Hydrogen and bioenergetics in the Yellowstone geothermal ecosystem

John R. Spear*, Jeffrey J. Walker*, Thomas M. McCollom†, and Norman R. Pace*t

*Department of Molecular, Cellular, and Developmental Biology and †Laboratory for Atmospheric and Space Physics, University of Colorado, Boulder, CO 80309

Contributed by Norman R. Pace, December 21, 2004

The geochemical energy budgets for high-temperature microbial ecosystems such as occur at Yellowstone National Park have been unclear. To address the relative contributions of different geochemicals to the energy demands of these ecosystems, we draw together three lines of inference. We studied the phylogenetic compositions of high-temperature (>70°C) communities in Yellowstone hot springs with distinct chemistries, conducted parallel chemical analyses, and carried out thermodynamic modeling. Results of extensive molecular analyses, taken with previous results, show that most microbial biomass in these systems, as reflected by rRNA gene abundance, is comprised of organisms of the kinds that derive energy for primary productivity from the oxidation of molecular hydrogen, H2. The apparent dominance by H2-metabolizing organisms indicates that H2 is the main source of energy for primary production in the Yellowstone high-temperature ecosystem. Hydrogen concentrations in the hot springs were measured and found to range up to >300 nM, consistent with this hypothesis. Thermodynamic modeling with environmental concentrations of potential energy sources also is consistent with the proposed microaerophilic, hydrogen-based energy economy for this geothermal ecosystem, even in the presence of high concentrations of sulfide.

geothermal springs | phylogenetic study | primary productivity | Yellowstone National Park | hydrogen metabolism

Microbial communities associated with volcanic hot springs have attracted broad interest because of the unique thermophilic properties of the constituent organisms. However, little attention has been given to hot spring communities as whole microbial ecosystems. One fundamental consideration in understanding any ecosystem is the energy budget: the relative contributions of different energy sources that fuel primary productivity, the conversion of carbon dioxide into biomass. Most of Earth’s biomass is considered to be the product of photosynthesis. However, at temperatures higher than ~70°C, photosynthesis is not known to occur, but thermophilic microbial communities develop well beyond that temperature (1–5). Consequently, high-temperature primary productivity must derive from chemosynthesis based on the oxidation of reduced inorganic or organic sources. A variety of lithotrophic microorganisms (which use inorganic energy sources) and heterotrophic organisms (which use reduced carbon) have been cultured from hot spring communities (6–11). However, the relative contributions of different potential energy sources to particular communities have not been systematically addressed.

Previous chemical analyses of Yellowstone hot springs have not provided satisfactory explanations of the energy sources that fuel the communities. Potential energy sources detected in different hot springs included sulfide, CH4, and other short-chain hydrocarbons, and reduced metals such as As(III), Fe(II), and Mn(II) (12, 13). However, none of these chemicals is ubiquitous in the hot springs, and robust microbial communities occur in some hot springs with little or none of these potential energy sources.

We propose that one way to gain insight into the relative contributions of potential energy sources available to microbial habitats is to assess the relative abundances of organisms that make up the communities. Microorganisms that engage in primary productivity are expected to be conspicuous in an autotrophic system. If the relative abundances of particular physiological types of organisms are taken to reflect the relative amounts of different energy sources that are drawn on for primary productivity, then a census of the physiologies that comprise a microbial community would correspond to a biological assessment of the energy demands of the particular ecosystem. The phenotypes of different microbes often are revealed by their phylogenetic associations, so a phylogenetic survey of the organisms that comprise a microbial community is expected to yield information on the community bioenergetics. Such a survey cannot be achieved with traditional culture-based methods, because most naturally occurring microbes resist cultivation with standard techniques (14). With the advent of molecular methods for the phylogenetic identification of organisms without the requirement for culture, the relative abundances of microbial community constituents can be estimated (15).

In the most commonly used culture-independent analysis of microbial community composition, small subunit rRNA genes are amplified by PCR from natural community DNA and then cloned and sequenced for phylogenetic identifications. The collection of rRNA gene sequences is a census of the phylogenetic types of organisms that comprise the community. If the organisms indicated by the sequences fall into relatedness groups with predictable forms of energy metabolism, based on cultured representatives, then the probable energy sources for the environmental organisms can be inferred. Microbial communities associated with high-temperature hot springs in Yellowstone National Park and elsewhere have been analyzed to some extent using these culture-independent methods (10, 11, 16–22). One finding of all studies has been the abundant occurrence of microorganisms from the Aquificales bacterial phylogenetic division (8, 10, 18, 20, 21, 23). All known representatives of Aquificales exclusively or preferentially use molecular hydrogen, H2, as an energy source. This dominance by Aquificales members suggested that H2 could be a main energy source in these hot spring ecosystems. However, the occurrence of H2 in Yellowstone hot springs had not been documented, and the few communities previously analyzed were from settings with limited variation in chemical composition.

