sp., a bacterium isolated from soil: *Applied Microbiology*, v. 4, p. 310–315.
- de Bont, J. A., 1976, Oxidation of ethylene by soil bacteria: *Antonie van Leeuwenhoek*, v. 42, p. 59–71.
- Delwiche, C. F., and Palmer, J. D., 1996, Rampant horizontal transfer and duplication of rubisco genes in eubacteria and plastids: *Molecular Biology and Evolution*, v. 13, p. 873–882.
- Des Marais, D. J., Strauss, H., Summons, R. E., and Hayes, J. M., 1992, Carbon isotope evidence for the stepwise oxidation of the Proterozoic environment: *Nature*, v. 359, p. 605–609.
- D'Hondt, S. L., Jorgensen, B. B., Miller, D. J., et al., 2003, Leg 201 Summary: *Proceedings of the Ocean Drilling Program, Initial Reports*: College Station, Texas, Ocean Drilling Program, v. 201, p. 1–81.
- Dismukes, G. C., 1996, Manganese enzymes with binuclear active sites: *Chemical Review*, v. 96, p. 2909–2926.
- Donat, J. R., and Bruland, K. W., 1995, Trace Elements in the Oceans, in Salbu, B., and Steinnes, E., editors, *Trace Elements in Natural Waters*: Boca Raton, CRC Press, p. 247–281.

- Eady, R. R., 1996, Structure-function relationships of alternative nitrogenases: *Chemical Review*, v. 96, p. 3013–3030.
- Eitenger, T., and Friedrich, B., 1997, Microbial nickel transport and incorporation into hydrogenases, in Winklermann, G., and Carrano, C. J., editors, *Transition Metals in Microbial Metabolism*: Reading, United Kingdom, Hardwood Academic Publishers, p. 235–256.
- Emerson, S. R., and Husted, S. S., 1991, Ocean anoxia and the concentrations of molybdenum and vanadium in seawater: *Marine Chemistry*, v. 34, p. 177–196.
- Felsenstein, J., 1993, PHYLIP (Phylogeny Inference Package), Distributed by the Author: Department of Genetics, University of Washington, Seattle.
- Finn, M. W., and Tabita, F. R., 2003, Synthesis of catalytically active form III ribulose 1,5-biphosphate carboxylase/oxygenase in archaea: *Journal of Bacteriology*, v. 185, p. 3049–3059.
- Fitz-Gibbon, S. T., and House, C. H., 1999, Whole genome-based phylogenetic analysis of free-living microorganisms: *Nucleic Acids Research*, v. 27, p. 4218–4222.
- Fox, G. E., Stackebrandt, E., Hespell, R. B., Gibson, J., Maniloff, J., Dyer, T. A., Wolfe, R. S., Balch, W. E., Tanner, R. S., Magrum, L. J., Zablen, L. B., Blakemore, R. P., Gupta, R. S., Bunsen, L., Lewis, B. J., Stahl, D. A., Luehrs, K. R., Chen, K. N., and Woese, C. R., 1980, The phylogeny of prokaryotes: *Science*, v. 209, p. 457–463.
- Frausto da Silva, J. J. R., and Williams, R. J. P., 2001, *The Biological Chemistry of the Elements*: New York, Oxford University Press, 575 p.
- Fu, L. M., and Fu-Liu, C. S., 2002, Is *Mycobacterium tuberculosis* a closer relative to Gram-positive or Gram-negative bacterial pathogens?: *Tuberculosis*, v. 82, p. 85–90.
- Fuchs, G., and Stupperich, E., 1978, Evidence for an incomplete reductive carboxylic acid cycle in *Methanobacterium thermoautotrophicum*: *Archives of Microbiology*, v. 118, p. 121–125.
- Fuchs, G., Thauer, R., Ziegler, H., and Stüchler, W., 1979, Carbon isotope fractionation by *Methanobacterium thermoautotrophicum*: *Archives of Microbiology*, v. 120, p. 135–139.
- Fuchs, G., Stupperich, E., and Eden, G., 1980, Autotrophic CO₂ fixation in *Chlorobium limicola*. Evidence for the operation of a reductive tricarboxylic acid cycle in growing cells: *Archives of Microbiology*, v. 151, p. 64–71.
- Gerstein, M., 1998, Patterns of protein-fold usage in eight microbial genomes: a comprehensive structural census: *PROTEINS: Structure, Function, and Genetics*, v. 33, p. 518–534.
- Gerstein, M., and Hegyi, H., 1998, Comparing genomes in terms of protein structure: surveys of a finite parts list: *FEMS Microbiology Reviews*, v. 22, p. 277–304.
- Gogarten, J. P., Doolittle, W. F., and Lawrence, J. G., 2002, Prokaryotic evolution in light of gene transfer: *Molecular Biology Evolution*, v. 19, p. 2226–2238.
- Golubic, S., Sergeev, V. N., and Knoll, A. H., 1995, Mesoproterozoic *Archaeoellipsoides*: akinetes of heterocystous cyanobacteria: *Lethaia*, v. 28, p. 285–298.
- Gordon, R. M., Martin, J. H., and Knauer, G. A., 1982, Iron in north-east Pacific waters: *Nature*, v. 299, p. 611–612.
- Gorlenko, V. M., and Pivovarova, T. A., 1977, On the assignment of the blue-green alga *Oscillatoria coerulea* Gickhorn 1921 to the new genus of Chlorobacteria *Oscillochloris* nov. gen.: *Proceedings of the Academy of Sciences of the Estonian SSR: Biology*, v. 3, p. 396–409.