To explore further the bioenergetics that underpin Yellowstone hot spring communities, we conducted extensive additional characterizations of microbial communities that thrive at >70°C in different chemical regimes. In parallel, we determined the chemical compositions of the hot springs, including the first systematic measurements of aqueous molecular hydrogen in the Yellowstone geothermal system. We then used thermodynamic modeling based on the hot spring chemistries to evaluate the

Data deposition: Nucleotide sequences have been deposited in the GenBank database (accession nos. AY861719–AY862082).

†To whom correspondence should be addressed. E-mail: nrpace@colorado.edu.


© 2005 by The National Academy of Sciences of the USA
bioenergetics of the available fuels. The results collectively provide new perspective on the energetics of high-temperature ecosystems.

**Materials and Methods**

**Molecular Phylogenetic Analyses.** Microbial communities associated with hot springs with high H\textsubscript{2} concentrations were surveyed with molecular phylogenetic methods previously described (19). Sediment samples were collected from Washburn Spring 1 (WB1), Washburn Spring 3 (WB3), and Cinder Pool (CPC) and frozen immediately on liquid nitrogen. In springs with little or no sediment, we collected and froze biomass that colonized glass growth slides placed in hot springs for periods of time from 48 h [Obsidian Pool Prime (OPP) Mud Volcano region] to 2 months [West Thumb Pool (WTP), 44°25’26”, W 110°34’36”]. Community DNA was extracted from ~1 g of sample by bead beating (24), which yielded an average of ~18 μg of DNA per g of sample. For OPP and WB1, we used the UltraClean fecal DNA extraction kit (MoBio Laboratories, Carlsbad, CA); for Cinder Pool, we used the 10-g UltraClean Mega Soil DNA extraction kit (MoBio Laboratories).

PCR primers used in this study included 515F, 1391R, 27F, 1492R, 4Fa, 333Fa, 360Fe, 82Fe, and 1391R (26). PCRs were incubated at 94°C for 2 min followed by 29 cycles at 94°C for 30 sec, 55.5°C for 1 min, and 72°C for 1.5 min, followed by a single 72°C step for 12 min. Each 30-μl reaction contained 1× PCR buffer, 0.2 mM each dNTP, 2.5 mM MgCl\textsubscript{2}, ~0.2 μM each primer, 1 mg/ml BSA, 1 M betaine, 0.5 units of Taq polymerase, and ~200 ng of template. PCR products were gel-purified with the QIAquick Gel Extraction Kit (Qiagen) and cloned with the TOPO TA Cloning Kit for Sequencing (Invitrogen). Unique clones were identified by restriction fragment-length polymorphism (24) and sequenced on a MegaBACE 1000 DNA Sequencer (Amersham Pharmacia). Sequences were aligned and analyzed with the ARB software package (www.mikrobiologie.tu-muenchen.de).

Nucleotide sequences have been deposited in the GenBank database under accession nos. AY861719–AY862082. Clone sequences deposited are coded for site (CPC, OPP, WB1, and WB3). PCR primer pair used (A, 515F and 1391R; B, 8F and 1492R; C, 333Fa and 1391R; D, 4Fa and 1391R; E, 360Fe and 1391R; F, 82Fe and 1391R), and three-digit clone number (e.g., OPPD012 = Obsidian Pool Prime, archaeal primer pair 4Fa&1391R, clone number 12).

**Hydrogen and Water Chemistry.** We measured aqueous H\textsubscript{2} concentrations in Yellowstone waters [hot springs, streams, geo-thermal vents, and a well (27)] with a modified bubble-stripping method (28). Source waters were pumped with a 12-V portable peristaltic pump through insulated, H\textsubscript{2}-impermeable polyethylene tubing for 20 min at a flow rate of 200 ml/min, through a 250-ml glass bottle bubble stripping device (28). Twenty milliliters of atmospheric air was introduced into the water-filled bottle. The temperature of the bubble was measured with a thermistor. Bubbles were collected with an air-tight syringe and transferred to nitrogen-charged, H\textsubscript{2}-impermeable, glass septum vials and sent to Microceps (Pittsburgh) for analysis of H\textsubscript{2}, CH\textsubscript{4}, CO\textsubscript{2}, and light hydrocarbons on a RGA3 reduction gas analyzer (Trace Analytical, Newark, DE) (28).