- Guss, J. M., and Freeman, H., 1983, The blue copper electron transfer proteins: *Journal of Molecular Biology*, v. 169, p. 521–563.
- Habicht, K. S., Gade, M., Thamdrup, B., Berg, P., and Canfield, D. E., 2002, Calibration of sulfate levels in the Archean ocean: *Science*, v. 298, p. 2372–2374.
- Hallam, S. J., Girguis, P. R., Preston, C. M., Richardson, P. M., and DeLong, E. F., 2003, Identification of methyl coenzyme M reductase A (mcrA) genes associated with methane-oxidizing bacteria: *Applied and Environmental Microbiology*, v. 69, p. 5483–5491.
- Hansmann, S., and Martin, W., 2000, Phylogeny of 33 ribosomal and six other proteins encoded in an ancient gene cluster that is conserved across prokaryotic genomes: influence of excluding poorly alignable sites from analysis: *International Journal of Systematic and Evolutionary Microbiology*, v. 50, p. 1655–1663.
- Hanson, R. S., and Hanson, T. E., 1996, Methanotrophic bacteria: *Microbiological Reviews*, v. 60, p. 439.
- Hanson, T. E., and Tabita, F. R., 2001, A ribulose-1,5-biphosphate carboxylase/oxygenase (RubisCO)-like protein from *Chlorobium tepidum* that is involved with sulfur metabolism and the response to oxidative stress: *Proceedings of the National Academy of Sciences*, v. 98, p. 4397–4402.
- Hausinger, R. P., 1987, Nickel utilization in microorganisms: *Microbiological Reviews*, v. 51, p. 22–42.
- Hayes, J. M., 2002, Practice and principles of isotopic measurements in organic geochemistry, revision 2: http://www.nosams.who.edu/research/staff_hayes.html.
- Hayes, J. M., Kaplan, I. R., and Wedeking, K. W., 1983, Precambrian organic geochemistry, preservation of the record, in Schopf, J. W., editor, *Earth's Earliest Biosphere*: Princeton, Princeton University Press, p. 93–134.
- Hayes, J. M., Des Marais, I. B., Lambert, H., Strauss, H., and Summons, R. E., 1992, Proterozoic biogeochemistry, in Schopf, J. W., and Klein, C., editors, *The Proterozoic Biosphere*: New York, Cambridge University Press, p. 81–134.
- Hersman, L., Lloyd, T., and Sposito, G., 1995, Siderophore-promoted dissolution of hematite: *Geochimica et Cosmochimica Acta*, v. 59, p. 3327–3330.
- Hider, R. C., 1984, Siderophore mediated absorption of Fe, in Clarke, M. J., Ibers, J. A., Mingos, D. M. P., Palmer, G. A., Sadler, P. J., and Williams, R. J. P., editors, *Structure and bonding: siderophores from microorganisms and plants*: Berlin, Springer-Verlag, v. 58, p. 25–87.
- Hille, R., 1996, The mononuclear molybdenum enzymes: *Chemical Review*, v. 96, p. 2757–2816.

- Holland, H. D., 1973, The Oceans: A possible source of iron in iron-formations: *Economic Geology and the Bulletin of the Society of Economic Geologists*, v. 68, p. 1169–1172.
- 1984, *The Chemical Evolution of the Atmosphere and Oceans*: Princeton, New Jersey, Princeton University Press, 560 p.
- Holland, H. D., and Beukes, N. J., 1990, A paleoweathering profile from Griqualand West, South Africa; evidence for a dramatic rise in atmospheric oxygen between 2.2 and 1.9 yBP: *American Journal of Science*, v. 290-A, p. 1–34.
- Holm, R. H., Kennepohl, P., and Solomon, E. I., 1996, Structural and functional aspects of metal sites in biology: *Chemical Review*, v. 96, p. 2239–2314.
- Holo, H., 1989, *Chloroflexus aurantiacus* secretes 3-hydroxypropionate, a possible intermediate in the assimilation of CO₂ and acetate: *Archives of Microbiology*, v. 151, p. 252–256.
- Holo, H., and Sirevåg, R., 1986, Autotrophic growth and CO₂ fixation of *Chloroflexus aurantiacus*: *Archives of Microbiology*, v. 145, p. 173–180.
- House, C. H., ms, 1999, Carbon isotopic fractionation by diverse extant and fossil prokaryotes and microbial diversity revealed through genomics: Ph.D. Dissertation, The University of California at Los Angeles, 208 p.
- House, C. H., and Fitz-Gibbon, S. T., 2002, Using homolog groups to create a whole-genomic tree of free-living organisms: An update: *Journal of Molecular Evolution*, v. 54, p. 539–547.
- House, C. H., Runnegar, B., and Fitz-gibbon, S. T., 2003b, Geobiological analysis using whole-genome based tree building applied to the Bacteria, Archaea, and Eukarya: *Geobiology*, v. 1, p. 15–26.
- House, C. H., Schopf, J. W., McKeegan, K. D., Coath, C. D., Harrison, T. M., and Stetter, K. O., 2000, The carbon isotopic composition of individual Precambrian microfossils: *Geology*, v. 28, p. 707–710.
- House, C. H., Schopf, J. W., and Stetter, K. O., 2003a, Carbon isotopic fractionation by Archaeans and other thermophilic prokaryotes: *Organic Geochemistry*, v. 34, p. 345–356.
- Howard, J. B., and Rees, D. C., 1996, Structural basis of biological nitrogen fixation: *Chemical Review*, v. 96, p. 2965–2982.