To determine the actual H\textsubscript{2} concentration in hot springs, we adjusted the measured values to account for the solubility of H\textsubscript{2} at high temperatures with Henry's law: \( C_{H_2} = C_{H_2}^o \frac{P_{H_2}}{R T} \), where \( C_{H_2} \) is the concentration of the gas in the water, \( C_{H_2}^o \) is the concentration (partial pressure) of the gas in the bubble, and \( H_2 \) is Henry's constant for that gas (the solubility of a gas at a given temperature). The temperature effect on Henry's constant is relatively small for most gases. However, for hydrogen, Henry's constant decreases by 28% from 0 to 100°C. Values of \( H_2 \) at high temperature were estimated with Ostwald's expression: \( H_2 = P(H_2)/RT \), where \( P(H_2) \) is 1.22 × 10^{-5} atm (1 atm = 101.3 kPa), \( R \) is the gas constant (0.0821 liters atm mol^{-1} K^{-1}), \( T \) is the measured bubble temperature (K), and \( C_{H_2}^o \) is 1 × 10^{-8} mol/liter.

Sulfide measurements were conducted with a colorimetric assay (CHEMetrics, Calverton, VA). Samples for water chemistry were filtered by syringe through a 0.2-μm filter, acidified with ultrapure nitric acid, and stored at 4°C until analysis. Anions, cations, and elemental analyses were conducted on a Series 4500 ICP (Dionex, Sunnyvale, CA), an ARL 3410+ ICP-AES (Thermo Electron), and a Varian ICP-MS (Varian, Palo Alto, CA).

**Thermodynamic Modeling.** The amounts of chemical energy available from lithoautotrophic reactions were quantified with the Gibbs free energy equation: \( \Delta G = \Delta G^o + RT \ln Q \), where \( \Delta G \) is the change in free energy of the reaction, \( \Delta G^o \) is the standard Gibbs free energy, and \( Q \) is the activity quotient of compounds involved in the reaction. Values of \( \Delta G^o \) were calculated with the computer program SUPCRT92 and thermodynamic data contained therein (29). Values of \( Q \) were determined with the measured composition of hot spring fluids. Because these are dilute solutions, activity coefficients were assumed to equal one for all dissolved compounds. Distributions of dissolved CO\textsubscript{2} and sulfide were calculated from the measured concentrations of these compounds, appropriate dissociation constants, and the measured pH, assuming the species were in equilibrium.

**Results**

**Chemistry of Yellowstone Hot Springs.** To provide a chemical context for interpretation of the results of microbiological studies, we conducted chemical analyses of selected hot springs in geologically distinct areas of Yellowstone with evidence of significant biomass (Fig. 1). The results of these analyses and other available data are presented in Tables 1 and 2. Collectively, the sites analyzed are representative of geothermal springs worldwide. Hot springs in Upper Geyser Basin, for instance, contain little sulfide and tend toward alkalinity (pH 8–9), with high concentrations of silica. Hot springs in Norris Geyser Basin and the Mud Volcano area contain relatively high concentrations of sulfides and low-to-neutral pH.

Concentrations of potential energy sources other than H\textsubscript{2},
such as sulfide and reduced metals, are highly variable in different hot springs (12, 13). Particularly notable, however, is our finding of ubiquitous H₂ at concentrations appropriate for energy metabolism, >5–10 nM (30–33). H₂ concentrations ranged to > 300 nM and were spring-dependent but seasonally constant in three springs tested (Queens Laundry, Octopus Spring, and Cinder Pool) (see Tables 1 and 2). Other potential energy sources, such as Fe[II], Mn[II], and NH₄, occur variably (Table 2 and refs. 12 and 13). However, the energy yield from microbial oxidation of such compounds is low relative to other (Tables 1 and 2). To determine the composition of microbial communities associated with these chemical settings, we amplified, cloned, and sequenced rRNA genes from crust and sediment communities as well as pioneer communities scraped from glass slides incubated in hot springs. Overall, ~2,500 randomly chosen rRNA gene clones were surveyed by restriction fragment-length polymorphism, and ~400 new sequences were determined and submitted to the GenBank database.