- Huber, C., and Wächtershauser, G. 1997, Activated acetic acid by carbon fixation on (Fe, Ni)S under primordial conditions: *Science*, v. 276, p. 245–247.
- Hughes, M. N., and Poole, R. K., 1989, *Metals and Microorganisms*: London, Chapman and Hall, 412 p.
- Hügler, M., Huber, H., Stetter, K. O., and Fuchs, G., 2003, Autotrophic CO₂ fixation pathways in archaea (Crenarchaeota): *Archives of Microbiology*, v. 179, p. 160–173.
- Ishii, M. T., 1998, Reductive TCA cycle - special emphasis on the CO₂ and energy metabolism in a thermophilic, aerobic and obligately autotrophic hydrogen-oxidizing bacterium *Hydrogenobacter thermophilus* strain TK-6: *Nippon Nogeikagaku Kaishi*, v. 72, p. 1165–1170.
- Ishii, M. T., Miyake, T., Satoh, T., Sugiyama, H., Oshima, Y., Kodama, T., and Igarashi, Y., 1996, Autotrophic carbon dioxide fixation in *Acidianus brierleyi*: *Archives of Microbiology*, v. 166, p. 368–371.
- Ivanovsky, R. N., Fal, Y. I., Berg, I. A. Ugolkova, N. V., Krasilnikova, E. N., Keppen, O. I., Zakharchuc, L. M., and Zyakun, A. M., 1999, Evidence for the presence of the reductive pentose phosphate cycle in a filamentous anoxygenic photosynthetic bacterium, *Oscillochloris trichoides* strain DG-6: *Microbiology*, v. 145, p. 1743–1748.
- Jahnke, L. L., Eder, W., Huber, R., Hope, J. M., Hinrichs, K. U., Hayes, J. M., Des Marais, D. J., Cady, S. L., and Summons, R. E., 2001, Signature lipids and stable carbon isotope analysis of Octopus Spring hydrothermophilic communities compared with those of Aquificales representatives: *Applied and Environmental Microbiology*, v. 67, p. 5179–5189.
- Jahnke, L. L., Summons, R. E., Hope, J. M., and Des Marais, D. J., 1999, Carbon isotope fractionation in lipids from methanotrophic bacteria II: The effects of physiology and environmental parameters on the biosynthesis and isotopic signatures of biomarkers: *Geochimica et Cosmochimica Acta*, v. 63, p. 79–93.
- Johnson, M. K., Rees, D. C., and Adams, M. W. W., 1996, Tungstoenzymes: *Chemical Review*, v. 96, p. 2817–2839.
- Kalinowski, B. E., Liermann, L. J., Givens, S., and Brantley, S. L., 2000, Rates of bacteria-promoted solubilization of Fe from minerals: a review of problems and approaches: *Chemical Geology*, v. 169, p. 357–370.
- Kandler, O., 1994, The early diversification of life, in Benston, S., editor, *Early Life on Earth*: New York, Columbia University Press, p. 152–160.
- Kaufman, A. J., and Xiao, S., 2003, High CO₂ levels in the Proterozoic atmosphere estimated from analyses of individual microfossils: *Nature*, v. 425, p. 279–282.
- Keppen, O. I., Baulina, O. I., and Kondratieva, E. N., 1994, *Oscillochloris trichoides* neotype strain DG-6: *Photosynthesis Research*, v. 41, p. 29–33.
- Keppen, O. I., Tourova, T. P., Kuznetsov, B. B., Ivanovsky, R. N., and Gorlenko, V. M., 2000, Proposal of Oscillochloridaceae fam. nov. on the basis of a phylogenetic analysis of the filamentous anoxygenic phototrophic bacteria, and emended description of *Oscillochloris* and *Oscillochloris trichoides* in comparison with further new isolates: *International Journal of Systematic Evolutionary Microbiology*, v. 50, p. 1529–1537.
- Kisker, C., Schindelin, H., and Rees, D. C., 1997, Molybdenum-cofactor-containing enzymes: Structure and mechanism: *Annual Review in Biochemistry*, v. 66, p. 233–267.
- Klappenback, J. A., and Pierson, B. K., 2004, Phylogenetic and physiological characterization of a filamentous anoxygenic photoautotrophic bacterium '*Candidatus Chlorothrix halophila*' gen. nov., sp. nov., recovered from hypersaline microbial mats: *Archives of Microbiology*, v. 181, p. 17–25.
- Kletzin, A., Mukund, S., Kelley-Crouse, T. L., Chan, M. K., Rees, D. C., and Adams, M. W. W., 1995, Molecular characterization of the genes encoding the tungsten-containing aldehyde ferredoxin oxidoreductase from *Pyrococcus furiosus* and formaldehyde ferredoxin oxidoreductase from *Thermococcus litoralis*: *Journal of Bacteriology*, v. 177, p. 4817–4819.

- Knoll, A. H., 2003, The geological consequences of evolution: *Geobiology*, v. 1, p. 3–14.
- Knoll, A. H., and Canfield, D. E., 1998, Isotopic inferences on early ecosystems: *The Paleontological Society Papers*, v. 4, p. 211–243.
- Knoll, A. H., Hayes, J. M., Kaufman, A. J., Swett, K., and Lambert, I. B., 1986, Secular variation in carbon isotope ratios from upper Proterozoic successions of Svalbard and East Greenland: *Nature*, v. 321, p. 832–838.
- Korbel, J. O., Snel, B., Huynen, M. A., and Bork, P., 2002, SHOT: a web server for the construction of genome phylogenies: *Trends in Genetics*, v. 18, p. 158–162.