To determine the phylogenetic types of organisms present, we compared the sequences to sequences of known organisms in public databases. We also compared the compositions of the communities. Although the detail of compositions varied, all of the communities contained sequences representative of the same phylogenetic groups. Samples obtained on artificial growth

<table>
<thead>
<tr>
<th>Location</th>
<th>Site</th>
<th>Temp., °C</th>
<th>pH, Eh, V</th>
<th>D.O., mg/liter</th>
<th>SSU* source</th>
<th>Sulfide, µM</th>
<th>Sulfate, µM</th>
<th>H₂, nM</th>
<th>CH₄, µM</th>
<th>CO₂, mM</th>
<th>C₃H₈, ng/liter</th>
</tr>
</thead>
<tbody>
<tr>
<td>2X Distilled Water (Control)†</td>
<td>—</td>
<td>23</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0</td>
<td>0</td>
<td>2.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Dragon Pool, (Norris Basin)</td>
<td>1</td>
<td>72</td>
<td>3.1</td>
<td>5.8</td>
<td>ND</td>
<td>2.9</td>
<td>1.1</td>
<td>22.5</td>
<td>2.0</td>
<td>164</td>
<td></td>
</tr>
<tr>
<td>Well V-7, Biscuit Basin</td>
<td>2</td>
<td>57</td>
<td>7</td>
<td>27</td>
<td>ND</td>
<td>3.8</td>
<td>8.6</td>
<td>2.7</td>
<td>0.36</td>
<td>673</td>
<td></td>
</tr>
<tr>
<td>Yellowstone Lake</td>
<td>4</td>
<td>7</td>
<td>7</td>
<td>Several</td>
<td>0</td>
<td>4.3</td>
<td>0.1</td>
<td>0.1</td>
<td>0.00</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Canary Spring, Mammoth</td>
<td>5</td>
<td>68</td>
<td>8.5</td>
<td>ND</td>
<td>ND</td>
<td>11.0</td>
<td>3.8</td>
<td>8.0</td>
<td>0.39</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Octopus Spring</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall 2000</td>
<td>6</td>
<td>92</td>
<td>8.5</td>
<td>0.018</td>
<td>0.92</td>
<td>&lt;0.47</td>
<td>0.22</td>
<td>15.0</td>
<td>±0.28</td>
<td>±0.15</td>
<td>±0.08</td>
</tr>
<tr>
<td>Summer 2001</td>
<td>92</td>
<td>8.5</td>
<td></td>
<td></td>
<td></td>
<td>14.0</td>
<td>1.7</td>
<td>1.5</td>
<td>±0.08</td>
<td>±0.15</td>
<td>±0.08</td>
</tr>
<tr>
<td>West Thumb Pool</td>
<td>7</td>
<td>89</td>
<td>7.3</td>
<td></td>
<td>This study</td>
<td>0</td>
<td>0.25</td>
<td>15.5</td>
<td>±0.00</td>
<td>±0.41</td>
<td>±0.27</td>
</tr>
<tr>
<td>Washburn Spring #3</td>
<td>8</td>
<td>86</td>
<td>6.2</td>
<td>0.223</td>
<td>ND</td>
<td>167</td>
<td>44</td>
<td>18.5</td>
<td>±0.60</td>
<td>±0.30</td>
<td>±0.54</td>
</tr>
<tr>
<td>Queen’s Laundry</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall 2000</td>
<td>9</td>
<td>89</td>
<td>8</td>
<td>18</td>
<td>This study</td>
<td>2.2</td>
<td>0.042</td>
<td>28.0</td>
<td>±0.59</td>
<td>±0.73</td>
<td>±0.03</td>
</tr>
<tr>
<td>Summer 2001</td>
<td>89</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td>30.4</td>
<td>0.40</td>
<td>±0.95</td>
<td>±0.04</td>
<td>±0.27</td>
<td>±0.04</td>
</tr>
<tr>
<td>Cinder Pool</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall 2000</td>
<td>11</td>
<td>88</td>
<td>4.2</td>
<td>0.022</td>
<td>0.5</td>
<td>This study</td>
<td>47</td>
<td>1.0</td>
<td>±27.8</td>
<td>±0.21</td>
<td>±16.6</td>
</tr>
<tr>
<td>Summer 2001</td>
<td>88</td>
<td>4.3</td>
<td></td>
<td></td>
<td></td>
<td>13.7</td>
<td>±0.23</td>
<td>1.85</td>
<td>±0.18</td>
<td>±24.5</td>
<td>±3.18</td>
</tr>
<tr>
<td>Washburn Spring 1</td>
<td>12</td>
<td>76</td>
<td>6.7</td>
<td>0.067</td>
<td>0.3</td>
<td>This study</td>
<td>235</td>
<td>32.5</td>
<td>±103.1</td>
<td>±0.36</td>
<td>±17.2</td>
</tr>
<tr>
<td>Obsidian Pool</td>
<td>13</td>
<td>80</td>
<td>6.5</td>
<td>19</td>
<td>17.6</td>
<td>0.33</td>
<td>133.2</td>
<td>5.80</td>
<td>±0.14</td>
<td>±0.36</td>
<td>±0.04</td>
</tr>
<tr>
<td>Obsidian Pool Prime</td>
<td>14</td>
<td>74</td>
<td>5.7</td>
<td>This study</td>
<td>17.6</td>
<td>0.52</td>
<td>325.3</td>
<td>40.0</td>
<td>±0.1</td>
<td>±12.8</td>
<td>±0.13</td>
</tr>
</tbody>
</table>

D.O., dissolved oxygen; Temp., temperature; ND, not determined; C₃H₈, totaled ethane, ethene, propane, propene, n-butane, isobutene.

aSmall subunit rRNA gene analysis.

bWe take these values as baseline for assays.

cAll values are from ref. 13.

dAll values are from ref. 12.

eData are from ref. 53.

Table 1. Hydrogen in Yellowstone National Park hot springs

<table>
<thead>
<tr>
<th>Location</th>
<th>Site</th>
<th>Temp., °C</th>
<th>pH, Eh, V</th>
<th>D.O., mg/liter</th>
<th>SSU* source</th>
<th>Sulfide, µM</th>
<th>Sulfate, µM</th>
<th>H₂, nM</th>
<th>CH₄, µM</th>
<th>CO₂, mM</th>
<th>C₃H₈, ng/liter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Octopus Spring</td>
<td>6</td>
<td>512</td>
<td>1,380</td>
<td>262</td>
<td>3,420</td>
<td>3.4</td>
<td>0.0014</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>West Thumb Pool</td>
<td>7</td>
<td>66</td>
<td>1,265</td>
<td>153</td>
<td>1,384</td>
<td>13.2</td>
<td>DL</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Washburn Spring 1</td>
<td>8</td>
<td>68,000</td>
<td>&lt;1</td>
<td>6.7</td>
<td>&lt;70</td>
<td>340</td>
<td>65</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Queen’s Laundry</td>
<td>9</td>
<td>282</td>
<td>1,313</td>
<td>239</td>
<td>1,996</td>
<td>0.87</td>
<td>DL</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cinder Pool</td>
<td>11</td>
<td>1,130</td>
<td>2,400</td>
<td>601</td>
<td>4,700</td>
<td>&lt;6</td>
<td>0.088</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Washburn Spring 1</td>
<td>12</td>
<td>34,000</td>
<td>&lt;1</td>
<td>10</td>
<td>50</td>
<td>510</td>
<td>23.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obsidian Pool</td>
<td>13</td>
<td>349</td>
<td>DL</td>
<td>25</td>
<td>199</td>
<td>427</td>
<td>0.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obsidian Pool Prime</td>
<td>14</td>
<td>206</td>
<td>526</td>
<td>305</td>
<td>1,171</td>
<td>50</td>
<td>0.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
surfaces generally overlapped with the environmental sediment samples. Fig. 2 summarizes the census results. The phylogenetic distribution of rRNA genes amplified with the universal PCR primers (Fig. 2A) provides some perspective on the overall microbial composition of the Yellowstone geothermal ecosystem. Communities were dominated by bacterial rRNA genes. Archaea are considered common in geothermal and other “extreme” environments, but these and all previous surveys indicate that such organisms are not more abundant than bacteria (19). Most of the archaeal sequences encountered were related to environmental crenarchaeote sequences previously observed in Obsidian Pool (16, 17), none with a specific relationship to a cultured organism (Fig. 2B).

Fig. 2 shows the main phylogenetic groups identified in springs with five different chemical compositions. Although several hundred unique bacterial sequences were determined, these fell into only a few phylogenetic groups. Sequences representative of *Aquificales* were most abundant in the communities, and sequences representative of *Thermotogales*, *Thermus/Deinococcus*, and *Thermodesulfobacteria* also were common.

Collectively, ≈90% of sequences obtained were representatives of these phylogenetic groups. These results are consistent with earlier findings from more limited studies (16–20). Most *Aquificales* sequences were closely related to cultured organisms that rely on H₂ as an energy source, including *Thermocrinis ruber* (8), *Hydrogenobacter* spp. (34), *Hydrogenobaculum* spp. (35), and *Hydrogenothermus* spp. (36). Because representatives of a relatedness group are expected to have properties that are uniformly present in known members of the group, the environmental hot spring organisms represented by the dominant sequences are predicted to engage in hydrogen oxidation.