- Krüger, M., Meyerdierks, A., Glockner, F. O., Amann, R., Widdel, F., Kube, M., Reinhardt, R., Kahnt, J., Bocher, R., Thauer, R. K., and Shima, S., 2003, A conspicuous nickel protein in microbial mats that oxidize methane anaerobically: *Nature*, v. 426, p. 878–881.
- Landing, W. M., and Bruland, K. W., 1987, The contrasting biogeochemistry of iron and manganese in the Pacific Ocean: *Geochimica et Cosmochimica Acta*, v. 51, p. 29–43.
- Lepland, A., van Zuilten, M. A., Arrhenius, G., Whitehouse, M. J., and Fedo, C. M., 2005, Questioning the evidence for Earth's earliest life – Akilia revisited: *Geology*, v. 33, p. 77–79.
- Lewis, B. L., and Landing, W. M., 1991, The biogeochemistry of manganese and iron in the Black Sea: *Deep-Sea Research*, v. 38, p. 773–803.
- 1992, The investigation of dissolved and suspended-particulate trace metal fractionation in the Black Sea, in Millero, F. J., editor, *Marine Chemistry*: Amsterdam, Elsevier, p. 105–141.
- Li, W., Fang, W., Ling, W., Wang, J., Xuan, Z., and Chen, R., 2002, Phylogeny based on whole genome as inferred from complete information set analysis: *Journal of Biological Physics*, v. 28, p. 439–447.
- Liermann, L. J., Kalinowski, B. E., Brantley, S. L., and Ferry, J. G., 2000, Role of bacterial siderophores in dissolution of hornblende: *Geochimica et Cosmochimica Acta*, v. 64, p. 587–602.
- Lin, J., and Gerstein, M., 2000, Whole genome trees based on the occurrence of folds and orthologs: implications for comparing genomes on different levels: *Genome Research*, v. 10, p. 808–818.
- Lipscomb, W. N., and Sträter, N., 1996, Recent advances in zinc enzymology: *Chemical Review*, v. 96, p. 2375–2433.
- Londry, K. L., and Des Marais, D. J., 2003, Stable carbon isotope fractionation by sulfate-reducing bacteria: *Applied and Environmental Microbiology*, v. 69, p. 2942–2949.
- Londry, K. L., Jahnke, L. L., and Des Marais, D. J., 2004, Stable carbon isotope ratios of lipid biomarkers of sulfate-reducing bacteria: *Applied and Environmental Microbiology*, v. 70, p. 745–751.
- L'vov, N. P., Nosikov, A. N., and Antipov, A. N., 2002, Tungsten-containing enzymes: *Biochemistry (Moscow)*, v. 67, p. 196–200.
- Macalady, J., and Banfield, J. F., 2003, Molecular geomicrobiology: genes and geochemical cycling: *Earth and Planetary Science Letters*, v. 209, p. 1–17.
- Madigan, M. T., Takigiku, R., Lee, R. G., Gest, H., and Hayes, J. M., 1989, Carbon isotope fractionation by thermophilic phototrophic sulfur bacteria: Evidence for autotrophic growth in natural populations: *Applied and Environmental Microbiology*, v. 55, p. 639–644.
- Marsh, E. N., 1999, Coenzyme B12-dependent enzymes: *Essays in Biochemistry*, v. 34, p. 139–154.
- Martin, J. H., and Gordon, R. M., 1988, Northeast Pacific iron distributions in relation to phytoplankton productivity: *Deep-Sea Research*, v. 35, p. 177–196.
- Martin, J. H., Gordon, R. M., Fitzwater, S., and Broenkow, W. W., 1989, Vertex - Phytoplankton Iron Studies in the Gulf of Alaska: *Deep-Sea Research Part A-Oceanographic Research Papers*, v. 36, p. 649–680.
- Martin, W., and Russell, M. J., 2004, The rocky roots of the acetyl-CoA pathway: *Trends in Biochemical Sciences*, v. 29, p. 358–363.
- Martinez-Argudo, I., Little, R., Shearer, N., Johnson, P., and Dixon, R., 2004, The NifL-NifA system: a multidomain transcriptional regulatory complex that integrates environmental signals: *Journal of Bacteriology*, v. 186, p. 601–610.
- Matte-Tailliez, O., Brochier, C., Forterre, P., and Philippe, H., 2002, Archaeal phylogeny based on ribosomal proteins: *Molecular Biology and Evolution*, v. 19, p. 631–639.
- Meijer, W. G., Arnberg, A. C., Enequist, H. G., Terpstra, P., Lidstrom, M. E., and Dijkhuizen, L., 1991, Identification and organization of carbon dioxide fixation genes in *Xanthobacter flavus* H4-14: *Molecular and General Genetics*, v. 225, p. 320–330.
- Menendez, C., Bauer, Z., Huber, H., Gad'on, N., Stetter, K. O., and Fuchs, G., 1999, Presence of acetyl coenzyme A (CoA) carboxylase and propionyl-CoA carboxylase in autotrophic crenarchaeota and indication for operation of a 3-hydroxypropionate cycle in autotrophic carbon fixation: *Journal of Bacteriology*, v. 181, p. 1088–1098.
- Mirkin, E., and Ruby, E. G., 1991, Stable carbon isotope fractionation in cultured chemoautotrophic bacteria: *General Meeting of the American Society for Microbiology*, v. 91, 199 p.
- Mizutani, H., and Wada, E., 1982, Effect of high atmospheric CO₂ concentration of δ¹³C of algae: *Origins of Life*, v. 12, p. 377–390.
- Mojzsis, S. J., Arrhenius, G., McKeegan, K. D., Harrison, T. M., Nutman, A. P., and Friend, C. R. L., 1996, Evidence for life on Earth before 3,800 million years ago: *Nature*, v. 384, p. 55–59.
- Mojzsis, S. J., Harrison, T. M., and Pidgeon, R. T., 2001, Oxygen-isotope evidence from ancient zircons for liquid water at the Earth's surface 4,300 Myr ago: *Nature*, v. 409, p. 178–181.
- Morel, F. M. M., and Price, N. M., 2003, The biogeochemical cycles of trace metals in the oceans: *Science*, v. 300, p. 944–947.
- Morford, J. L., and Emerson, S., 1999, The geochemistry of redox sensitive trace metals in sediments: *Geochimica et Cosmochimica Acta*, v. 63, p. 1735–1750.
- Morowitz, H. J., Kostelnik, J. D., Yang, J., and Cody, G. D., 2000, The origin of intermediary metabolism: *Proceedings of the National Academy of Sciences*, v. 97, p. 7704–7708.

- Mukund, S., and Adams, M. W. W., 1991, The novel tungsten-iron-sulfur protein of the hyperthermophilic archaeobacterium, *Pyrococcus furiosus*, is an aldehyde ferredoxin oxidoreductase - evidence for its participation in a unique glycolytic pathway: *Journal of Biological Chemistry*, v. 266, p. 14208–14216.
- Neilands, J. B., 1984, Methodology of siderophores, Structure and Bonding: Berlin, Springer-Verlag, p. 1–24.
- Normand, P., and Bousquet, J., 1989, Phylogeny of the nitrogenase sequences in Frankia and other nitrogen-fixing microorganisms: *Journal of Molecular Evolution*, v. 29, p. 436–447.
- Oremland, R. S., Whiticar, M. J., Strohmaier, F. E., and Keine, R. P., 1988, Bacterial ethane formation from reduced ethylated compounds in anoxic sediments: *Geochemica et Cosmochimica Acta*, v. 52, p. 1895–1904.
- Orphan, V. J., Hinrichs, K. U., Ussler, W., III, Paull, C. K., Taylor, L. T., Sylva, S. P., Hayes, J. M., and DeLong, E. F., 2001a, Comparative analysis of methane-oxidizing Archaea and sulfate-reducing bacteria in anoxic marine sediments: *Applied and Environmental Microbiology*, v. 67, p. 1922–1934.
- Orphan, V. J., House, C. H., Hinrichs, K. U., McKeegan, K. D., and DeLong, E. F., 2001b, Coupled isotopic and phylogenetic characterization of single cells: direct evidence for a methane-consuming archaeal/bacterial consortium: *Science*, v. 293, p. 484–487.
- 2002, Multiple archaeal groups mediate methane oxidation in anoxic cold seep sediments: *Proceedings of the National Academy of Sciences of the United States of America*, v. 99, p. 7663–7668.
- Orphan, V. J., Ussler, W. I., Naehr, T. H., House, C. H., Hinrichs, K. U., and Paull, C. K., 2004, Geological, geochemical, and microbiological heterogeneity of the seafloor around methane vents in the Eel River Basin, offshore California: *Chemical Geology*, v. 205, p. 265–289.
- Outten, C. E., and O'Halloran, T. V., 2001, Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis: *Science*, v. 292, p. 2488–2492.
- Pardue, J. W., Scalan, R. S., Van Baalen, C., and Parker, P. L., 1976, Maximum carbon isotope fractionation in photosynthesis by blue-green algae and a green alga: *Geochimica et Cosmochimica Acta*, v. 40, p. 309–312.
- Paull, C. K., Lorensen, T. D., Borowski, W. S., Ussler III, W., Olsen, K., and Rodriguez, N. M., 2000, Isotopic composition of CH₄, CO₂ species, and sedimentary organic matter within samples from the Blake Ridge: gas source implications, in Paull, C. K., Matsumoto, R., Wallace, P. J., Dillon, W. P., and others, editors, *Proceedings of the Ocean Drilling Program, Scientific Results 164: College Station, Texas, Ocean Drilling Program*, p. 67–78.
- Popp, B. N., Laws, E. A., Bidigare, R. P., Dore, J. E., Hanson, K. L., and Wakeham, S. G., 1998, Effect of phytoplankton cell geometry on carbon isotopic fractionation: *Geochimica et Cosmochimica Acta*, v. 62, p. 69–77.
- Poulton, S. W., Fralick, P. W., and Canfield, D. E., 2004, The transition to a sulphidic ocean ~1.84 billion years ago: *Nature*, v. 431, p. 173–177.
- Preuß, A., Schauder, R., Fuchs, G., and Stichler, W., 1989, Carbon isotope fractionation by autotrophic bacteria with three different CO₂ fixation pathways: *Zeitschrift für Naturforschung*, v. 44c, p. 397–402.
- Quandt, L., Gottschalk, H., Ziegler, H., and Stichler, W., 1977, Isotope discrimination by photosynthetic bacteria: *FEMS Microbiology Letters*, v. 1, p. 125–129.
- Ragsdale, S. W., and Kumar, M., 1996, Nickel-containing carbon monoxide dehydrogenase/Acetyl-CoA synthase: *Chemical Review*, v. 96, p. 2515–1539.
- Raven, J. A., 1988, The iron and molybdenum use efficiencies of plant growth with different energy, carbon and nitrogen sources: *New Phytologist*, v. 109, p. 279–287.