Comparison of community compositions in low- and high-sulfide samples (Fig. 4) indicates that organisms recognized for sulfur oxidation, such as relatives of *Thiobacillus* spp., do not dominate. Instead, β-proteobacterial sequences emerge in communities with higher sulfide concentrations. Many of these sequences are specifically related to β-proteobacteria known for sulfate reduction and commonly use H₂ as a reductant. Sulfate is often present in the hot springs (Table 1). Our results suggest that, when sulfate is present, sulfate-reducing bacteria can contribute significantly to the energy budget of the community.

**Thermodynamic Modeling.** We modeled the potential energy available to the microbial communities in hot springs. Because photosynthesis does not occur above about 70°C, most microbes in
Yellowstone hot spring environments must obtain their energy from reduced compounds in geothermal fluids. Dissolved sulfides, CH$_4$, and H$_2$ are the principal potential energy sources available to these communities. The relative potential energy yields of available energy sources can be estimated from the measured chemical compositions of the hot springs. This potential energy depends on available O$_2$, which is difficult to measure but is low because of the reduced nature of hot springs and the poor solubility of O$_2$ in hot water (12, 13). Therefore, we modeled potential energy available in four springs for oxygen-consuming lithoautotrophic reactions over a range of O$_2$ concentrations (Fig. 5) (37). Results show that H$_2$ oxidation was favored under oxygen-limited conditions observed in hot springs, which is consistent with the apparent dominance of putative hydrogen metabolizers in this study. The dominant abundance of predicted hydrogen-oxidizing organisms occurs in hot springs with high sulfide concentrations (e.g., Washburn Spring 1, site 12; see Fig. 1) and low sulfide concentrations (e.g., Obsidian Pool Prime, site 14, and West Thumb Pool, site 7).

**Discussion**

Microbiology historically has focused on single organisms, with limited attention to microbial ecosystems. Indeed, it is a challenge even to identify a microbial “ecosystem,” in the sense of an ecological unit. Microbial ecosystems are constrained not by geography or climate, but rather by local chemical and physical conditions. Innumerable microbial ecosystems collectively underpin and mold our biosphere, so it is important to strive to understand their biochemical webs.

In this study we considered a relatively simple and confined ecosystem setting, Yellowstone geothermal springs $>70^\circ$C, and we posed a simple question: What is the main source of metabolic energy that drives such communities? We draw together three lines of inference to propose that the main energy source for these communities is H$_2$, O$_2$ and, to a lesser extent, oxidized sulfur species serve as the main terminal electron acceptors. Sulfide, long considered an important energy source for hot spring communities, seems to play a minor role. Microbial sulfide oxidation may play a more prominent role further away from hot spring sources, where cooler waters allow higher O$_2$ solubility.

The phylogenetic composition of these communities is the first line of inference that leads us to the conclusion that hydrogen is their main energy source. These and previous molecular analyses of hot springs with varied chemistries show that the dominant rRNA genes are derived from close relatives of species known for hydrogen metabolism. The second line of inference, which corroborates the molecular results, is the finding of ubiquitous H$_2$ in Yellowstone hot springs, at concentrations sufficient for microbial bioenergetics. Finally, thermodynamic calculations based on O$_2$ limitation show that H$_2$ metabolism is favored in this ecosystem.

Our conclusions are based on several assumptions about the molecular approach to microbial community analysis. In principal, rRNA clone libraries provide a snapshot of the relative proportions of phylogenetic types in a community, and some properties of those individuals can be inferred from phylogenetic information. Representatives of a relatedness group are expected to have properties that are common to the group (38). However, we acknowledge that potential experimental artifacts can bias how well clone libraries represent the actual proportions of phylogenetic types in a sample (reviewed in refs. 39 and 40). Such artifacts include variable PCR amplification due to primer selectivity and differential extraction of genomic DNA from samples. Nonetheless, comparisons of results obtained from clone libraries and other methods, such as fluorescent in situ hybridization (41) and rRNA hybridization (42), show that careful application of molecular methods accurately identifies the abundant organisms in a sample.