- 1990, Predictions of Mn and Fe use efficiencies of phototrophic growth as a function of light availability for growth and of C assimilation pathway: *New Phytologist*, v. 116, p. 1–18.
- 1995, The early evolution of land plants: Aquatic ancestors and atmospheric interactions: *Botanical Journal of Scotland*, v. 47, p. 151–175.
- Rawal, N., Kelkar, S. M., and Altekar, W., 1988, Ribulose 1,5-biphosphate dependent CO₂ fixation in the halophilic archaeobacterium, *Halobacterium mediterraneum*: *Biochemical and biophysical research communications*, v. 156, p. 451–456.
- Roberts, S., Lane, T., and Morel, F. M. M., 1997, Carbonic anhydrase in the marine diatom *Thalassiosira weissflogii* (Bacillariophyceae): *Journal of Phycology*, v. 33, p. 845–850.
- Robinson, J. J., and Cavanaugh, C. M., 1995, Expression of form I and form II Rubisco in chemoautotrophic symbioses: implications for the interpretation of stable carbon isotope values: *Limnology and Oceanography*, v. 40, p. 1496–1502.
- Roeske, C. A., and O'Leary, M. H., 1984, Carbon isotope effects on the enzyme-catalyzed carboxylation of ribulose biphosphate: *Biochemistry*, v. 23, p. 6275–6284.
- 1985, Carbon isotope effects on carboxylation of ribulose biphosphate catalyzed by ribulosebiphosphate carboxylase from *Rhodospirillum rubrum*: *Biochemistry*, v. 24, p. 1603–1607.
- Romano, A. H., and Conway, T., 1996, Evolution of carbohydrate metabolic pathways: *Research in Microbiology*, v. 147, p. 448–455.
- Rosing, M. T., 1999, C-13-depleted carbon microparticles in >3700-Ma sea-floor sedimentary rocks from west Greenland: *Science*, v. 283, p. 674–676.
- Ruby, E. G., Jannasch, H. W., and Deuser, W. G., 1987, Fractionation of stable carbon isotopes during chemoautotrophic growth of sulfur-oxidizing bacteria: *Applied and Environmental Microbiology*, v. 53, p. 1940–1943.
- Rue, E. L., and Bruland, K. W., 1995, Complexation of Iron (Iii) by Natural Organic-Ligands in the Central North Pacific as Determined by a New Competitive Ligand Equilibration Adsorptive Cathodic Stripping Voltammetric Method: *Marine Chemistry*, v. 50, p. 117–138.
- Ryder, G., 2003, Bombardment of the Hadean earth: Wholesome or deleterious: *Astrobiology*, v. 3, p. 3–6.
- Rye, R., and Holland, H. D., 1998, Paleosols and the evolution of atmospheric oxygen: a critical review: *American Journal of Science*, v. 298, p. 621–672.

- Saito, M. A., Sigman, D. M., and Morel, F. M. M., 2003, The bioinorganic chemistry of the ancient ocean: the co-evolution of cyanobacterial metal requirements and biogeochemical cycles at the Archean-Proterozoic boundary?: *Inorganica Chimica Acta*, v. 356, p. 308–318.
- Schauder, R., Widdel, F., and Fuchs, G., 1987, Carbon assimilation pathways in sulfate-reducing bacteria. II. Enzymes of a reductive citric acid cycle in the autotrophic *Desulfurobacter hydrogenophilus*: *Archives of Microbiology*, v. 148, p. 218–225.
- Schauder, R., Preuß, A., Jetten, M., and Fuchs, G., 1989, Oxidative and reductive acetyl-CoA/carbon monoxide dehydrogenase pathway in *Desulfobacterium autotrophicum*. 2. Demonstration of the enzymes of the pathway and comparison of CO dehydrogenase: *Archives of Microbiology*, v. 151, p. 84–89.
- Schidlowski, M., 2001, Carbon isotopes as biogeochemical recorders of life over 3.8 Ga of Earth history: evolution of a concept: *Precambrian Research*, v. 106, p. 117–134.
- Schidlowski, M., Hayes, J. M., and Kaplan, I. R., 1983, Isotopic Inferences of Ancient Biochemistries: Carbon, Sulfur, Hydrogen, and Nitrogen, in Schopf, J. W., editor, *The Earth's Earliest Biosphere*: Princeton, Princeton University Press, p. 149–185.
- Schopf, J. W., 1993, Microfossils of the Early Archean Apex chert: new evidence of the antiquity of life: *Science*, v. 260, p. 640–646.
- Schopf, J. W., and Packer, B. M., 1987, Early Archean (3.3-billion to 3.5-billion-year-old) microfossils from Warrawoona Group, Australia: *Science*, v. 237, p. 70–73.
- Schopf, J. W., Kudryavtsev, A. B., Agresti, D. G., Wdowiak, T. J., and Czaja, A. D., 2002, Laser-Raman imagery of Earth's earliest fossils: *Nature*, v. 416, p. 73–76.
- Schouten, S., Strous, M., Kuypers, M. M. M., Rijpstra, W. I. C., Baas, M., Schubert, C. J., Jetten, M. S. M., and Sinninghe Damste, J. S., 2004, Stable carbon isotopic fractionations associated with inorganic carbon fixation by anaerobic ammonium-oxidizing bacteria: *Applied and Environmental Microbiology*, v. 70, p. 3785–3788.
- Scott, A. I., 1998, How nature synthesizes B-12 without oxygen. Discoveries along the ancient, anaerobic pathway: *Heterocycles*, v. 47, p. 1051–1066.