We endeavored to minimize potential experimental artifacts by analysis of clone libraries prepared using different suites of PCR primers with broad specificities, an approach used successfully in other studies (43). Obsidian Pool Prime, for example, was

![Fig. 4. Low- and high-sulfide communities compilation. (A) The phylogenetic distribution of rRNA gene sequences obtained from the two low-sulfide springs of this study (West Thumb Pool and Obsidian Pool Prime) combined with the five low-sulfide springs studied by Blank et al. in ref. 18 (Octopus Spring, Queens Laundry, Eclipse Geyser, Spindle Spring, and Boulder Spring). (B) The phylogenetic distribution of rRNA gene sequences obtained from the three high-sulfide springs in this study (Cinder Pool, Washburn Spring 1, and Washburn Spring 3).](image)

![Fig. 5. Results of thermodynamic models. Shown are the amounts of energy available from O$_2$-consuming metabolic reactions, expressed in terms of available energy per mole of limiting O$_2$ for comparative purposes. The available energy is shown for a range of hypothetical O$_2$ concentrations because accurate O$_2$ concentrations in hot waters are difficult to assess.](image)
examined with eight different PCR primer pairs. Although different libraries always contained some unique sequences not seen in other libraries from the same environmental DNA, we see no systematic difference in the proportions of phylogenetic groups in the libraries. Regardless of potential biases, >93% of rRNA sequences characteristic of H$_2$-oxidizing microbes dominate both low- and high-sulfide springs. This finding provides strong evidence that such organisms constitute the main component of these communities. Also, each of the communities is probably more complex than we detect. At most, we analyzed several hundred randomly chosen clones, and in no case did we exhaust the diversity in a library. Thus, our analysis captures only the most abundant rRNA genes. Substantial diversity remains to be uncovered in these and other geothermal systems.

The Yellowstone hot spring communities are relatively simple from the perspective of rRNA gene sequences. However, microbes that have even identical rRNA sequences may not be entirely identical. Because hot spring geochemistry varies, organisms are expected to evolve adaptations to local conditions. For instance, organisms in settings with little reduced iron (e.g., Octopus Spring) may have mechanisms for the acquisition or utilization of iron that are not required in high-iron hot springs (e.g., Obsidian Pool). Previous studies have shown genetic variation among microbes with identical rRNA sequences from different hot springs, indicated by variation in sequences of rRNA internal transcribed spacers (18, 44).

The importance of H$_2$-metabolizing organisms in environmental microbiology has long been recognized (1). The nM concentrations of H$_2$ reported here are consistent with those reported for other oxygen-poor environments such as lake sediment (36 nM), rice paddies (28 nM), and sewage sludge (203 nM) (30). Hydrogen in Yellowstone geothermal waters is likely geochemical in origin. Sources of geochemical H$_2$ are not well understood in general (45), but in the Yellowstone environment they probably derive from subsurface interaction of water with Fe$_3$O$_4$ (46–51). Life in the subsurface probably is limited more by the availability of oxidant than of fuels such as H$_2$ (52). This theme of hydrogen as a main fuel in Yellowstone hot springs likely resonates to other geohydrothermal ecosystems, where H$_2$ probably is common in anoxic water.

This article is dedicated to the memory of Tommy Gold, a pioneer in thought on carbon and energy sources in the Earth’s crust. We gratefully acknowledge the help and assistance of the Yellowstone Center for Resources. Dr. Kirk Nordstrom (U.S. Geological Survey, Boulder, CO) provided useful insight and feedback throughout the life of the project. We thank members of the N.R.P. laboratory for collegiality, review, and thoughtful comments on the manuscript. Funds for this work were provided by the National Science Foundation’s Astrobiology grant DEB-9870880 (to N.R.P.) the National Aeronautics and Space Administration Astrobiology Institute (to N.R.P.), a National Science Foundation Microbial Biology Postdoctoral fellowship (to J.R.S.), and an Agouron Institute Postdoctoral Fellowship (to J.R.S.).
Hydrogen and energy flow as “sensed” by molecular genetics

Kenneth H. Nealson*
Department of Earth Sciences, University of Southern California, Los Angeles, CA 90089

Perhaps just a quick whiff of the rarified air in Yellowstone Park might lead one to the apparently wrong conclusion; namely, that this is a sulfur-driven ecosystem. With all of that smelly hydrogen sulfide emanating from the hot water, it is easy for a microbiologist to leap to the conclusion that sulfur must dominate this ecosystem. Perhaps even more so since the discovery of the deep sea hydrothermal vents and their remarkable symbiotically driven ecosystems (1, 2), we have become accustomed to the notion of microbial systems powered by sulfide and/or sulfur oxidation and expect them to be operating here. For this reason, it was somewhat surprising to read the conclusion of the Spear et al. (3) in a recent issue of PNAS, who reported that these smelly boiling pools were in fact running on hydrogen rather than sulfide metabolism. The implications of this work go beyond this apparently straightforward conclusion.