- 2001, Reflections on the discovery of nature's pathways to vitamin B12: *Chemical Record*, v. 1, p. 212–227.
- 2003, Discovering nature's diverse pathways to vitamin B-12: A 35-year odyssey: *Journal of Organic Chemistry*, v. 68, p. 2529–2539.
- Shen, Y., Buick, R., and Canfield, D. E., 2001, Isotopic evidence for microbial sulphate reduction in the early Archaean era: *Nature*, v. 410, p. 77–81.
- Shen, Y., Canfield, D. E., and Knoll, A. H., 2002, Middle Proterozoic ocean chemistry: Evidence from the McArthur Basin, Northern Australia: *American Journal of Science*, v. 302, p. 81–109.
- Shen, Y., Knoll, A. H., and Walter, M. R., 2003, Evidence for low sulphate and anoxia in a mid-Proterozoic marine basin: *Nature*, v. 423, p. 632–635.
- Shively, J. M., Keulen, G., and Meijer, W. G., 1998, Something from almost nothing: Carbon dioxide fixation in chemoautotrophs: *Annual Review in Microbiology*, v. 52, p. 191–230.
- Silver, V. S., and Postgate, J. R., 1973, Evolution of symbiotic nitrogen fixation: *Journal of Theoretical Biology*, v. 40, p. 1–10.
- Sirevåg, R., 1974, Further studies on carbon dioxide fixation in *Chlorobium*: *Archives of Microbiology*, v. 98, p. 3–18.
- Sirevåg, R., Buchanan, B. B., Berry, J. A., and Troughton, J. H., 1977, Mechanisms of CO₂ fixation in bacterial photosynthesis studied by carbon isotope fractionation technique: *Archives of Microbiology*, v. 112, p. 35–38.
- Slesarev, A. I., Mezhevaya, K. V., Makarova, K. S., Polushin, N. N., Shcherbinina, O. V., Shakhova, V. V., Belova, G. I., Aravind, L., Natale, D. A., Rogozin, I. B., Tatusov, R. L., Wolf, Y. I., Stetter, K. O., Malykh, A. G., Koonin, E. V., and Kozyavkin, S. A., 2002, The complete genome of hyperthermophile *Methanopyrus kandleri* AV19 and monophyly of archaeal methanogens: *Proceedings of the National Academy of Sciences*, v. 99, p. 4644–4649.
- Smith, K. S., Jakubzick, C., Whittam, T. S., and Ferry, J. G., 1999, Carbonic anhydrase is an ancient enzyme widespread in prokaryotes: *Proceedings of the National Academy of Sciences*, v. 96, p. 15184–15189.
- Snel, B., Bork, P., and Huynen, M. A., 1999, Genome phylogeny based on gene content: *Nature Genetics*, v. 21, p. 108–110.
- Solomon, E. I., Sundaran, V. M., and Machonkin, T. E., 1996, Multi copper oxidases and oxygenases: *Chemical Reviews*, v. 96, p. 2563–2603.
- Strauss, G., and Fuchs, G., 1993, Enzymes of a novel autotrophic CO₂ fixation pathway in the phototrophic bacterium *Chloroflexus aurantiacus*, the 3-hydroxypropionate cycle: *European Journal of Biochemistry*, v. 215, p. 633–643.
- Stumm, W., and Morgan, J. J., 1996, *Aquatic Chemistry*: New York, John Wiley and Sons, Inc., 1022 p.
- Summons, R. E., Jahnke, L. L., and Roksandic, Z., 1994, Carbon isotopic fractionation in lipids from methanotrophic bacteria - relevance for interpretation of the geochemical record of biomarkers: *Geochimica et Cosmochimica Acta*, v. 58, p. 2853–2863.
- Summons, R. E., Jahnke, L. L., Hope, J. M., and Logan, G. A., 1999, 2-Methylhopanoids as biomarkers for cyanobacterial oxygenic photosynthesis: *Nature*, v. 400, p. 554–557.
- Swofford, D. L., 2002, *PAUP* Phylogenetic Analysis Using Parsimony (*and Other Methods)*: Sunderland, MA, Sinauer Associates.
- Taha, T. S. M., Brayman, T. G., Karplus, P. A., and Hausinger, R. P., 1997, Urease nickel metalocenter: Structure and assembly, in Winkelmann, G., and Carrano, C. J., editors, *Transition Metals in Microbial Metabolism*: Amsterdam, Harwood Academic Publishers, p. 391–414.
- Tatusov, R. L., Koonin, E. V., and Lipman, D. J., 1997, A genomic perspective on protein families: *Science*, v. 278, p. 631–637.

- Tekaia, F., Lazcano, A., and Dujon, B., 1999, The genomic tree as revealed from whole proteome comparisons: *Genome Research*, v. 9, p. 550–557.
- Tice, M. M., and Lowe, D. R., 2004, Photosynthetic microbial mats in the 3,416-Myr-old ocean: *Nature*, v. 431, p. 549–552.
- Turner, D., Whitfield, M., and Dickson, A. G., 1981, Inorganic speciation of trace metals: *Geochimica et Cosmochimica Acta*, v. 45, p. 855.
- Ueno, Y., Isozaki, Y., Yurimoto, H., and Maruyama, S., 2001, Carbon isotopic signatures of individual Archean microfossils(?) from Western Australia: *International Geology Review*, v. 43, p. 196–212.