The ability to identify and classify microbes by using molecular genetics techniques (16S rRNA sequence analysis) ushered in a new era in microbiology, making it possible for the first time to assess “who’s there” even when the organisms could not be cultivated (4–6). Based on previously published results showing that members of the Aquificales were abundant (7–11), the authors suspected that hydrogen and not sulfur should be the important nutrient. This logic comes from the knowledge that cultivated members of the Aquificales use hydrogen either exclusively or preferentially. To test this hypothesis, Spear et al. (3) examined the populations in a variety of Yellowstone ponds, with levels of sulfides ranging from nondetectable to >200 μM. They also measured several key chemical variables, including sulfate, hydrogen, oxygen, pH, Eh (reducing potential), and methane. With these data in hand, it was possible to ask, using a thermodynamic modeling approach, whether it was feasible that hydrogen was a major energy source in each environment.

This study showed very nicely how a combination of approaches can lead one to an explanation that is compatible with all of the data and yet in contrast to what might have been expected. Thus, the impact of the work may lie as much as in the implementation of the multifaceted approach as in the particular nature of the conclusions reached: Molecular genetics, environmental geochemistry, and geochemical modeling are brought together to begin to unravel the workings of the ecosystem. Although none of these three approaches on its own might have been sufficient to lead to this inference, taken together, they make a good case for these boiling sulfurous ponds being examples of hydrogen-driven ecosystems. In essence, such evidence can be used to move us from the “who’s there?” era to the “what’s going on?” era—a move of great importance to biologists and geochemists alike.

The boiling pools in Yellowstone Park are running on hydrogen rather than sulfide metabolism.

But does it really matter whether hydrogen or sulfur metabolism dominates these boiling ecosystems? Yes! There is little doubt that one of the most abundant energy sources on almost any geologically active body is expected to be molecular hydrogen. Hydrogen is not only the most abundant element in the universe, it is also sequestered in many geological reservoirs from which can be released as a function of either magmatic degassing, as in so-called magma-hosted systems, and/or the exothermic reaction of highly reduced magmatic rocks (peridotites) releasing high-pH fluids containing H₂ and CH₄ but containing much less CO₂ because of the high pH (11 or higher; Fig. 1). In the latter case, if the water–rock interaction takes place at high temperature and pressure, the product can be primarily methane, with organic carbon produced as well (16). Although there are many variations on these themes (i.e., based on the water source interacting with the reduced rocks), both are capable of producing hydrogen for subsequent powering of subsurface ecosystems.

To this end, in 1995, Stevens and McKinley (17) proposed the existence of a hydrogen-driven microbial community in the deep subsurface of the Columbia River Basin, a claim that was hotly contested by Anderson et al. (18) based on the expected rates of hydrogen production. This latter group then proposed a hydrogen-powered ecosystem of their own (19): a subsurface Archaea-dominated microbial community in the groundwater system beneath the Lidy Hot Springs in Yellowstone National Park.

See companion article on page 2555 in issue 7 of volume 102.

*E-mail: knealson@usc.edu.

© 2005 by The National Academy of Sciences of the USA
Idaho. Although molecular methods were employed to detect the presence of methanogens as members of the microbial community, there was no molecular probing to indicate abundance, no isotopic analysis of the methane to indicate a biological origin, and no suggested mechanism for how nanomolar levels of hydrogen could be producing millimolar levels of methane. Herein lies an example of where more molecular genetic and geochemical data, along with some modeling of the type done by Spear et al. (3), would have been very helpful. Subsequently, several other workers have proposed hydrogen-driven ecosystems in the subsurface vent sediments found near mid-oceanic ridge (MOR) environments: communities hypothesized on the basis of observations of abundant biological material being identified in vent fluids or associated with chimney material of black smokers (20–28). More recently, a hydrogen-driven community was hypothesized to exist in the subvent region in the Central Indian Ridge (13). Of interest here is that many of the same approaches used by Spear et al. (3) were employed by this group: phylogenetic profiling and geochemical measurements, all of which were consistent with the presence of a methane generating community dominated by hydrogen-utilizing Methanococcales as primary producers and Thermococcus as fermenters.

What is seen from the above discussion of the Spear et al. article (3), is that with a good dose of molecular phylogeny, sufficient knowledge of cultivated organisms, a bit of geochemistry, and some clever modeling, it is possible to make strong inferences with regard to how energy flows through microbial ecosystems. Such inferences can and will be tested by using even more tools and approaches, such as actual flux measurements, and stable isotopic fractionation patterns. Will we one day be able to infer processes from molecular data alone? Will we be able to look into past processes using these methods? Will we be able to unambiguously fingerprint energy processes in extreme environments, and identify the role of hydrogen (or other energy sources) in the present and past Earth? This reviewer remains the optimist: It is easy to imagine that as we learn how to read the Rosetta stone of genomic information, it will reveal all of these things and more.