- Ueno, Y., Yurimoto, H., Yoshioka, H., Komiya, T., and Maruyama, S., 2002, Ion microprobe analysis of graphite from ca. 3.8 Ga metasediments, Isua supracrustal belt, West Greenland: Relationship between metamorphism and carbon isotopic composition: *Geochimica et Cosmochimica Acta*, v. 66, p. 1257–1268.
- Valentine, D. L., Chidthaisong, A., Rice, A., Reeburgh, W. S., and Tyler, S. C., 2004, Carbon and hydrogen isotope fractionation by moderately thermophilic methanogens: *Geochimica et Cosmochimica Acta*, v. 68, p. 1571–1590.
- van der Meer, M. T., Schouten, S., de Leeuw, J. W., and Ward, D. M., 2000, Autotrophy of green non-sulphur bacteria in hot spring microbial mats: biological explanations for isotopically heavy organic carbon in the geological record: *Environmental Microbiology*, v. 2, p. 428–435.
- van der Meer, M. T. J., Schouten, S., Rijpstra, W. I. C., Fuchs, G., and Damste, J. S. S., 2001a, Stable carbon isotope fractionations of hyperthermophilic crenarchaeon *Metallosphaera sedula*: *FEMS Microbiology Letters*, v. 196, p. 67–70.
- van der Meer, M. T. J., Schouten, S., van Dongen, D. E., Rijpstra, W. I. C., Fuchs, G., Damste, J. S. S., de Leeuw, J. W., and Ward, D. M., 2001b, Biosynthetic controls on the C-13 contents of organic components in the photoautotrophic *Chloroflexus aurantiacus*: *Journal of Biological Chemistry*, v. 276, p. 10971–10976.
- Vorholt, J., Kunow, J., Stetter, K. O., and Thauer, R. K., 1995, Enzymes and coenzymes of the carbon monoxide dehydrogenase pathway for autotrophic CO₂ fixation in *Archaeoglobus lithotrophicus* and the lack of carbon monoxide in the heterotrophic *A. profundus*: *Archives of Microbiology*, v. 163, p. 112–118.
- Wächtershauser, G., 1990, Evolution of the first metabolic cycles: *Proceedings of the National Academy of Sciences*, v. 87, p. 200–204.
- Wahlund, T. M., and Tabita, F. R., 1997, The reductive tricarboxylic acid cycle of carbon dioxide assimilation: Initial studies and purification of ATP-citrate lyase from the green sulfur bacterium *Chlorobium tepidum*: *Journal of Bacteriology*, v. 179, p. 4859–4867.
- Waseda, A., and Didyk, B. M., 1995, Isotopic compositions of gases in sediments from the Chile continental margin, in Lewis, S. D., Behrmann, J. H., Musgrave, R. J., Candy, S. C., and others, editors, *Proceedings of the Ocean Drilling Program, Scientific Results 141*: College Station, Texas, Ocean Drilling Program, p. 307–312.
- Watson, G. M. F., Yu, J. P., and Tabita, F. R., 1999, Unusual ribulose 1,5-biphosphate carboxylase/oxygenase of anoxic Archaea: *Journal of Bacteriology*, v. 181, p. 1569–1575.
- Whiticar, M. J., 1999, Carbon and hydrogen isotopic systematics of bacterial formation and oxidation of methane: *Chemical Geology*, v. 161, p. 291–314.
- Wilde, S. A., Valley, J. W., Peck, W. H., and Graham, C. M., 2001, Evidence from detrital zircons for the existence of continental crust and oceans on the Earth 4.4 Gyr ago: *Nature*, v. 409, p. 175–178.
- Windhovel, U., and Bowien, B., 1990, On the operon structure of the *cfx* gene clusters in *Alcaligenes eutrophus*: *Archives of Microbiology*, v. 154, p. 85–91.
- Winkelmann, G., 1991, Specificity of iron transport in bacteria and fungi, in Winkelmann, editor, *Handbook of Microbial Iron Chelates*: Boca Raton, Florida, CRC Press, p. 366.
- Wirsen, C. O., Sievert, S. M., Cavanaugh, C. M., Molyneux, S. J., Ahmad, A., Taylor, L. T., DeLong, E. F., and Taylor, C. D., 2002, Characterization of an autotrophic sulfide-oxidizing marine *Arcobacter* sp. that produces filamentous sulfur: *Applied and Environmental Microbiology*, v. 68, p. 316–325.
- Woese, C. R., 1987, Bacterial evolution: *Microbiological Reviews*, v. 51, p. 221–271.
- Wolf, Y. I., Rogozin, I. B., Grishin, N. V., Tatusov, R. L., and Koonin, E. V., 2001, Genome trees constructed using five different approaches suggest new bacterial clades: *BMC Evolutionary Biology*, v. 1, p. 8.
- Wong, W., Sackett, W. M., and Benedict, C. R., 1975, Isotope fractionation in photosynthetic bacteria during carbon dioxide assimilation: *Plant Physiology*, v. 55, p. 475–479.
- Yachandra, V. K., Sauer, K., and Klein, M. P., 1996, Manganese cluster in photosynthesis: Where plants oxidize water to dioxygen: *Chemical Review*, v. 96, p. 2927–2950.
- Yang, S., Doolittle, R. F., and Bourne, P. E., 2005, Phylogeny determined by protein domain content: *Proceedings of the National Academy of Sciences*, v. 102, p. 373–378.
- Young, J. P. W., 1992, Phylogenetic classification of nitrogen-fixing organisms, in Stacey, G., Burris, R. H., and Evans, H. J., editors, *Biological Nitrogen Fixation*: Routledge, Chapman and Hall, p. 43–86.